基于渐进式混合对比学习的无监督领域自适应行人再识别

赵 宇[1],舒巧媛[2]
(1. 重庆第二师范学院人工智能学院,重庆 400065;2. 重庆第二师范学院数学与大数据学院,重庆 400065)

摘 要: 无监督领域自适应(Unsupervised Domain Adaptation,UDA)行人再识别(person Re-IDentification,Re-ID)旨在利用有标注的源域数据来解决无标注目标域数据的无监督 Re-ID 任务. 近期,对比学习在该领域引起关注,但现有方法存在正样本对差异较小以及忽略负代理采样偏差的问题. 为解决这些问题,本文提出一种渐进式混合对比学习(Progressive Hybrid Contrastive Learning,PHCL)方法. 在每个训练轮次,PHCL方法通过聚类和渐进细化两个步骤,将无标签数据集划分为带伪标签的聚类样本和未聚类的独立实例. 基于聚类划分结果,PHCL方法在两个层次实施对比学习:通过将同一聚类(目标域)或同一身份标签(源域)中的相似样本拉近,指导模型学习类内相似性,同时通过在未聚类的实例间施加排斥作用,挖掘实例间差异性. 此外,PHCL方法通过最近邻挖掘为未聚类的实例生成正代理,增大正样本对的差异性,学习更丰富的语义信息. 同时,PHCL方法在负代理采样过程中去偏差,减轻假负代理对训练的不利影响. 实验结果表明:PHCL方法在Market-1501和MSMT17数据集上的平均精度均值(mean Average Precision,mAP)分别为85.9%与42.3%,比基线模型分别提高4.3个百分点和13.5个百分点. 上述实验结果验证了PHCL方法在UDA Re-ID任务中的有效性.

关键词: 无监督领域自适应(UDA)行人再识别(Re-ID);对比学习;伪标签;最近邻挖掘;去偏差

基金项目: 重庆市自然科学基金面上项目(No.CSTB2022NSCQ-MSX0404);重庆市教育委员会科学技术研究项目(No.KJQN202401624)

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2025)06-1829-18

电子学报 URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250110

第二十七届中国科协年会学术论文

Unsupervised Domain Adaptive Person Re-Identification Based on Progressive Hybrid Contrastive Learning

ZHAO Yu[1], SHU Qiao-yuan[2]
(1. *School of Artificial Intelligence*, *Chongqing University of Education*, *Chongqing* 400065, *China*;
2. *School of Mathematics and Big Data*, *Chongqing University of Education*, *Chongqing* 400065, *China*)

Abstract: Unsupervised domain adaptive (UDA) person re-identification (Re-ID) seeks to leverage labeled source domain data to address the task of unsupervised Re-ID in unlabeled target domain data. Recently, contrastive learning has attracted attention in this field. However, current methods suffer from small differences in positive sample pairs and overlook biases in negative proxy sampling. To resolve these challenges, this paper presents a progressive hybrid contrastive learning (PHCL) method. In each training epoch, the PHCL method divides the unlabeled dataset into clustered samples with pseudo-labels and un-clustered independent instances through two steps: clustering and progressive refinement. Based on the clustering results, PHCL implements contrastive learning at two different levels: to learn intra-category similarity through bringing together similar samples within the same cluster (target domain) or identity label (source domain) and explores inter-instance discrimination by applying repulsion among un-clustered individual instances. Moreover, the PHCL method generates positive proxies for anchor samples through nearest neighbor mining, increasing the differences among positive sample pairs to learn richer semantic information. Additionally, the PHCL method performs debiasing in the negative proxy sampling process, mitigating the adverse impact of false negative proxies on model training. Experimental results show that the PHCL method achieves mean average precision (mAP) of 85.9% and 42.3% on the Market-1501 and

收稿日期:2025-02-13;修回日期:2025-06-03;责任编辑:覃怀银

MSMT17 datasets, respectively, which are improvements of 4.3 percentage points and 13.5 percentage points over the baseline model. These results validate the efficacy of the PHCL method for UDA Re-ID.

Key words: unsupervised domain adaptive (UDA) person re-identification (Re-ID); contrastive learning; pseudo labels; nearest neighbor mining; debiasing

Foundation Item(s): General Program of Chongqing Natural Science Foundation (No.CSTB2022NSCQ-MSX0404); Science and Technology Research Program of Chongqing Municipal Education Commission (No.KJQN202401624)

1 引言

行人再识别(Re-IDentification,Re-ID)旨在基于特定查询样本,在非重叠摄像头间检索匹配的行人目标.它对于智能监控、智能安防等应用场景具有重要意义.在深度学习的蓬勃发展下[1~3],受人工标注信息的指导,监督模式下的Re-ID方法取得了令人瞩目的进展[4~7].然而,Re-ID领域仍然存在一些挑战.首先,由于人体姿态不匹配、背景杂乱和光照变化等原因,特定行人的外观在不同摄像头下可能发生剧烈变化.此外,对每一个新的监控场景获取大规模的人工标注数据成本高昂,这在实际应用中是不可行的.因此,近年来,无监督Re-ID引起越来越多的关注.

一些Re-ID方法基于纯无监督学习(Fully Unsupervised Learning,FUL).此类方法通常利用图像的内在特性来指导模型训练.例如,时空相似性学习[8,9]被用于挖掘摄像头内部轨迹的关联性和摄像头间轨迹的相关性.一些方法[10,11]利用聚类算法挖掘图像的相似性,从而生成伪标签作为监督信号.为了更好地表示样本之间的关系,一些研究人员放弃硬伪标签(hard pseudo label),使用软化相似性学习来挖掘软标签信息[12].然而,这些方法的性能容易受到不稳定的伪标签学习所产生的噪声指导信号的影响.

近年来,无监督领域自适应(Unsupervised Domain Adaptation,UDA)在计算机视觉领域取得显著进展.其核心目标是利用源域(有标注)的知识提升目标域(无标注)的模型性能,同时克服域间分布差异带来的挑战.早期的UDA工作主要依赖对抗性训练(adversarial training)[13]和最大均值差异(Maximum Mean Discrepancy,MMD)[14]等策略来进行特征对齐.然而,这类方法在开放环境下面临局部结构失配问题.随着自训练(self-training)策略的兴起,基于伪标签(pseudo label)的迭代优化方法成为主流.例如,文献[15]利用目标域数据的类内相似度来优化伪标签的生成,并使用改进后的伪标签对齐源域和目标域.然而,伪标签的可靠性高度依赖初始模型的预测质量.为此,文献[16]提出基于证据学习(evidential learning)的无监督域自适应框架,通过结合证据预测不确定性(Evidential Prediction Uncertainty,EPU)和证据邻接不确定性(Evidential Adjacency-consistent Uncertainty,EAU)实现细粒度知识聚合与可信语义传播.文献[17]设计了基于小波数据增广的Transformer模型与不确定性感知自训练机制,通过隐式数据增广和伪标签校正提升跨域超分辨率任务的性能.文献[18]通过构建中间域概率分布桥接源域与目标域,利用博弈论驱动的特征混合策略来解决域间差异问题,显著提升无监督域适应任务的性能.鉴于UDA方法的巨大潜力,一些研究人员也开始将UDA方法引入Re-ID领域,将监督信息从有标注的源域数据传递到无标注的目标域数据[19~21].然而,这些UDA Re-ID方法通常先在源域上进行预训练,再在目标域进行微调,未能充分结合源域数据的标注信息和目标域数据的内在特性.

近年来,对比学习(contrastive learning)在无监督视觉表示学习任务上取得了显著的进展[22,23].这些方法的主要目标是学习一个合适的特征空间,使得图像表示在同一样本的不同数据增广(data augmentation)视图上具有不变性.在对比学习框架中,给定锚样本(anchor sample)后,正代理的构建和负代理的采样对于模型性能至关重要.在正代理构建方面,文献[24~27]采用对同一输入数据进行两种不同的数据增广操作,生成两个数据增广副本,以构建正样本对.这类方法的正代理来源于锚样本的直接数据增广,因而正样本对之间的差异较小,可能限制模型对语义不变性的学习能力.在负代理采样方面,文献[25]使用一个训练批次中除了正样本对以外的样本作为负代理,文献[24,26,27]使用内存库中除正代理以外的所有特征缓存作为负代理.然而,这些方法在进行负代理采样时,有可能会将与锚样本相似的样本列为负代理,为模型训练带来偏差.此外,一些研究将对比学习引入到无监督Re-ID领域[28~30].其中,文献[28]通过在摄像头内和摄像头间进行聚类和对比学习,挖掘摄像头内和摄像头间的图像相似性.一些研究人员结合生成对抗网络和对比学习模块,用于学习合成图像的视角不变特征,并用聚类算法生成的伪标签指导对比学习[30].这些工作已构建多种用于无监督Re-ID的对比学习方法,取得了一定的性能提升.然而,这些方法仍然面临着不可忽视的问题:

(1)实例级对比学习的正样本对之间差异较小,难以学习到丰富的语义信息.

(2)忽略了负代理采样中的偏差.由于缺少标注信息,一些与锚样本具有相同身份的图像可能会被指定为负代理,这些假负的负代理采样是偏差的源头,将会

给对比学习过程带来负面影响.

基于以上分析,本文在构建跨域对比学习框架时作出如下改进:

(1)设计一种基于最近邻挖掘的正代理生成方法,以增强正代理与锚样本之间的差异性.

(2)将去偏差机制引入负代理采样中,以缓解假负代理对模型训练的负面影响.

本文提出一种用于UDA Re-ID的渐进式混合对比学习(Progressive Hybrid Contrastive Learning,PHCL)方法. 为了学习一个有判别性的特征空间,本文所提出的PHCL方法首先使用聚类算法和渐进细化策略将目标域数据集划分为两个子集:一个子集包含带有伪标签的聚类样本,而另一个子集由未聚类的独立实例组成. 然后,为确保大量样本相似度的计算能够在基于小批次(mini-batch)的训练过程中有效进行,PHCL方法建立一个样本级的内存库,用于存储所有训练样本特征的近似值,即代理. 基于源域数据集和目标域的两个子集,PHCL方法旨在同时应用两个层次的对比学习,挖掘行人样本的类内相似性和实例间判别性. 具体而言,PHCL方法不仅通过将同一聚类(目标域)或同一身份标签(源域)中的相似样本彼此拉近,以指导特征提取网络学习类内相似性(intra-category similarity),而且还通过在未聚类的独立实例间施加排斥作用来挖掘实例间差异性(inter-instance discrimination). 此外,一方面PHCL方法通过最近邻挖掘的方式,为未聚类的实例生成正代理,增大正样本对的差异性,使得模型能够真正拉近两个不同的相似样本,学习到更丰富的语义信息. 另一方面,PHCL方法还在负代理采样的过程中实现去偏差操作,从而减轻假负代理对模型训练的不利影响. 综上所述,本文的主要贡献如下:

(1)本文为UDA Re-ID任务引入两种不同的学习策略,即正代理最近邻挖掘和负代理的去偏差,从而增大正样本对之间的差异以及降低假负代理对模型训练的不利影响.

