DOI: 10. 13382/j. jemi. B2306791

细化特征引导对抗性解纠缠学习的 无监督行人重识别*

陈元妹^{1,2} 王凤随^{1,2} 王路遥^{1,2}

(1. 安徽工程大学电气工程学院 芜湖 241000;2. 高端装备先进感知与智能控制教育部重点实验室 芜湖 241000)

摘 要:无监督行人重识别旨在无监督设置下从非重叠的相机中识别出同一行人。针对现有的无监督行人重识别网络不能充分提取行人特征以及相机之间的差异导致行人检索错误的问题,提出了一种细化特征引导对抗性解纠缠学习的无监督行人重 识别方法,设计特征细化信息融合模块嵌入 ResNet50 网络的不同层,用以增强网络提取关键信息的能力。设计特征解耦学习 方法最小化行人特征和相机特征之间的互信息,减少相机差异对网络的负面影响,同时设计对抗性解纠缠损失函数进行无监督 联合学习。在两个公共数据集 Market-1501 和 DukeMTMC-reID 上对所提方法进行评估,平均精度均值分别提升了4.6%、3.1%, 相较于基线算法具备较强的鲁棒性,满足在无监督背景下对行人的识别需求。

关键词:行人重识别;无监督;特征细化;相机差异;对抗性解纠缠

中图分类号: TP391.4 文献标识码: A 国家标准学科分类代码: 510.40

Unsupervised person re-identification of adversarial disentangling learning guided by refined features

Chen Yuanmei^{1,2} Wang Fengsui^{1,2} Wang Luyao^{1,2}

(1. School of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China; 2. Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Ministry of Education, Wuhu 241000, China)

Abstract: Unsupervised person re-identification aims to identify the same person from non-overlapping cameras under unsupervised settings. Aiming at the problem that the existing unsupervised person re-identification network cannot fully extract pedestrian features and the difference between cameras leads to pedestrian retrieval errors, we propose an unsupervised person re-identification of adversarial disentangling learning guided by refined features. A feature refinement information fusion module is designed and embedded into different layers of ResNet50 network to enhance the ability of the network to extract key information. A disentangled feature learning method is designed to minimize the mutual information between pedestrian features and camera features, and reduce the negative impact of camera differences on the network. At the same time, the adversarial disentangling loss function is designed for unsupervised joint learning. Using the Market-1501 and DukeMTMC-reID public datasets, we tested the proposed method. The mean average precision increased by 4.6% and 3.1% respectively. Compared with the baseline algorithm, it has strong robustness and meets the needs of pedestrian recognition in unsupervised background.

Keywords: person re-identification; unsupervised; feature refinement; camera differences; adversarial disentangling

0 引 言

随着计算机视觉领域技术的不断提高以及深度学习 的快速发展,智能安防和智慧城市等应用使得行人重识 别获得越来越多的关注^[1]。行人重识别^[23]是一项通过 给定摄像机中被检测人的图像或视频来识别指定行人的 技术。目前,基于深度学习的有监督行人重识别方法被 广泛应用,且取得出色性能。但有监督行人重识别需要 完整的标签,成本高昂且标注耗时,已经不能满足现实需

收稿日期:2023-07-31 Received Date: 2023-07-31

^{*}基金项目:安徽省自然科学基金(2108085MF197)、安徽高校省级自然科学研究重点项目(KJ2019A0162)、安徽工程大学国家自然科学基金预研项目(Xjky2022040)资助

要。无监督行人重识别弥补了有监督行人重识别技术的 缺陷,它不需要人工标注的身份信息,更适合实际应用。 目前无监督行人重识别方法有两种:完全无监督(fully unsupervised, FU)和无监督域自适应(unsupervised domain adaptation, UDA)行人重识别。其中,UDA方法 需要一个标记源域数据集,且源域数据集和目标域数据 集之间的类间方差难以补偿。因此,FU方法广受关注。 FU方法在训练过程中无标记数据信息的指导,如何构建 高效的无监督行人重识别模型成为研究的难点之一。