(2)本文将正样本最近邻挖掘和负代理去偏差整合到两个不同层次的对比学习中,这使得模型能够有效利用差异足够大的正样本对以及真负代理,来更好地学习类内相似性和实例间差异性.

(3)为了验证所提出方法的有效性,本文在两个广泛使用的Re-ID数据集上进行充分的实验和评估. 综合实验结果表明,本文所提出的方法在UDA Re-ID领域具有有效性.

2　相关工作

2.1　无监督Re-ID

现有的无监督Re-ID方法大致可以分为两类:FUL方法和UDA方法. FUL方法只使用未标记的数据训练模型,而UDA方法一般使用已标注的源域数据集和未标注的目标域数据集共同训练模型.

FUL方法的目标是通过探索其他监督信号,在没有任何人工标注的情况下训练一个深度神经网络. 文献[12]使用软标签来训练一个分类网络,从而更好地估计图像对之间的相似性. 文献[31]利用图卷积网络(Graph Convolutional Network,GCN)来估计成对图像的伪标签. 文献[32]提出将层次化聚类和三元组损失(triplet loss)函数集成到一个统一的框架,从而更加鲁棒地学习伪标签. 伪标签的质量依赖于输入图像的特征表示. 但是,在模型训练的早期,图像特征表示还不够好,不能生成高质量的伪标签,而低质量的伪标签会污染网络训练. 因此,在输入伪标签训练网络之前,需要设计一种聚类细化方法来提高聚类质量. 文献[33]为样本分配多个标签,并设计一个可以使用多标签作为监督信号的损失函数,以便进行模型训练.

UDA方法利用已标记的源域数据集来辅助模型训练. 一些方法旨在通过对齐特征分布来减小标记的源域数据集和未标记的目标数据集之间的域差异[34]. 此外,也有研究人员使用生成对抗网络,将图像风格从源域转移到目标域,从而训练Re-ID模型[35,36]. 文献[37]提出一种组合损失函数,该函数使用源域和目标域中的样本进行联合训练,使得模型迁移时能有效利用源域信息和目标域聚类异常样本信息. 文献[38]建议通过线性变换对齐两个域中分布的二阶统计量,以减少域偏移,从而进行模型参数的有效迁移.

2.2　对比学习

近年来,随着孪生网络(siamese network)的发展和应用,对比学习(contrastive learning)开始出现在无监督学习领域. 对比学习的目标是学习良好的图像表征,它通过对一组正样本和负样本进行数据扩充来学习具有不变性的特征表示. 目前已有的对比学习方法可以分为两类:基于实例的方法[24,25,39,40]和基于聚簇的方法[26,27,41~45].

基于实例的方法将每个图像视为单独的类别,并将同一实例的两个增广图像视为正样本对,将同一个训练批次或者全局内存库中的其他图像视为负样本. 文献[24]使用字典实现对比学习,将对比学习的一个分支转换为动量编码器(momentum encoder),缓解大批量训练带来的内存压力,提升对比学习的效率. 文献[25]则将当前批次中除了正样本对以外的样本视为负样本,以此来构建实例级的对比学习. 文献[39]提出一种基于实例的特征嵌入方法,利用实例级的自监督信号来指导模型学习数据增广下的特征不变性. 文献[40]提出一种非对称网络结构,通过在线网络(online

network)预测目标网络(target network)对同一图像不同增广视图的特征表示,并利用在线网络参数的移动平均来更新目标网络的参数,从而在对比学习过程中,移除对负样本的依赖. 基于实例的对比学习方法构造相对简单,利用数据增广和对比损失函数能够有效地进行无监督学习,可广泛应用于图像、文本等多种数据类型. 然而,此类方法通过对锚样本进行数据增广直接构造正代理,使得两者的差异较小,难以贡献丰富的语义信息. 此外,此类方法的负样本中可能包含与锚样本相似的实例,从而为训练过程带来偏差.

基于聚簇的方法将同一聚类中的样本视为正样本,将其他样本视为负样本. 文献[26]利用权值共享深度神经网络,从具有不同数据增广的样本对中提取特征,并对行空间和列空间中的特征进行对比聚类. 文献[27]提出一种基于聚类引导的非对称对比学习框架,利用对图像特征的聚类结果来指导对比学习过程,降低图像色彩变化带来的不利影响. 文献[41]采用多裁剪数据增广方法,增强网络的鲁棒性,并提出一种可扩展的在线聚类方法来探索聚类的互不变性. 文献[42]提出一种基于自适应滤波器的图聚类对比学习方法,该方法首先使用自适应滤波器挖掘超出低频成分的整体性信息,并将拓扑结构信息编码到特征中;然后,使用对比学习机制优化图结构,提升图的可辨识性. 文献[43]在图像的强增广样本和弱增广样本之间构建基于聚类的对比学习,从而提升学习到的图像特征的可判别性. 文献[44]通过高置信度聚类结果指导对比学习,构建权重不共享的孪生编码器生成图的双视图,从而提升图聚类的判别性与可靠性. 文献[45]提出一种聚类监督对比学习损失函数,用于更好地聚类正常特征,同时分离异常特征,从而达到量化驾驶员异常行为的目标. 基于聚簇的对比学习方法将锚样本对应的聚簇设置为正代理,有效解决了正样本对差异较小的问题. 然而,这类方法将其他聚簇作为负样本,高度依赖于图像特征聚类的准确性,依然无法有效应对负样本采样中的偏差问题.

3 本文方法

3.1 问题定义

给定一个拥有 N_t 张行人图像的无标注目标域数据集 $X^t=\{x_i\}(i=1,2,\cdots,N_t)$,一个拥有 N_s 张行人图像的有标注源域数据集 $X^s=\{x_i\}(i=1,2,\cdots,N_s)$,UDA Re-ID 的目标是使用 X^t 和 X^s 训练一个行人图像特征提取模型 $f(w;x)$,使得其在目标域测试集上能取得较高的 Re-ID 准确率. 目标域测试集一般被划分为两个子集:查询集 $X^q=\{x_i\}(i=1,2,\cdots,N_q)$ 和候选集 $X^g=\{x_i\}(i=1,2,\cdots,N_g)$. 对于任意一个查询样本 $\boldsymbol{x}_q\in X^q$,根据其对应的图像特征 $f(w;x_q)$ 在候选集 X^g 中检索含有相同行人目标的图像. 为确保 Re-ID 的准确性,训练后的模型应使得 x_q 与属于同一个行人的图像之间的特征信息相似性大于与 X^g 中来自不同行人的其他样本之间的相似性.

3.2 特征提取网络

如图1所示,本文采用的特征提取网络包含三个模块:(1)基于卷积神经网络的主干网络;(2)下采样模块;(3)特征嵌入模块. 主干网络采用在 ImageNet 数据集[46]上预训练的 ResNet-50 网络[47]. 输入行人图像经过主干网络提取中间特征后,使用全局平均池化层进行下采样. 下采样后的中间特征输入特征嵌入模块. 特征嵌入模块由一个全连接层和一个一维的批次归一化层(batch normalization layer)组成. 最终的行人图像特征由特征嵌入模块输出.

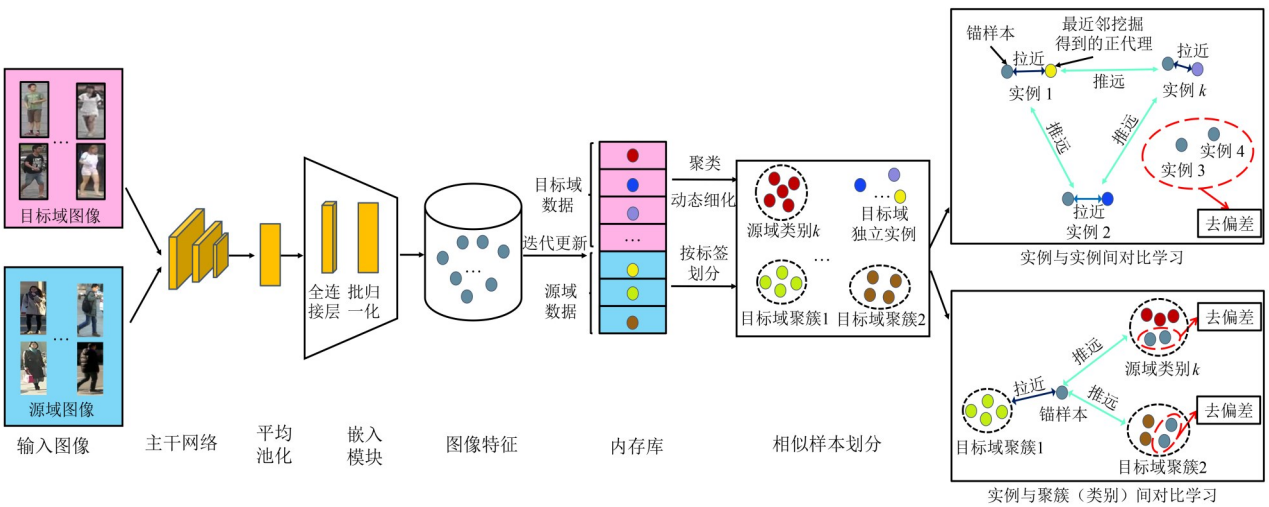


图1 本文提出的PHCL方法示意图

3.3 混合对比学习过程

为了将目标域和源域统一到端到端的对比学习框架中,本文根据在目标域数据集X^t上的聚类结果将其划分为两个子集:未聚类的独立实例集合X_I^t、已聚类的样本集合X_p^t. 直观上,每个聚簇的聚类标签可以看作属于该聚簇的样本的伪标签. 对于未聚类的样本,本文将其作为独立的实例进行模型训练. 基于这两个目标域子集以及源域数据集,本文提出一种混合对比学习方法,通过吸引相似样本和拉远不同独立实例,来指导网络学习有判别性的特征空间.