目前采用聚类策略^[4-5]的无监督行人重识别方法精 度较高,即通过对未标记的行人图像进行聚类生成伪标 签,再利用伪标签对网络进行训练和优化,通过利用伪标 签生成和网络训练迭代运行的方法提高深度模型的精 度。如Li等^[6]提出了一种聚类引导的非对称对比学习 方法 (cluster-guided asymmetric contrastive learning, CACL),该方法利用聚类结果指导非对称对比学习框架 中的特征进行无监督学习。在 CACL 中,采用实例级和 聚类级对比学习,分别用于帮助孪生网络学习不同数据 增强视图内和视图间关于聚类结果的判别性特征。Ge 等[7]提出了一个统一的对比学习框架以整合源域和目标 域中的可用信息进行联合特征学习。它为源域类级、目 标域簇级和非簇实例级提供监督信号,并动态更新混合 内存。此外还设计了一种自步对比学习策略(self-paced contrastive learning, SPCL),以防止噪声标签引起训练误 差放大。Sridhar 等^[8]提出了一种新的基于时空关联规 则的深度无标签聚类(spatio-temporal association rule based deep annotation-free clustering, STAR-DAC)框架,该 框架基于视觉特征对未标注的行人图像进行增量聚类, 并通过挖掘的时空关联规则进行聚类微调。然而,上述 方法在聚类时将每个聚类视为一个伪身份类,忽略了聚 类过程中会产生噪声标签,且存在相机差异导致不同相 机的类内样本被分配到不同的伪标签上。2022年, Chen 等^[9]提出了一种基于集群记忆的元学习(cluster memorybased meta learning, CMML)策略,该策略采用基于集群 记忆的加性边际损失来处理由相机变化引起的噪声标签 和模型训练问题,以及一种基于集群记忆的抗噪声损失 来处理噪声标签,进一步提升了模型训练精度。Cao 等^[10]提出了一种语义特征提取-相机感知重分配 (semantic feature extraction-camera-aware re-allocation, SFE-CR)框架,通过解析模型提取训练样本的语义局部 特征,消除冗余信息。此外,使用相机 ID 对样本进行分 割,并重新分配由通用聚类方法生成的伪标签。

在无监督 Re-ID 中,由于真实环境中存在相机视角 变化、行人遮挡等因素的影响,导致模型识别精度不高。 因此,Xuan 等^[11]提出了一种相机内和相机间的相似度 计算(intra-inter camera similarity, IICS)方法,通过生成相 机内和相机间的伪标签来迭代优化相机内相似度,有效 降低相机之间的分布差异,获得更鲁棒的特征表示。 Chen等^[12]提出了一种联合生成和对比学习框架的无监 督行人重识别方法,通过使用三维网格引导的生成式对 抗网络(generative adversarial network, GAN)生成不同视 图,然后将这些生成的新视图与基于记忆的对比学习中 的原始图像相结合,以学习视图不变表示,从而提高生成 质量。Hui等^[13]提出了一种基于注意力块和精细化聚类 的无监督跨域行人重识别方法。将坐标注意力和三元组 注意力串联嵌入到骨干网络中挖掘细粒度信息,并提出 一种精细化聚类策略来剔除噪声样本,实现由粗到精的 聚类过程。

上述方法虽然在一定程度上提升了无监督行人重识 别性能,但仍然存在以下几个问题:1)采集的行人图像中 所含特征信息较少,现有算法难以精确提取有效的特征 信息;2)由于相机之间存在差异,网络提取的特征信息中 可能包含一些无关特征,现有算法未考虑这些无关特征 对网络的负面影响,从而导致模型的识别准确率不高。

针对上述问题,本文从提高网络特征细化能力以及 降低相机分布差异,减少噪声信息,提高识别精度出发, 提出了一种细化特征引导对抗性解纠缠学习的无监督行 人重识别方法。在完全无标记信息的情况下,主要贡献 如下:

1)设计特征细化信息融合模块(feature refinement information fusion module, FRIFM),自适应定位到行人图像中的潜在区域,增强网络对关键信息的细化学习,提高特征提取能力。

2)设计了特征解耦学习方法对行人特征和相机特征 进行解耦,减少无关特征对模型训练的影响;同时设计对 抗性 解 纠 缠 损 失 函 数 (adversarial disentangling loss, ADL)进行无监督联合学习,有效解决相机差异导致行人 检索错误的问题,提升模型的识别精度。

3) 在行人重识别的两个公共数据集上与近年多种主 流模型对比,并在 Market-1501 数据集上进行可视化实 验,实验结果表明本文所改进的方法能够有效提升无监 督行人重识别的识别准确率。

1 本文方法

1.1 网络整体框架

本文提出了一种细化特征引导对抗性解纠缠学习的 无监督行人重识别方法,网络总体框架如图 1 所示。首 先,本文使用嵌有特征细化信息融合模块的 ResNet50 提 取未标记行人样本的全局特征。其次,利用提取的样本 特征分别在相机内与相机间训练阶段进行聚类,为所有 训练实例分配伪标签,并计算相应的相机内损失和相机 间损失。同时,在相机内训练阶段,将聚类后的特征 F₁ 作为特征提取器 E 的输入,并获得相机特征 F₂ = z^c。然 后将相机特征 z^c 与行人特征 z^p 的联合分布 P 和边际分 布 Q 作为判别器 D 的输入,利用判别器 D 估计两个分布 的 JS 散度,衡量两种分布的相似性。最终,通过最小化 两种分布之间的 JS 散度获得对抗性解纠缠损失函数。 在整个训练过程中,相机内训练阶段和相机间训练阶段 交替执行,以此优化 CNN 网络提取的特征。



图 1 整体网络结构 Fig. 1 Overall network structure diagram

1.2 FRIFM

通常 Re-ID 模型的性能会受到行人遮挡或行人姿态 等问题的影响,为缓解此类问题,本文设计了 FRIFM 并 嵌入 ResNet50 网络的不同层定位行人图像的重要区域, 消除背景噪声的干扰。

FRIFM 从空间结构和通道维度重新加权特征映射中的每个元素,加权后的高层特征富含行人语义信息,这些信息能够帮助网络模型充分提取行人图像中的判别性特征,同时有利于后续对抗性解纠缠学习的开展。 FRIFM 结构图如图 2 所示,具体工作如下:将 FRIFM 的输入特征记为 $F \in R^{H \times W \times C}$,输出特征记为 $F_2 \in R^{H \times W \times C}$,其中,H、W、C分别表示特征图的高度、宽度以及通道的大小。

首先,将输入特征图分别在宽度和高度两个维度进行平均池化操作,并用 F_w 和 F_H 表示,其中, $F_w \in R^{C\times1\times W}$, $F_H \in R^{C\times1\times H}$ 。其次,为提升空间特征表示能力,将 F_w 和 F_H 融合,然后分别经过 Conv1×1 卷积层和 ReLU 激活函数层进行不同维度的压缩。再次,为获取下一步的中间特征图,将融合的特征进行切割,并依次经过



Fig. 2 FRIFM module structure diagram

堆叠的 Conv1×1 卷积层和 Sigmoid 激活函数层进行不同 维度的扩展,捕获相应位置信息,用 F'_w 和 F'_H 表示,其 中, $F'_w \in R^{C\times1\times W}$, $F'_H \in R^{C\times H\times 1}$ 。 F'_w 或 F'_H 中的每个元素表 示特征图 F 中相应列或行的局部特征对识别行人的重要 性。最后,通过用 F'_w 和 F'_H 重新加权 F,得到中间特征图 F_1 。计算公式如下:

$$F_1 = F \otimes F'_W \otimes F'_H \tag{1}$$

其中,⊗表示元素级乘法。

为进一步细化行人特征,将特征 F₁ 在通道维度进行 平均池化,然后将池化后的特征依次经过 Conv1d 卷积层 和 Sigmoid 激活函数层进行细化学习,最后用细化特征重 新加权 F_1 ,得到特征 F_2 。计算公式如下:

*F*₂ = (Sigmoid(*C*1*D*(GAP(*F*₁)))) ⊗ *F*₁ (2)
 其中,*C*1*D* 表示通过一维卷积过程,GAP 表示全局
 平均池化。

1.3 损失函数

1) 在相机间训练阶段, 通过采用域自适应策略训练 分类器计算相机间相似度, 然后对其进行聚类。在聚类 结果基础之上, 进一步优化带有相机间损失的 *f*, 计算公 式如下:

$$L_{\text{inter}} = \frac{1}{|B|} \sum_{I_n \in B} l(F(w, f_n), m) + \alpha L_{\text{triplet}}$$
(3)

其中, *B* 是一个训练小批, *I*_n 是由相机捕获的行人图像。*l*(•)计算软最大交叉熵损失, *m* 是由聚类结果分配的伪标签, α 是损失权重, *L*_{triplet} 是硬批三元组损失。本文从每个聚类中随机选择 *J* 个聚类和 *K* 个样本来构建训练小批 *B*。

2) 在相机内训练阶段, 对每个相机内特征进行聚类 生成伪标签, 然后根据聚类结果训练 f。假设第 c 个摄像 机的聚类结果为 T^{*}, 第 c 个摄像机的训练损失可以表 示为:

$$L_{\text{intra}}^{c} = \sum_{I_{n} \in X^{c}, I_{n} \in T_{m}^{c}} loss^{c}(f_{n}, m)$$
(4)

其中,X^e 是行人图像的集合。为保证f在不同摄像 机下对复杂的行人姿态、照明等的鲁棒性,式(4)通过共 享相同的f在不同相机上计算。

通过聚类步骤,为每个集群分配具有相同标签的图像,从而将每个 X^e 变成一个有标记的数据集,则式(4)中的函数 *loss*^e(•)可以表示为:

$$loss^{c}(f_{n},m) = l(F(w^{c},f_{n}),m)$$
(5)

其中, $F(w^{e}, \cdot)$ 表示一个具有可学习参数 w^{e} 的分类器。 $l(\cdot)$ 计算分类器输出和标签 m 的 softmax 交叉熵损失。因此,相机内整体训练损失可表示为:

$$L_{\rm intra} = \sum_{c=1}^{C} L_{\rm intra}^c \tag{6}$$

其中,*C*为相机的总数。最终,将相机内和相机间总 损失定义为:

$$L_{\rm v} = L_{\rm inter} + L_{\rm intra} \tag{7}$$

3)考虑到因不同相机之间存在差异,网络提取的特征中通常包含一些无关信息,从而导致行人检索出现错误,影响模型性能。本文通过最小化行人特征和相机特征之间的互信息进行对抗性解纠缠学习,对抗性解纠缠损失定义如下:

$$L_{\text{ADL}} = \min_{P \in \mathcal{D}} \left(D_{\text{JS}}(P \mid \mid Q) \right)$$
(8)

$$D_{JS}(P \mid \mid Q) = \sup_{D:\Omega \to R} (E_{x \sim P} [\log\sigma(I(x))] + E_{x \sim Q} [\log(1 - \sigma(I(x)))])$$
(9)