为了进行混合对比学习,需要计算所有实例与实例之间的相似度以及实例与中心之间的相似度. 然而,在基于批处理的训练过程中难以直接进行大量的相似度计算. 为了解决这个问题,本文提出的方法采用一个内存库$M=\{m_i\}(i=1,2,\cdots,N)$来存储所有训练样本特征的近似值,其中$N=N_s+N_t$. 对于训练图像$x_i$,假设其对应的特征为$f_i=f(w;x_i)$,则混合对比学习的优化目标为

$$L_{\mathrm{CL}}=-\frac{1}{N_b}\sum_{i=1}^{N_b}\ln\frac{\mathrm{e}^{f_i^+f_i/\tau}}{D} \tag{1}$$

其中,

$$D=\sum_{j=1}^{N_I^t}\mathrm{e}^{m_jf_i/\tau}+\sum_{j=1}^{N_c^t}\mathrm{e}^{c_jf_i/\tau}+\sum_{j=1}^{N_c^s}\mathrm{e}^{v_jf_i/\tau} \tag{2}$$

其中,N_b为批次的大小;N_I^t为目标域未聚类的独立实例个数;N_c^t为目标域聚簇的个数;N_c^s为源域行人身份标签个数;参数τ用于调节数据分布的集中程度;m_j为第j个独立实例特征的内存库近似值;c_j为目标域数据中第j个聚簇中心;v_j为源域数据中第j个行人的所有特征数据的中心;m_j、c_j以及v_j在这里被当作对比学习中f_i的负代理;f_i^+为图像特征f_i的正代理,其取值为

$$f_i^+=\begin{cases}\mathrm{NN}(m_i,M), & x_i\in X_I^t\\ c_l, & x_i\in X_p^t\wedge x_i\in C_l\\ v_k, & x_i\in X^s\wedge x_i\in V_k\end{cases} \tag{3}$$

其中,C_l为目标域第l个聚簇;c_l为该聚簇所有图像对应的特征数据的中心;V_k为源域数据中第k个行人的所有图像;v_k为这些图像对应的特征数据的中心. $\mathrm{NN}(m_i,M)$代表m_i在内存库M中的最近邻. 如果直接使用m_i作为未聚类独立实例特征f_i的正代理,由于m_i采用如式(7)所示的动量更新(momentum update)方式生成,则其与f_i差异性较小,二者构成的是较容易的正样本对,难以为对比学习提供有效对比信息. 而使用$\mathrm{NN}(m_i,M)$为f_i生成的正代理,其得到的是与f_i差异性较大的样本,例如,同一行人在不同摄像头下的样本,或者相同摄像头下同一行人不同时刻的样本. 此时,$\mathrm{NN}(m_i,M)$与f_i构成的是较难的正样本对. 因此,通过最近邻挖掘的方式为独立实例构建正代理,可以使得模型能够真正拉近两个不同的相似样本,从而学习到更丰富的语义信息. 关于此方法的有效性验证,具体见4.7节实验分析.

虽然式(1)能够有效指导模型的跨领域对比学习,但其面临着一个不可忽视的问题:如果没有标签信息的帮助,上述对比学习过程可能会引入一些偏差. 例如,对于没有标注信息的数据,如果一个被指定为f_i的负代理实际上与f_i属于同一个真实类别,则模型的训练可能会向错误的方向偏移. 事实上,这样的负代理是假负的. 为了解决这个问题,本文在对比学习过程中引入去偏差操作,降低假负代理带来的负面影响.

为了区分使用的负代理是真负还是假负,本文将目标域数据集X^t上的分布$p(x)$分解为

$$p(x)=\tau^+p^+(x)+\tau^-p^-(x) \tag{4}$$

其中,$\tau^+=p(y)$为类别为y概率,并假设其是均匀分布的;$\tau^-=1-\tau^+$为与y不同类别的概率;$p^+(x)$为正样本的分布;$p^-(x)$为负样本的分布,其也可以被表示为$\left(p(x)-\tau^+p^+(x)\right)\big/\tau^-$. 在对比学习过程中,训练数据中含有和锚样本相似的正样本和不相似的负样本,所以全体样本分布$p(x)$必然由正样本分布$p^+(x)$和负样本分布$p^-(x)$组成,这是对$p(x)$做如式(4)所示的分解的前提. 然而,在无监督域自适应设定下,训练阶段目标域的标注信息是未知的,所以引入超参数τ^+和τ^-来调节$p^+(x)$和$p^-(x)$的比例.

需要注意的是,在式(1)描述的对比学习过程中,由于缺少标注信息,目标域数据集的负样本分布$p^-(x)$是未知的,负代理m_j和c_j是使用$p(x)$而不是$p^-(x)$来计算的. 这表明在负样本采样中存在偏差. 为了解决这个问题,本文替换式(1)中与负代理相关的项,从而得到去偏差后的对比学习损失函数:

$$L_{\mathrm{DCL}}=-\frac{1}{N_b}\sum_{i=1}^{N_b}\ln\frac{\mathrm{e}^{f_i^+f_i/\tau}}{\hat{D}} \tag{5}$$

其中,

$$\hat{D}=\frac{1}{1-\tau^+}\left(\sum_{j=1}^{N_I^t}\mathrm{e}^{m_jf_i/\tau}+\sum_{j=1}^{N_c^t}\mathrm{e}^{c_jf_i/\tau}-\tau^+\mathrm{e}^{f_i^+f_i/\tau}\right)+\sum_{j=1}^{N_c^s}\mathrm{e}^{v_jf_i/\tau} \tag{6}$$

通过最小化式(5)来指导模型训练,一方面能够使得模型能够真正拉近两个不同的相似样本,学习到更丰富的语义信息,另一方面能够降低负代理偏差的不利影响. 以上措施可以确保训练后的特征网络具备如下特点:(1)群组内特征数据的内聚;(2)不同实例间特征的相互排斥. 具体而言,假设$\tau^+=0$,式(6)将退化为式(2),对比学习损失函数L_{DCL}将退化为L_{CL}. 此时,$p^-(x)=\left(p(x)-\tau^+p^+(x)\right)\big/\tau^-=p(x)$,这表明当使用$L_{\mathrm{CL}}$指导对比学习过程时,会将目标域全体训练样本分布

$p(x)$当作负样本分布$p^-(x)$,从中采样样本来计算目标域的负代理m_j和c_j. 然而,在实际应用场景中,τ^+不可能为0,此时$p^-(x)=\left(p(x)-\tau^+p^+(x)\right)\big/\tau^- \neq p(x)$,这表明使用$L_{\mathrm{CL}}$指导对比学习过程时可能会把和锚样本$f_i$相似的样本定义为负样本,从而产生负代理采样偏差. 所以,对比学习的实现需要从$p^-(x)$中采样样本来计算负代理m_j和c_j. 然而,由于训练过程中目标域数据的标注信息是未知的,本文使用式(5)和式(6)定义的损失函数L_{DCL}来实现对比学习过程中的负代理采样去偏差操作. 式(6)中前三项与目标域无标注数据的负代理计算相关,第四项与源域有标注数据的负代理计算相关. 其中,前两项代表使用从$p(x)$中采样的样本来计算负代理m_j和c_j,第三项代表消除前述采样中可能引入的正样本,从而实现负代理采样去偏差操作. 由于目标域数据分布$p(x)$中实际的负样本分布$p^-(x)$是未知的,上述去偏差操作由超参数τ^+来调节. 通过调整超参数τ^+,对比学习中负代理m_j和c_j的计算可以得到优化,从而消除正样本的潜在干扰,更好地逼近负代理的真实值. 关于超参数τ^+的选择,详见4.3.1节以及4.3.3节中的参数分析.

3.4 内存库的更新策略

式(5)需要在每一次迭代中计算所有实例与实例之间的相似度以及实例与聚簇中心之间的相似度. 为了达到这一计算目的,本文提出的方法使用内存库$M=\{m_i\}(i=1,2,\cdots,N)$来存储所有训练样本特征数据的近似值.

在第一轮训练开始时,使用所有行人图像的特征值来初始化内存库M中的值. 在接下来的每一轮训练中,内存库中的元素的更新策略为

$$m_i=\begin{cases}f_{i,r}, & r=0\\ \alpha m_i+(1-\alpha)f_{i,r}, & r>0\end{cases} \tag{7}$$

其中,r为当前的训练轮次;$f_{i,r}$为在第r轮训练中,行人图像x_i的特征;α为内存库元素的更新率. 每轮训练过程中,基于更新后的内存库M,可计算出目标域中第j个聚簇的中心,计算式为

$$c_j=\frac{1}{N_{C_j}}\sum_{m_i\in C_j}m_i \tag{8}$$

其中,C_j为目标域已聚类数据X_p^t中的第j个聚簇;N_{C_j}为C_j中的数据个数.

同理,基于更新后的内存库M,源域中第j个行人的所有特征数据中心的计算式为

$$v_j=\frac{1}{N_{V_j}}\sum_{m_i\in V_j}m_i \tag{9}$$

其中,V_j为源域数据X^s中第j个行人的所有图像;N_{V_j}为V_j中的数据个数.

内存库M存储所需的空间复杂度为$O(N)$,一个训练轮次里M中所有元素使用式(7)进行更新,所需的时间复杂度为$O(N)$,其中,$N=N_s+N_t$为训练样本个数. 使用式(8)计算目标域第j个聚簇中心所需的空间复杂度为$O(N_{C_j})$,时间复杂度为$O(N_{C_j})$. 使用式(9)计算源域中第j个行人的所有特征数据的中心所需的空间复杂度为$O(N_{V_j})$,时间复杂度为$O(N_{V_j})$.

3.5 聚类结果的渐进式细化策略

内存库M中保存所有训练数据的特征近似值,在训练的每一个轮次中,本文使用基于密度的带有噪声的空间聚类应用(Density-Based Spatial Clustering of Applications with Noise,DBSCAN)算法[48]对M中目标域特征进行聚类,从而根据聚类结果,将目标域数据集X^t划分为X_I^t和X_p^t两个子集. 本文提出的混合对比学习的有效性依赖于内存库M上的聚类结果. 然而,由于网络参数w和内存库M是迭代更新的,因此,聚类结果在每个训练轮次都会发生变化. 此外,并非所有聚类结果都是可靠的. 将实例分配到错误的聚簇将对混合对比学习产生不利影响. 因此,需要在每个训练轮次对M上的聚类结果进行动态细化,将属于可靠聚类的样本划分到子集X_p^t中,其他样本被划分到子集X_I^t中.

为了改进聚类结果,本文提出如式(10)所示的类别一致性准则:

$$\begin{aligned}H(C_j)&=-\sum_{i=1}^{N_{C_j}}P(c_j|f_i)\ln\left(P(c_j|f_i)\right)\\&=-\sum_{i=1}^{N_{C_j}}\frac{\mathrm{e}^{c_jf_i/\tau}}{\sum_{k=1}^{N_{C_j}}\mathrm{e}^{c_jf_k/\tau}}\ln\left(\frac{\mathrm{e}^{c_jf_i/\tau}}{\sum_{k=1}^{N_{C_j}}\mathrm{e}^{c_jf_k/\tau}}\right)\end{aligned} \tag{10}$$

其中,$P(c_j|f_i)$为样本x_i被划分到第j聚簇的概率;参数τ用于调节数据分布的集中程度. 类别一致性准则$H(C_j)$是基于熵的度量,$H(C_j)$值小意味着类别一致性高. 具体来说,$H(C_j)$较小,表示聚簇C_j的样本位于密度较低的区域,这表明这些样本更容易与其他聚簇的样本区分开来. 相反,较大的$H(C_j)$意味着聚簇C_j处于样本难以正确识别的密集区域. 在每个训练轮次中,根据类别一致性准则$H(C_j)$对所有聚簇进行递增排序,并将top-K的聚簇视为可靠聚类.

在上述聚类结果渐进式细化过程中,每一轮细化需要对N个训练样本以及N_c^t个聚簇的类别一致性数据进行存储,其对应的空间复杂度为$O(N+N_c^t)$. DBSCAN算法对N个训练样本进行聚类所需的时间复杂度为$O(N^2)$;使用式(10)对每一个聚簇计算类别一致性所需的时间复杂度为$O(N_{C_j})$,则对所有聚簇进行类别一致性计算所需的时间复杂度为$O(N_c^tN_{C_j})$;此外,对聚簇按照类别一致性排序所需的时间复杂度为$O\left(N_c^t\ln(N_c^t)\right)$. 综上,每一轮细

化的时间复杂度为 $O\left(N^2+N_c^t N_{C_j}+N_c^t \ln(N_c^t)\right)$.

4　实验结果与分析

4.1　数据集与评价指标

为了验证所提出方法在UDA Re-ID任务上的性能,本文在三个主流的数据集Market-1501[49]、MSMT17[50]以及DukeMTMC-reID[51]上进行实验.

Market-1501数据集由来自6个摄像头的32 668张图像组成,这些图像属于1 501个行人.Market-1501数据集由两个子集组成:训练集包含来自751个行人的12 936张图像,测试集包含来自750个行人的19 732张图像.在测试阶段,使用来自750个行人的33 368张图像作为查询(query)样本.这些查询图像用于在包含19 732张图像的图库(gallery)中检索具有相同身份的行人.

MSMT17数据集包含来自15个摄像头的126 441张图像,这些图像属于4 101个行人.训练子集包含来自1 041个行人的32 621张图像,而测试子集包含来自3 060个行人的93 820张图像.查询样本包含11 659张图像.

DukeMTMC-reID包含从1 404人收集的36 411张标注图像,这些图像通过8个同步摄像头采集.其中,属于702人的16 522张图像用于训练,另外702人的2 228张图像作为查询样本.此外,在图库中还包含来自其他408人的大量干扰图像,使得测试阶段的图库拥有来自1 110人的17 661张图像.3个数据集的统计信息如表1所示.

表1　实验所用数据集的统计信息

数据集	训练行人数	训练图像数	查询图像数	图库图像数
Market-1501	751	12 936	3 368	19 732
MSMT17	1 041	32 621	11 659	82 161
DukeMTMC-reID	702	16 522	2 228	17 661

本文使用平均精度均值(mean Average Precision,mAP)和累积匹配特性(Cumulative Matching Characteristic,CMC)精度作为算法性能的评价指标,其中,CMC精度报告首次命中准确率(Rank-1准确率).

mAP的定义为

$$\mathrm{mAP}=\frac{1}{Q}\sum_{q=1}^{Q}\mathrm{AP}_q \tag{11}$$

其中,Q为查询个数;AP_q为单个查询q的平均精度.当前n个检索结果中有m个是正确匹配时,其可以表示为

$$\mathrm{AP}_q=\frac{1}{m}\sum_{k=1}^{n}P(k)\cdot\mathrm{rel}(k) \tag{12}$$

其中,$P(k)$为前k个检索结果的精度(precision);$\mathrm{rel}(k)$为指示函数,当第k个结果是正确的时取1,否则取0.

CMC的定义为

$$\mathrm{CMC@}k=\frac{1}{Q}\sum_{q=1}^{Q}\mathrm{rel}(r_q\leqslant k) \tag{13}$$

其中,CMC@k为在前k个检索结果中,正确匹配首次出现的概率,如CMC@1代表Rank-1准确率;Q为查询个数;r_q为第q个查询中,正确匹配首次出现在检索结果中的位置;rel(·)为指示函数,若条件成立则值为1,否则值为0.

4.2　实验环境设置

本文采用在ImageNet[46]数据集上预训练的ResNet-50网络[47]作为PHCL方法的主干网络.在输入网络之前,图像尺寸被调整为256×128,并且进行水平翻转、填充和随机擦除等数据增广操作.优化器采用Adam[52],其权重衰减(weight decay)系数为0.000 5.学习率被初始化为0.003 5,然后每隔50个训练轮次(epoch)衰减为原始值的1/10,总的训练轮次为100.内存库元素的更新率α(见式(7))被设置为0.3.式(4)中的参数τ^+被设置为0.2,式(5)与式(10)中的参数τ被设置为0.06.可靠聚类系数K被设置为80%.在训练过程中,每个小批次(mini-batch)数据包含8个行人,每个行人16张图像.源域样本基于真实标注信息构建mini-batch,目标域样本则依据聚类生成的伪标签构建mini-batch.若某个行人图像不足16张(如目标域中未聚类的独立实例),则通过重复采样补足.由于采用可缓存源域与目标域全部特征近似值的内存库,mini-batch无须同时包含源域与目标域样本.在一个训练轮次中,源域和目标域训练数据依次以mini-batch的形式输入模型即可.

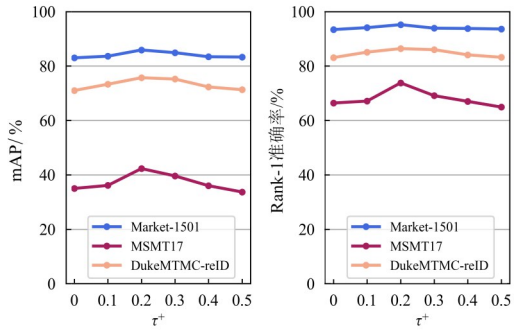
本文采用的UDA实验范式如下:在训练阶段,如果一个数据集作为目标域时,其标注信息对模型是不可见的,此时,这个数据集在训练阶段是完全无标注的,符合无监督学习的定义.而在测试阶段,目标域数据集的标注信息只作为检索匹配准确率计算时的金标准,不会泄露到训练阶段.当一个数据集在训练阶段作为源域时,其标注信息对模型而言是可见的,此时源域数据集的标注信息可以用于领域自适应.所有实验均在搭载NVIDIA A800图形处理单元(Graphics Processing Unit,GPU)的服务器上进行.

4.3　参数分析

4.3.1　参数τ^+

式(4)中的参数τ^+表示去偏差过程中正样本的类别概率.图2显示关于τ^+的参数分析过程.从图2可得,当τ^+从0增大到0.2时,本文提出的PHCL方法的性能在逐渐提升,当τ^+从0.2开始继续在增大时,性能开始逐步下降.因此,本文在Market-1501、MSMT17以及

DukeMTMC-reID数据集上将参数τ^+设置为0.2,从而使得模型取得最佳性能.

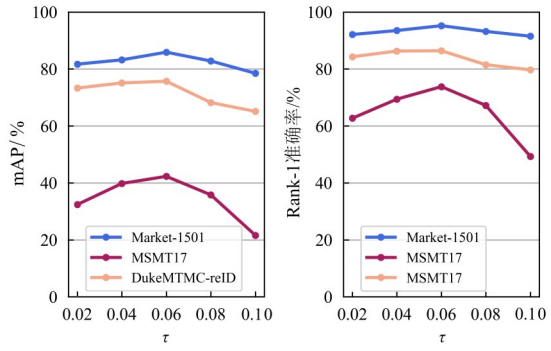


(a) mAP (b) Rank-1准确率

图2 参数τ^+的分析结果

4.3.2 参数τ

式(5)与式(10)中参数τ的大小会影响特征空间中数据的聚集程度.当τ较大时,数据分布会比较平滑;而当τ比较小时,数据分布会比较尖锐.如图3所示,当τ从0.02增大到0.06时,本文提出的PHCL方法在三个数据集上的mAP指标和Rank-1准确率开始逐步增加.当τ的值继续增大,三个数据集上的mAP和Rank-1准确率开始呈现出下降趋势.因此,本文在Market-1501、MSMT17以及DukeMTMC-reID数据集上将参数τ的值设置为0.06,从而使得模型取得最佳性能.



(a) mAP (b) Rank-1准确率

图3 参数τ的分析结果

4.3.3 参数τ^+和τ的协同作用分析

为了分析参数τ^+和τ的协同作用,本文在Market-1501数据集上对τ^+和τ进行联合网格搜索实验.具体而言,本文选取$\tau^+\in\{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$以及$\tau\in\{0.02, 0.04, 0.06, 0.08, 0.1\}$的多组参数组合进行实验,并记录模型的mAP指标以及Rank-1准确率.从表2可得,当固定τ^+和τ中的一个参数时,PHCL方法的性能随着另一个参数的增加而出现先增加后下降的趋势.例如,当τ固定在0.1时,随着τ^+从0增加到0.3,PHCL方法的mAP指标以及Rank-1准确率逐渐增加;当τ^+从0.3增加到0.5时,PHCL方法的性能开始下降.从表2还可以看出,当τ^+取0.2,τ取0.06时,PHCL方法取得85.9%的mAP指标以及95.2%的Rank-1准确率;而当τ^+取0.3,τ取0.04时,PHCL方法取得与前述参数组合比较接近的性能,mAP指标为85.6%,Rank-1准确率为95.1%.此实验结果表明:存在多种τ^+和τ的组合可以使得模型表现良好,二者存在协同作用.为了使模型取得最佳性能,本文将参数τ^+设置为0.2,将参数τ设置为0.06.

表2 Market-1501数据集上参数τ^+和τ的协同作用分析 单位:%

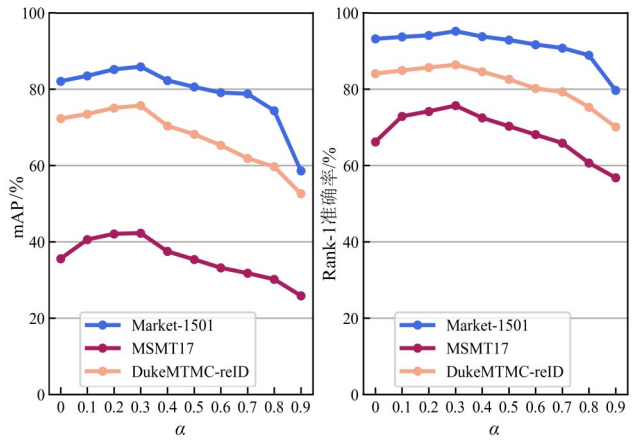
τ \ τ^+	0	0.1	0.2	0.3	0.4	0.5
0.02	79.3/90.8	80.9/91.7	81.7/92.1	80.2/91.4	79.6/90.9	78.8/90.3
0.04	81.9/92.7	82.6/93.3	83.2/93.5	85.6/95.1	83.3/93.8	82.5/93.2
0.06	83.0/93.4	83.2/93.8	85.9/95.2	84.7/93.9	83.5/93.6	83.3/93.5
0.08	81.7/92.6	83.7/94.1	82.8/93.2	82.4/92.8	81.8/93.0	81.3/92.4
0.10	76.2/90.3	77.8/91.1	78.5/91.5	79.6/92.1	76.9/90.3	75.4/88.6

4.3.4 参数α

式(7)中的参数α表示内存库元素的更新率.如图4所示,当α从0开始增加到0.3时,本文提出的PHCL方法在三个数据集上的mAP指标和Rank-1指标呈现缓慢上升的趋势.当α继续增大时,PHCL方法在三个数据集上的性能开始下降.当α取0.9时,PHCL方法的性能下降明显,这是因为较大的α不利于前一次迭代时特征信息的保留,无法为对比学习提供有效的全量样本近似特征值.因此,本文在Market-1501、MSMT17以及DukeMTMC-reID数据集上将参数α的值设置为0.3,从而使得模型取得最佳性能.

4.3.5 参数K

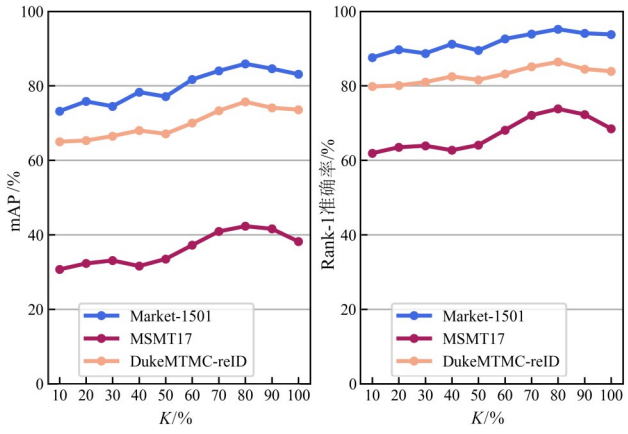
参数K代表聚类结果的渐进式细化策略中可靠聚



(a) mAP (b) Rank-1准确率

图4 参数α的分析结果

类的比例. 从图5可得,当K从10%增加到80%时,PHCL方法在三个数据集上的mAP指标和Rank-1准确率呈现振荡上升的趋势. 当K取80%时,PHCL方法取得最佳性能. 随着K继续增大,PHCL方法的性能开始下降. 其中,K取100%时,等价于未进行渐进式细化,全部聚簇都被认为是可靠聚类. 此时,相较K取80%时,PHCL方法的性能下降明显. 因此,本文在Market-1501、MSMT17以及DukeMTMC-reID数据集上,将参数K的值设置为80%,从而使得模型取得最佳性能.



(a) mAP　　(b) Rank-1准确率

图5　参数K的分析结果

4.4　与现有方法的对比实验

本文将提出的PHCL方法与一些现有的无监督Re-ID方法进行对比,这些参与对比实验的方法大致可以分为两大类:UDA和FUL. 对比实验结果如表3、表4以及表5所示,粗体数据为最优结果.

4.4.1　与FUL方法的对比

如表3~表5所示,本文提出的PHCL方法在Market-1501、MSMT17以及DukeMTMC-reID数据集上,超过所有参与对比的FUL方法. 具体而言,在Market-1501数据集上,与置信度引导的聚类中心(Confidence-Guided Centroids,CGC)[53]方法相比,PHCL方法的mAP指标和Rank-1准确率分别领先0.3个百分点以及0.9个百分点. CGC方法虽然利用聚类算法为无标注图像生成伪标签,然后构建基于聚簇的对比学习,然而,此方法只进行单一层次的对比学习,且未能解决负代理采样中的偏差问题. 在Market-1501数据集上,与层次化相机感知对比扩展(Hierarchical Camera-Aware Contrast Extension,HCACE)[54]方法相比,PHCL方法在mAP指标上提升了2.5个百分点,Rank-1准确率提升了1.5个百分点. 在MSMT17数据集上,PHCL方法的mAP指标超过HCACE方法0.7个百分点,Rank-1准确率超过HCACE方法1.4个百分点. 而在DukeMTMC-reID数据集上,与HCACE方法相比,PHCL方法在mAP指标上领先4.2个百分点,在Rank-1指标上领先2.3个百分点. HCACE方法利用摄像头标签作为粗略地标注信息,将无标注的图像特征聚簇进一步划分为更小的子集,从而构建多层次的对比学习过程. 然而,此方法在各个子集内以及子集之间的对比学习中依然存在负代理采样偏差. 此外,在DukeMTMC-reID数据集上,与混合动态对比和概率蒸馏(Hybrid Dynamic Contrast and Probability Distillation,HDCPD)[55]相比,PHCL方法将mAP指标和Rank-1准确率分别提升了2.2个百分点与1.0个百分点. HDCPD方法构建融合实例与聚簇的混合对比学习,然而,此方法在实例级对比学习中锚样本和正代理之间的差异较小;在负代理计算过程中,与锚样本语义上相似的样本依然可能会被划分为负样本.

表3~表5中展示的FUL方法聚类引导的非对称对比学习(Cluster-guided Asymmetric Contrastive Learning,CACL)[27]、联合生成式和对比学习(Joint Generative and Contrastive Learning,JGCL)[30]、CGC[53]、HCACE[54]、HDCPD[55]、基于部件的伪标签细化(Part-based Pseudo Label Refinement,PPLR)[56]、实例间对比编码(Inter-instance Contrastive Encoding,ICE)[57]、时空双重注意力(Spatial and Temporal Dual-Attention,STDA)[58]以及实时记忆(Real-Time Memory,RTMem)[59]与本文所提出的PHCL方法有着类似的思路:通过集成不同层次的对比学习来提升无监督Re-ID性能. 然而,PHCL方法在三个数据集上都展现出比这些方法更好的性能. 本文认为产生此实验现象的主要原因如下:

(1)这些方法采用FUL范式,无法利用有标注的源域数据. 而本文提出的PHCL方法采用UDA方法,在对比学习过程中,可以充分借助源数据集的标注信息,以强化学习到的特征空间的可分性.

(2)在构建对比学习过程中,这些方法不仅存在正样本对差异较小的问题,而且忽略负代理采样中的偏差. 而PHCL方法,一方面采用最近邻挖掘来构建正样本对,使得模型能够真正拉近两个不同的相似样本,学习到更丰富的语义信息;另一方面PHCL方法在负代理采样时进行去偏差操作,能有效降低假负代理带来的不利影响.

4.4.2　与UDA方法的对比

如表3~表5所示,相对于参与比较的UDA方法,本文提出的PHCL方法在Market-1501、MSMT17以及DukeMTMC-reID数据集上均展现出更好的性能. 例如,与多类中心动态对比学习(Multi-Class Center Dynamic Contrastive Learning,MCC-DCL)[60]方法相比,PHCL方法在Market-1501数据集上将mAP指标提升0.1个百分点,将Rank-1准确率提升0.4个百分点. 在

MSMT17数据集上，与MCC-DCL方法相比，PHCL方法性能提升6.1个百分点的mAP指标与7.1个百分点的Rank-1准确率. 在DukeMTMC-reID数据集上，PHCL方法领先MCC-DCL方法1.3个百分点的mAP指标与1.6个百分点的Rank-1准确率. MCCL-DCL方法利用多中心聚类生成伪标签，减少聚类过程产生的噪声标签. 然而，其在构建对比学习过程中，并没有关注负代理采样偏差问题. 与相机驱动对比学习(Camera-driven Contrastive Learning，CaCL)[61]方法相比，PHCL方法在Market-1501数据集上取得1.2个百分点的mAP指标以及1.4个百分点的Rank-1准确率提升. 而在MSMT17数据集上，相对于CaCL方法，PHCL方法的mAP指标领先5.8个百分点，Rank-1准确率领先7.2个百分点. CaCL方法基于相机标签提出了相机多样性损失(camera diversity loss)，从而减少同一相机下伪标签生成时的噪声. 然而，此方法无法有效解决跨相机下语义相似的行人图像的伪标签不一致问题. 此外，在DukeMTMC-reID数据集上，PHCL方法超过互补注意力驱动对比学习与难样本探索(Complementary Attention-driven Contrastive learning with Hard-sample Exploring，CACHE)[62]方法2.5个百分点的mAP指标与2.2个百分点的Rank-1准确率. CACHE方法将互补注意力机制引入对比学习中，减少伪标签噪声，提升图像特征的可判别性. 然而，此方法在实施对比学习时，与锚样本相似的样本被划分为负样本的问题依然存在.

在参与对比的UDA方法中，减少背景诱导的域偏移(Reducing Background Induced Domain Shift，RBIDS)[63]方法、属性判别与选择性样本丢弃相结合(Attribute Discrimination Combined with Selected Sample Dropout，ADC-SSD)[64]方法、稀疏注意力增强的域适应(Sparse-Attention Augmented Domain Adaptation，SAADA)[65]方法、跨域平衡网络(Inter-Domain Equilibrium Network，IDENet)[66]方法都通过尝试设计不同的网络结构，将源域的监督信息迁移到目标域. 然而，源域与目标域之间的领域差异(domain gap)是天然存在的，这类方法没有充分挖掘目标域的图像间的相似性信息，其性能严重依赖源域的数据质量. 而本文提出的PHCL方法将对比学习与UDA有机结合，在充分利用源域数据的标注信息的同时，还能有效挖掘目标域的类内相似性与实例间差异性，从而在UDA Re-ID任务上产生更优的性能. 而JGCL[30]、自步对比学习(Self-paced Contrastive Learning，SpCL)[37]、MCC-DCL[60]、多域联合学习(Multi-Domain Joint Learning，MDJL)[67]、多粒度一致性网络(Multi-Granularity Consistency Network，MGCN)[68]以及二次近似卷积与图采样(Quadratic Approximation Convolution with Graph Sampling，QAConv-GS)[69]这几种方法，虽然也采用对比学习的方法来解决UDA Re-ID问题，但这些方法存在以下问题：(1)正样本对差异较小；(2)忽略负代理采样中的偏差. 而本文提出的PHCL方法通过最近邻挖掘为未聚类的实例构建正样本对，以及负代理去偏差操作来优化对比学习过程的构建，从而在UDA Re-ID任务上取得更好的性能.

表3 Market-1501数据集上的对比实验结果 单位:%

方法	来源	类型	源数据集	mAP	Rank-1准确率
JGCL[30]	CVPR'2021	FUL	—	66.8	87.3
CACL[27]	TIP'2022	FUL	—	80.9	92.7
PPLR[56]	CVPR'2022	FUL	—	81.5	92.8
ICE[57]	ICCV'2021	FUL	—	82.3	93.8
STDA[58]	TITS'2023	FUL	—	82.7	93.1
RTMem[59]	TIP'2023	FUL	—	83.1	93.9
HCACE[54]	TMM'2024	FUL	—	83.4	93.7
CGC[53]	TIFS'2024	FUL	—	85.6	94.3
MDJL[67]	PR'2023	UDA	DukeMTMC-reID	59.8	80.3
JGCL[30]	CVPR'2021	UDA	DukeMTMC-reID	75.4	90.5
RBIDS[63]	TIINF'2023	UDA	DukeMTMC-reID	76.9	90.7
MGCN[68]	TOMM'2024	UDA	DukeMTMC-reID	81.9	93.2
ADC-SSD[64]	IVC'2024	UDA	DukeMTMC-reID	82.8	93.2
QAConv-GS[69]	CVPR'2022	UDA	MSMT17	56.9	82.4
SpCL[37]	NIPS'2020	UDA	MSMT17	77.5	89.7
SAADA[65]	PRL'2025	UDA	MSMT17	80.3	92.3
IDENet[66]	TIP'2025	UDA	MSMT17	83.6	93.3
CaCL[61]	ICCV'2023	UDA	MSMT17	84.7	93.8
MCC-DCL[60]	CEE'2024	UDA	MSMT17	85.8	94.8
PHCL	本文	UDA	DukeMTMC-reID	85.3	94.4
PHCL	本文	UDA	MSMT17	**85.9**	**95.2**

4.5 消融实验

为了验证所提出方法的各个组成部分的有效性，本文在Market-1501、MSMT17以及DukeMTMC-reID数据集上进行了多组消融实验. 在消融实验过程中，本文使用没有结合最近邻挖掘和负代理去偏差的混合对比学习损失函数来训练3.2节中设计的特征提取网络，以此作为基线模型. 消融实验结果如表6所示，粗体表示最优的结果. 其中NN代表使用最近邻挖掘构造正样本对，Debiasing代表负代理采样时的去偏差操作. "w/o细化"代表不使用聚类结果的渐进式细化策略. Market代表Market-1501数据集，MSMT代表MSMT17数据集，Duke代表DukeMTMC-reID数据集. "MSMT→Market"代表源域为MSMT17数据集，目标域为Market-1501数据集.

4.5.1 最近邻挖掘的有效性

如表6所示，第2行为在基线模型的基础上使用最

表4　MSMT17数据集上的对比实验结果　　单位:%

方法	来源	类型	源数据集	mAP	Rank-1准确率
JGCL[30]	CVPR'2021	FUL	—	21.3	45.7
CACL[27]	TIP'2022	FUL	—	23.0	48.9
PPLR[56]	CVPR'2022	FUL	—	31.4	61.1
STDA[58]	TITS'2023	FUL	—	31.8	62.6
CGC[53]	TIFS'2024	FUL	—	35.7	66.1
ICE[57]	ICCV'2021	FUL	—	38.9	70.2
RTMem[59]	TIP'2023	FUL	—	40.8	72.0
HCACE[54]	TMM'2024	FUL	—	41.6	72.4
MDJL[67]	PR'2023	UDA	DukeMTMC-reID	17.1	40.3
RBIDS[63]	TIINF'2023	UDA	DukeMTMC-reID	21.0	65.8
JGCL[30]	CVPR'2021	UDA	DukeMTMC-reID	29.7	54.4
ADC-SSD[64]	IVC'2024	UDA	DukeMTMC-reID	30.7	59.0
MGCN[68]	TOMM'2024	UDA	DukeMTMC-reID	30.7	61.8
QAConv-GS[69]	CVPR'2022	UDA	Market-1501	17.2	45.9
SpCL[37]	NIPS'2020	UDA	Market-1501	26.8	53.7
SAADA[65]	PRL'2025	UDA	Market-1501	31.4	60.0
IDENet[66]	TIP'2025	UDA	Market-1501	31.9	60.6
MCC-DCL[60]	CEE'2024	UDA	Market-1501	36.2	66.7
CaCL[61]	ICCV'2023	UDA	Market-1501	36.5	66.6
PHCL	本文	UDA	DukeMTMC-reID	42.1	73.3
PHCL	本文	UDA	Market-1501	**42.3**	**73.8**

表5　DukeMTMC-reID数据集上的对比实验结果　单位:%

方法	来源	类型	源数据集	mAP	Rank-1准确率
JGCL[30]	CVPR'2021	FUL	—	62.8	82.9
SECRET[70]	AAAI'2022	FUL	—	63.9	77.9
CACL[27]	TIP'2022	FUL	—	68.8	82.9
ICE[57]	ICCV'2021	FUL	—	69.9	83.3
HCACE[54]	TMM'2024	FUL	—	71.5	84.1
HDCPD[55]	TIP'2022	FUL	—	73.5	85.4
RBIDS[63]	TIINF'2023	UDA	Market-1501	61.0	79.2
MDJL[67]	PR'2023	UDA	Market-1501	62.8	78.6
SAADA[65]	PRL'2025	UDA	Market-1501	65.1	79.8
JGCL[30]	CVPR'2021	UDA	Market-1501	67.6	81.9
MGCN[68]	TOMM'2024	UDA	Market-1501	71.1	83.4
TAL-MIRN[71]	TCASVT'2021	UDA	MSMT17	48.7	68.4
DG-Net++[72]	ECCV'2020	UDA	MSMT17	58.2	75.2
CACHE[62]	TCASVT'2022	UDA	MSMT17	72.6	84.2
MCC-DCL[60]	CEE'2024	UDA	MSMT17	74.4	84.8
PHCL	本文	UDA	Market-1501	75.1	85.6
PHCL	本文	UDA	MSMT17	**75.7**	**86.4**

近邻挖掘构造正样本对后的实验数据. 从表6可以看出,与基线模型相比,"基线模型+NN"在Market-1501上将mAP指标从81.6%提升到82.8%,将Rank-1准确率从92.1%提升到93.5%. 在MSMT17数据集上,相对于基线模型,"基线模型+NN"的mAP指标和Rank-1准确率提升分别为7.9个百分点和11.9个百分点. 在DukeMTMC-reID数据集上,与基线模型相比,"基线模型+NN"将mAP指标从67.8%提升到71.0%,将Rank-1准确率从81.2%提升到83.1%. 这些实验结果表明:依靠最近邻挖掘构造正样本对可以有效加强对比学习的效果,提升训练后的模型在UDA Re-ID任务上的性能.

4.5.2　负代理去偏差的有效性

表6第3行数据代表在基线模型的基础上,对负代理采样进行去偏差操作后的实验结果. 如表6所示,与基线模型相比,"基线模型+Debiasing"在Market-1501数据集上取得2.9个百分点的mAP指标以及1.8个百分点的Rank-1准确率提升. 而在MSMT17数据集上,与基线模型相比,"基线模型+Debiasing"将mAP指标提升10.6个百分点,将Rank-1准确率提升13.6个百分点. 在DukeMTMC-reID数据集上,与基线模型相比,"基线模型+Debiasing"将mAP指标提升6.3个百分点,将Rank-1准确率提升3.1个百分点. 以上的实验结果表明负代理去偏差操作可以有效降低假负代理对模型训练的不利影响.

4.5.3　混合对比学习的有效性

表6最后一行中涉及的"基线模型+NN+ Debiasing"即为本文提出的采用混合对比学习的PHCL方法,其在

实例级与聚类级混合对比学习的基础上，引入基于最近邻挖掘的正样本构造方法以及负代理采样去偏差操作. 从表6可得，本文提出的PHCL方法在三个数据集上的性能都超过其他方法. 例如，相对于基线模型，PHCL方法在Market-1501数据集上将性能提升4.3个百分点的mAP指标以及3.1个百分点的Rank-1准确率. 相对于“基线模型+NN”，PHCL方法在MSMT17数据集上将mAP指标从36.7%提升到42.3%，将Rank-1准确率从67.6%提升到73.8%. 相对于“基线模型+Debiasing”，PHCL方法在DukeMTMC-reID数据集上的性能提升为1.6个百分点的mAP指标以及2.1个百分点的Rank-1准确率. 以上的实验结果验证了本文提出的混合对比学习在UDA Re-ID任务的有效性.

4.5.4 渐进式细化策略的有效性

表6第4行涉及“基线模型+NN+Debiasing w/o细化”为不使用渐进式细化策略的PHCL方法. 从表6可以看出，当使用渐进式细化策略时，Re-ID的准确率得到提升. 具体而言，相对于“基线模型+NN+Debiasing w/o细化”方法，“基线模型+NN+Debiasing”方法在Market-1501数据集上的mAP指标以及Rank-1准确率分别上升2.8个百分点和1.4个百分点. 在MSMT17数据集上，“基线模型+NN+Debiasing”方法将mAP指标提升4.1个百分点，将Rank-1准确率提升5.3个百分点. 而在DukeMTMC-reID数据集上，“基线模型+NN+Debiasing”方法领先“基线模型+NN+Debiasing w/o细化”方法2.1个百分点的mAP指标以及2.5个百分点的Rank-1准确率. 以上的实验结果验证了本文提出的渐进式细化策略的有效性.

表6 Market-1501、MSMT17以及DukeMTMC-reID数据集上的消融实验 单位: %

方法	MSMT→Market		Market→MSMT		MSMT→Duke	
	mAP	Rank-1	mAP	Rank-1	mAP	Rank-1
基线模型	81.6	92.1	28.8	55.7	67.8	81.2
基线模型+NN	82.8	93.5	36.7	67.6	71.0	83.1
基线模型+Debiasing	84.5	93.9	39.4	69.3	74.1	84.3
基线模型+NN+Debiasing w/o细化	83.1	93.8	38.2	68.5	73.6	83.9
基线模型+NN+Debiasing (本文)	**85.9**	**95.2**	**42.3**	**73.8**	**75.7**	**86.4**

4.6 数据可视化分析

为了更加直观地展示所提出方法的有效性，本文在Market-1501数据集上对4.5节中提到的“基线模型”“基线模型+NN”“基线模型+Debiasing”以及“基线模型+NN+Debiasing”这几种方法进行丰富的可视化实验，具体包括t分布随机邻域嵌入(t-distributed Stochastic Neighbor Embedding，t-SNE)[73]可视化以及前十检索结果可视化. 对于t-SNE可视化而言，其展示的是训练后模型在训练集上的特征空间可分性. 本文使用训练集里来自60个行人的共计1 000张图像进行t-SNE可视化分析，其结果如图6所示，其中相同颜色的特征点属于同一个行人. 对于前十检索结果可视化而言，其展示的是训练后的模型在测试集上Re-ID性能，其结果如图7所示，其中黄色边框表示查询样本，绿色边框表示正确匹配的结果，红色边框表示错误匹配的结果.

如图6所示，黑色实线椭圆内两组特征对应身份标签为658和659的两个行人. 这两名行人都穿着深色的T恤、黑色短裤. 图像反映出来的行人外观相似性对Re-ID的结果有着巨大的影响，而这两个行人由于其高度相似的衣着颜色和款式，可能会使得算法产生误判. 从图6可以看出，本文提出的PHCL算法能够有效分离出这两个行人的图像特征，而其他方法如“基线模型”“基线模型+NN”以及“基线模型+Debiasing”关于这两个行人生成的特征则十分接近或者重叠. 此外，红色虚线椭圆标注的特征点也有类似的实验现象. 例如，对于“基线模型”不能分离的身份标签为660和709的两行人特征，“基线模型+NN”和“基线模型+Debiasing”都能有效地将它们区分开. 这表明最近邻挖掘构造正样本对以及负代理去偏差这两种策略的有效性. 而“基线模型+NN+Debiasing”方法为两种策略的结合，即为本文提出的PHCL方法. 对于“基线模型+NN”不能区分的669和702两个行人，以及“基线模型+Debiasing”不能区分的700以及701两个行人，本文提出的PHCL方法在特征空间里均能有效地将它们分离. 上述t-SNE可视化结果表明本文方法在区分外观相似的不同行人图像时的有效性.

如图7所示，当使用相同的查询样本在测试集里进行检索时，本文提出的PHCL方法比“基线模型”“基线模型+NN”和“基线模型+Debiasing”召回更多正确匹配的图像(绿色边框). 大多数错误匹配的测试图像(红色边框)与查询样本有着相似的外观，如衣服颜色、款式等. 前十检索结果的可视化进一步表明本文方法在UDA Re-ID任务中区分视觉相似但身份不同的行人图像的潜力.

4.7 正样本对差异性分析

在增强未聚类独立实例集合X_I^t中，为了验证本文提出的PHCL方法正样本对差异性上的有效性，本文在

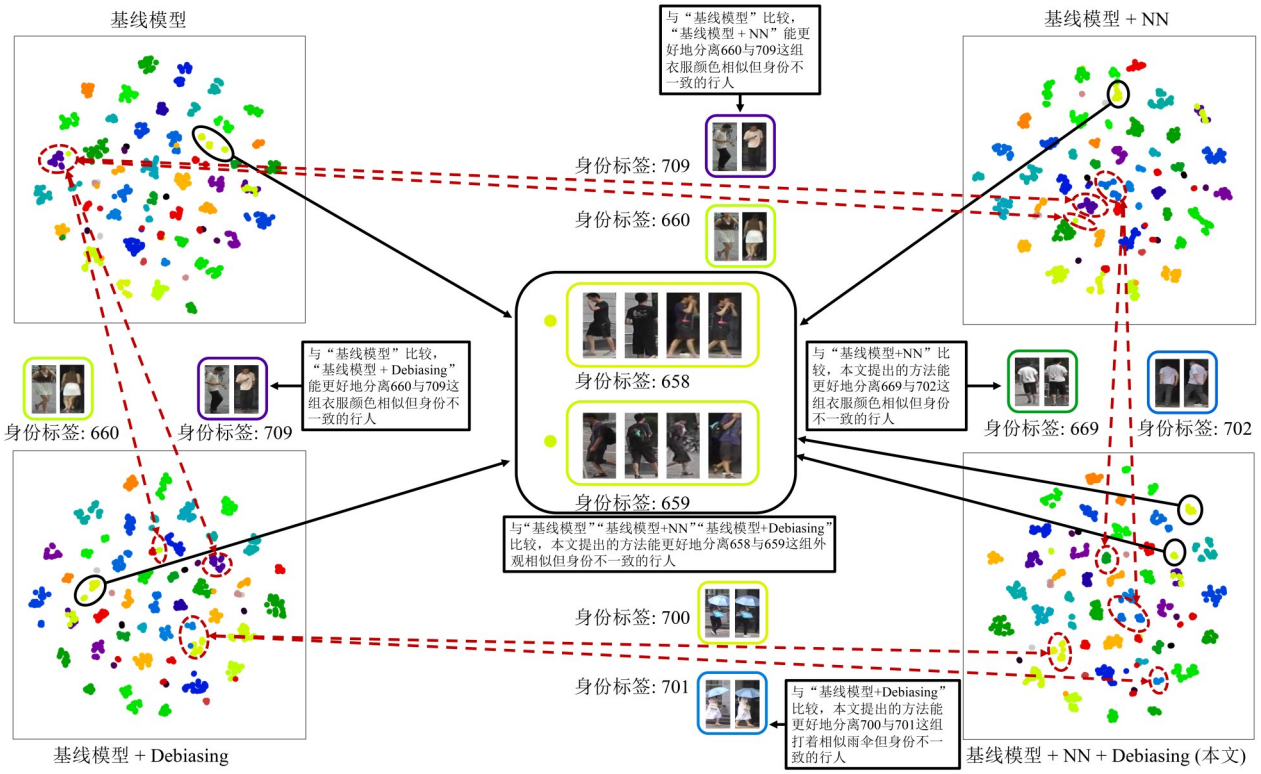


图 6　Market-1501 数据集上的 *t*-SNE 可视化

图 7　前十检索结果的可视化

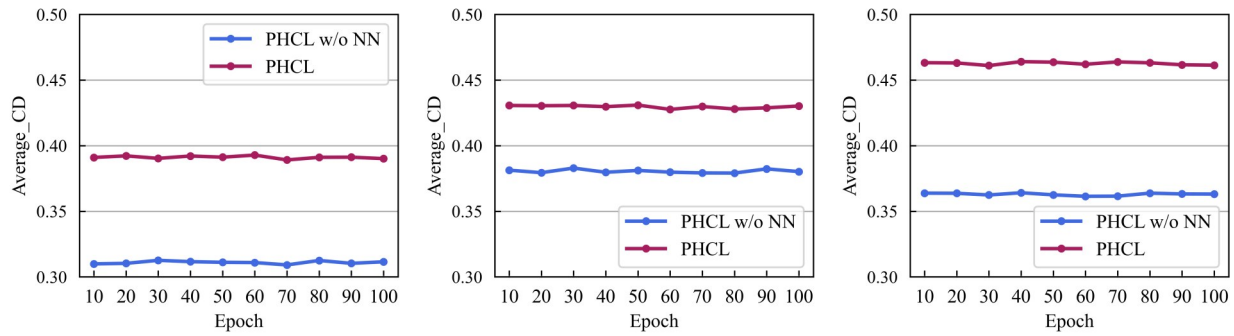
Market-1501、MSMT17以及DukeMTMC-reID数据集上对每10个训练轮次(epoch)的正样本对平均差异性进行统计. 具体而言,本文采用余弦距离(Cosine Distance,CD)度量锚样本f_i与正代理f_i^+之间的差异性:

$$\mathrm{CD}(f_i,f_i^+)=1-\frac{f_i\cdot f_i^+}{\|f_i\|\cdot\|f_i^+\|} \tag{14}$$

每10个训练轮次统计一次当前未聚类独立实例集合X_I^t中所有正样本对之间的CD,并求其平均值,作为当前X_I^t中正样本对差异性的度量:

$$\mathrm{Average_CD}=\frac{1}{N_I^t}\sum_{i=1}^{N_I^t}\mathrm{CD}(f_i,f_i^+) \tag{15}$$

如图8所示,“PHCL w/o NN”表示不使用最近邻挖掘为独立实例构建正代理时的PHCL方法. 此时,使用内存库M中与f_i对应的元素m_i作为正代理. 如图8所示,相对于“PHCL w/o NN”方法而言,本文所提出的PHCL方法在三个数据集上均能增大未聚类独立实例集合X_I^t中的正样本对差异性. 例如,在Market-1501数据集上,“PHCL w/o NN”方法的正样本对的差异性度量Average_CD在0.31左右波动,而PHCL方法的Average_CD值在0.39附近波动. 在MSMT17和DukeMTMC-reID数据集上,PHCL方法也取得了类似的正样本对差异性提升.



(a) Market-1501 (b) MSMT17 (c) DukeMTMC-reID

图8 正样本对差异性度量

5 结论

本文提出一种PHCL方法,旨在解决UDA Re-ID任务. 在每个训练轮次中,PHCL方法通过聚类和渐进细化两个步骤,有效地将无标注的目标域训练集划分为带伪标签的聚类样本以及未聚类的独立实例. PHCL方法通过将目标域中同一聚类或源域中同一身份标签的相似样本拉近,指导网络学习类内相似性,同时,通过在目标域中未聚类的实例间施加排斥作用,挖掘实例间差异性. 此外,PHCL方法使用最近邻挖掘生成正代理,增大正样本对的差异性,从而学习到更丰富的语义信息. 同时,PHCL方法在负代理采样过程中进行去偏差操作,减轻假负代理对模型训练的不利影响. 实验结果表明,PHCL方法在解决UDA Re-ID任务方面具有显著的有效性.

本研究存在一定的局限性,例如,在应对更加复杂、开放的户外场景时的泛化性能仍需进一步验证. 在后续研究中,本文研究者拟结合多模态数据,如视频时空信息、语音信息、人体姿态信息以及场景深度信息对模型进行迭代升级,增强算法对于复杂开放场景的适应性.

参考文献

[1] 周新民, 熊智谋, 史长发, 等. 基于多尺度卷积调制的医学图像分割[J]. 电子学报, 2024, 52(9): 3159-3171.
ZHOU X M, XIONG Z M, SHI C F, et al. Medical image segmentation based on multi-scale convolution modulation[J]. Acta Electronica Sinica, 2024, 52(9): 3159-3171. (in Chinese)

[2] 童康, 吴一全. 基于深度学习的小目标检测基准研究进展[J]. 电子学报, 2024, 52(3): 1016-1040.
TONG K, WU Y Q. Research advances on deep learning based small object detection benchmarks[J]. Acta Electronica Sinica, 2024, 52(3): 1016-1040. (in Chinese)

[3] 杨礼吉, 王家祺, 景丽萍, 等. 基于张量计算的卷积神经网络语义表示学习[J]. 计算机学报, 2023, 46(3): 568-578.
YANG L J, WANG J Q, JING L P, et al. Semantic representation learning of convolutional neural network based on tensor computation[J]. Chinese Journal of Computers, 2023, 46(3): 568-578. (in Chinese)

[4] 肖进胜, 吴婧逸, 郭浩文, 等. 精细化局部语义与属性学习的行人重识别[J]. 计算机学报, 2024, 47(10): 2387-2400.
XIAO J S, WU J Y, GUO H W, et al. Person re-identification network with fine-grained local semantics and attribute learning[J]. Chinese Journal of Computers, 2024,

47(10): 2387-2400. (in Chinese)

[5] 侯瑞兵, 常虹, 马丙鹏, 等. 基于时序多尺度互补特征的视频行人重识别[J]. 计算机学报, 2023, 46(1): 31-50.
HOU R B, CHANG H, MA B P, et al. Temporal multi-scale complementary feature for video person re-identification[J]. Chinese Journal of Computers, 2023, 46(1): 31-50. (in Chinese)

[6] 钱锦浩, 宋展仁, 郭春超, 等. 基于时空共现模式的视觉行人再识别[J]. 自动化学报, 2022, 48(2): 408-417.
QIAN J H, SONG Z R, GUO C C, et al. Visual person re-identification based on spatial and temporal co-occurrence patterns[J]. Acta Automatica Sinica, 2022, 48(2): 408-417. (in Chinese)

[7] 邵晓雯, 帅惠, 刘青山. 融合属性特征的行人重识别方法[J]. 自动化学报, 2022, 48(2): 564-571.
SHAO X W, SHUAI H, LIU Q S. Person re-identification based on fused attribute features[J]. Acta Automatica Sinica, 2022, 48(2): 564-571. (in Chinese)

[8] XIE Q K, ZHOU W G, QI G J, et al. Progressive unsupervised person re-identification by tracklet association with spatio-temporal regularization[J]. IEEE Transactions on Multimedia, 2020, 23: 597-610.

[9] SRIDHAR RAJ S, PRASAD M V N K, BALAKRISHNAN R. Spatio-temporal association rule based deep annotation-free clustering (STAR-DAC) for unsupervised person re-identification[J]. Pattern Recognition, 2022, 122: 108287.

[10] PANG Z Q, WANG C Y, ZHAO L L, et al. Cross-modality hierarchical clustering and refinement for unsupervised visible-infrared person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(4): 2706-2718.

[11] CHEN Z Q, CUI Z C, ZHANG C, et al. Dual clustering co-teaching with consistent sample mining for unsupervised person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(10): 5908-5920.

[12] LIN Y T, XIE L X, WU Y, et al. Unsupervised person re-identification via softened similarity learning[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 3390-3399.

[13] WANG Q, RAO W, SUN S N, et al. Unsupervised domain adaptation via domain adversarial training for speaker recognition[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE, 2018: 4889-4893.

[14] WANG W, LI H J, DING Z M, et al. Rethinking maximum mean discrepancy for visual domain adaptation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(1): 264-277.

[15] WANG J, ZHANG X L. Improving pseudo labels with intra-class similarity for unsupervised domain adaptation[J]. Pattern Recognition, 2023, 138: 109379.

[16] PEI J B, MEN A D, LIU Y, et al. Evidential multi-source-free unsupervised domain adaptation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(8): 5288-5305.

[17] AI Y A, ZHOU X Q, HUANG H B, et al. Uncertainty-aware source-free adaptive image super-resolution with wavelet augmentation transformer[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2024: 8142-8152.

[18] ZHU J J, BAI H T, WANG L. Patch-mix transformer for unsupervised domain adaptation: A game perspective[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 3561-3571.

[19] HAN G X, ZHANG X, LI C R. One-shot unsupervised cross-domain person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(3): 1339-1351.

[20] ANG E P W, LIN S, KOT A C. A unified deep semantic expansion framework for domain-generalized person re-identification[J]. Neurocomputing, 2024, 600: 128120.

[21] CHEN Y B, ZHU X T, GONG S G. Instance-guided context rendering for cross-domain person re-identification[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2019: 232-242.

[22] XIE J H, ZHAN X H, LIU Z W, et al. Delving into inter-image invariance for unsupervised visual representations[J]. International Journal of Computer Vision, 2022, 130(12): 2994-3013.

[23] CHEN X L, HE K M. Exploring simple Siamese representation learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 15750-15758.

[24] HE K M, FAN H Q, WU Y X, et al. Momentum contrast for unsupervised visual representation learning[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 9729-9738.

[25] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple

framework for contrastive learning of visual representations[C]//International Conference on Machine Learning. Cambridge MA: PMLR, 2020: 1597-1607.

[26] LI Y, HU P, LIU Z, et al. Contrastive clustering[C]//2021 AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021: 8547-8555.

[27] LI M K, LI C G, GUO J. Cluster-guided asymmetric contrastive learning for unsupervised person re-identification[J]. IEEE Transactions on Image Processing, 2022, 31: 3606-3617.

[28] XUAN S Y, ZHANG S L. Intra-inter camera similarity for unsupervised person re-identification[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 11926-11935.

[29] CHEN H, WANG Y H, LAGADEC B, et al. Learning invariance from generated variance for unsupervised person re-identification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(6): 7494-7508.

[30] CHEN H, WANG Y H, LAGADEC B, et al. Joint generative and contrastive learning for unsupervised person re-identification[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 2004-2013.

[31] JI H, WANG L, ZHOU S P, et al. Meta pairwise relationship distillation for unsupervised person re-identification[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 3641-3650.

[32] ZENG K W, NING M N, WANG Y H, et al. Hierarchical clustering with hard-batch triplet loss for person re-identification[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 13657-13665.

[33] WANG D K, ZHANG S L. Unsupervised person re-identification via multi-label classification[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 10981-10990.

[34] LIU W F, LI J F, LIU B D, et al. Unified cross-domain classification via geometric and statistical adaptations[J]. Pattern Recognition, 2021, 110: 107658.

[35] DAI C Q, PENG C, CHEN M. Selective transfer cycle GAN for unsupervised person re-identification[J]. Multimedia Tools and Applications, 2020, 79(17): 12597-12613.

[36] SONG J F, YANG Y X, SONG Y Z, et al. Generalizable person re-identification by domain-invariant mapping network[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 719-728.

[37] GE Y X, ZHU F, CHEN D P, et al. Self-paced contrastive learning with hybrid memory for domain adaptive object re-ID[J]. Advances in Neural Information Processing Systems, 2020, 33(1): 11309-11321.

[38] SUN B, FENG J, SAENKO K. Return of frustratingly easy domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2016: 10306.

[39] YE M, ZHANG X, YUEN P C, et al. Unsupervised embedding learning via invariant and spreading instance feature[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 6210-6219.

[40] GRILL J B, STRUB F, ALTCHÉ F, et al. Bootstrap your own latent: A new approach to self-supervised Learning[J]. Advances in Neural Information Processing Systems, 2020, 33(1): 21271-21284.

[41] CARON M, MISRA I, MAIRAL J, et al. Unsupervised learning of visual features by contrasting cluster assignments[J]. Advances in Neural Information Processing Systems, 2020, 33(1): 9912-9924.

[42] XIE X T, CHEN W Y, KANG Z, et al. Contrastive graph clustering with adaptive filter[J]. Expert Systems with Applications, 2023, 219: 119645.

[43] DENG X Z, HUANG D, CHEN D H, et al. Strongly augmented contrastive clustering[J]. Pattern Recognition, 2023, 139: 109470.

[44] YANG X H, LIU Y, ZHOU S H, et al. Cluster-guided contrastive graph clustering network[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(9): 10834-10842.

[45] HU Z X, XING Y, GU W H, et al. Driver anomaly quantification for intelligent vehicles: A contrastive learning approach with representation clustering[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(1): 37-47.

[46] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 248-255.

[47] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.

[48] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[J]. International Conference on Knowledge Discovery and Data Mining, 1996, 96(34): 226-231.

[49] ZHENG L, SHEN L Y, TIAN L, et al. Scalable person re-identification: A benchmark[C]//2015 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2015: 1116-1124.

[50] WEI L H, ZHANG S L, GAO W, et al. Person transfer GAN to bridge domain gap for person re-identification[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 79-88.

[51] RISTANI E, SOLERA F, ZOU R, et al. Performance Measures and A Aata Set for Multi-target, Multi-Camera Tracking[M]//Computer Vision-ECCV 2016 Workshops. Cham: Springer International Publishing, 2016: 17-35.

[52] KINGMA D P, BA J. Adam: A method for stochastic optimization[C]//International Conference on Learning Representations. California: ICLR, 2015: 1-15.

[53] MIAO Y Q, DENG J K, DING G G, et al. Confidence-guided centroids for unsupervised person re-identification[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 6471-6483.

[54] LUO X, JIANG M, KONG J, et al. Hierarchical camera-aware contrast extension for unsupervised person re-identification[J]. IEEE Transactions on Multimedia, 2024, 26: 7636-7648.

[55] CHENG D, ZHOU J Y, WANG N N, et al. Hybrid dynamic contrast and probability distillation for unsupervised person re-id[J]. IEEE Transactions on Image Processing, 2022, 31: 3334-3346.

[56] CHO Y, KIM W J, HONG S, et al. Part-based pseudo label refinement for unsupervised person re-identification[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 7298-7308.

[57] CHEN H, LAGADEC B, BREMOND F. ICE: Inter-instance contrastive encoding for unsupervised person re-identification[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 14940-14949.

[58] HE Q L, WANG Z H, ZHENG Z J, et al. Spatial and temporal dual-attention for unsupervised person re-identification[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(2): 1953-1965.

[59] YIN J H, ZHANG X Y, MA Z Y, et al. A real-time memory updating strategy for unsupervised person re-identification[J]. IEEE Transactions on Image Processing, 2023, 32: 2309-2321.

[60] TIAN Q, DU X X. Multi-class center dynamic contrastive learning for unsupervised domain adaptation person re-identification[J]. Computers and Electrical Engineering, 2024, 116: 109155.

[61] LEE G, LEE S, KIM D, et al. Camera-driven representation learning for unsupervised domain adaptive person re-identification[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2023: 11419-11428.

[62] LIU Y X, GE H W, SUN L, et al. Complementary attention-driven contrastive learning with hard-sample exploring for unsupervised domain adaptive person re-ID[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(1): 326-341.

[63] LEI J J, QIN T Y, PENG B, et al. Reducing background induced domain shift for adaptive person re-identification[J]. IEEE Transactions on Industrial Informatics, 2023, 19(6): 7377-7388.

[64] LIU C, LI H Q, HE X, et al. Attribute discrimination combined with selected sample dropout for unsupervised domain adaptive person re-identification[J]. Image and Vision Computing, 2024, 146: 105038.

[65] ZHANG W, YE P J, SU T, et al. Sparse-attention augmented domain adaptation for unsupervised person re-identification[J]. Pattern Recognition Letters, 2025, 187: 8-13.

[66] YANG X, DONG W J, ZHENG G, et al. IDENet: An inter-domain equilibrium network for unsupervised cross-domain person re-identification[J]. IEEE Transactions on Image Processing, 2025, 34: 2133-2146.

[67] CHEN F, WANG N, TANG J, et al. Unsupervised person re-identification via multi-domain joint learning[J]. Pattern Recognition, 2023, 138: 109369.

[68] ZHU Y C, ZHENG Y F, LIU J W, et al. Noise-resistance learning via multi-granularity consistency for unsupervised domain adaptive person re-identification[J]. ACM Transactions on Multimedia Computing, Communications and Applications, 2025, 21(1): 1-23.

[69] LIAO S C, SHAO L. Graph sampling based deep metric learning for generalizable person re-identification[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 7349-7358.

[70] HE T, SHEN L Q, GUO Y C, et al. SECRET: Self-consistent pseudo label refinement for unsupervised domain adaptive person re-identification[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(1): 879-887.
[71] LI H F, DONG N, YU Z T, et al. Triple adversarial learning and multi-view imaginative reasoning for unsupervised domain adaptation person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(5): 2814-2830.
[72] ZOU Y, YANG X D, YU Z D, et al. Joint Disentangling and Adaptation for Cross-domain Person Re-identification[M]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 87-104.
[73] VAN DER MAATEN L, HINTON G. Visualizing data using *t*-SNE[J]. Journal of Machine Learning Research, 2008, 9(11): 2579-2605.

作者简介

赵 宇 男,1991年10月出生于重庆市.现为重庆第二师范学院人工智能学院讲师,博士研究生.主要研究方向为计算机视觉、模式识别以及主动健康.
E-mail: zhaoyu@cque.edu.cn

舒巧媛 女,1991年8月出生于四川省自贡市.现为重庆第二师范学院数学与大数据学院讲师,博士研究生.主要研究方向为多目标优化、机器学习以及计算机视觉.
E-mail: shuqy@cque.edu.cn