其中, σ 是 Sigmoid 型激活函数,I是一个映射函数。 E 为特征提取器,用于提取相机特征。使 $\mu = E_{x-p}[\log\sigma(I(x))] + E_{x-q}[\log(1-\sigma(I(x)))],则式(8)表明 JS 散度被转换为寻找<math>\mu$ 的上界。 L_{ADL} 可以表示为:

$$L_{\text{ADL}} = \min_{\mu \in \mathcal{D}} \max_{\mu}(\mu) \tag{10}$$

为了优化网络训练,本文将相机间与相机内总损失 L_v 和对抗性解纠缠损失 L_{ADL} 联合监督网络学习,定义为 最终的损失 L_o 其中, λ 用于平衡两项损失。

$$L = L_{\rm v} + \lambda L_{\rm ADL} \tag{11}$$

2 实验结果与分析

2.1 实验数据集与评价指标

本文在 Market-1501 和 DukeMTMC-reID 两个大规模 的数据集上进行一系列实验,并通过实验结果验证所提 方法的有效性。表 1 分别展示了两个数据集中的训练 集、测试集、查询集的图像数。

Market-1501 数据集是从 6 个不相交的摄像机中捕获图像,包含 1 501 个身份,训练集行人数量为 751,共有 12 936 张图像;测试集行人数量为 750,共有 19 732 张图像。

DukeMTMC-reID 数据集是从 8 个不相交的摄像机中 捕获图像,包含 1 404 个身份,训练集行人数量为 702,共 有 16 522 张图像;测试集行人数量为 702,共有 17 661 张 图像。

表 1	不同数据集中的训练集、测试集和查询集的数据统计

Table 1	Data statistics of training set,	test set and query set in different	t data sets
---------	----------------------------------	-------------------------------------	-------------

数据集	相机个数	训练集行人数	训练集图像数	测试集行人数	测试集图像数	Query 图像数	图像总数
Market-1501	6	751	12 936	750	19 732	3 368	32 668
DukeMTMC-reID	8	702	16 522	702	17 661	2 228	36 441

本文采用平均精度均值(mean average precision, mAP)和累积匹配特性曲线(cumulative matching

characteristics, CMC)评估模型性能。mAP反应检索的 行人在数据库中所有正确图片排在序列表前面的程度,

(%)

计算过程如式(12)所示。其中,AP 表示每个类别的平均 精度,C 表示类别数,每次实验采用随机测试集重复5次 以获取平均性能。对于 CMC 曲线,本文通过 Rank-1, Rank-5, Rank-10 的匹配精度给出,这意味着查询图片在 前 n 个列表中有匹配值。

$$mAP = \frac{\sum_{i=1}^{L} AP_i}{C}$$
(12)

2.2 实验设置

本文实验选用的 GPU 为 NVIDIAGeForce RTX 2080Ti(11 GiB),操作系统为 Ubuntu 16.04,处理器为英特尔 Core i9-10900@3.7 GiHz,深度学习框架为 Fytorch 1.2.0。本文采用 ImageNet 预训练的 ResNet50 作基础网络。删除池化层之后的所有层,并添加一个 BN 层。在测试和聚类过程中,本文利用池化层后的特征计算相似

度。在训练过程中,输入图像的大小为 256×128。采用 了随机翻转和随机擦除等图像增强策略。迭代训练过程 中,相机内和相机间阶段按顺序执行。epoch 大小设置为 40。在相机内训练阶段,每个摄像机的 batchsize 大小为 8。ResNet50 基层的学习率为 0.000 5,其他基层的学习 率为 0.005。在相机间训练阶段,batchsize 大小设置为 64。ResNet50 基层的学习率为 0.001,其他层的学习率 为 0.01。超参数 α=1,J=16,K=4。

2.3 与其他方法的对比实验

为验证所提方法的有效性,本节在 Market-1501 和 DukeMTMC-reID 两个大数据集上将提出的方法与近3年 无监督行人重识别方法进行对比,对比的方法包括完全 无监督方法与无监督域自适应方法。对比结果如表2所 示。其中"—"表示原论文中没有报告的数据。

表 2	与其他方法的对比实验结果

Table 2E	xperimental	results	compared	with	other	methods
----------	-------------	---------	----------	------	-------	---------

方法	佐田樹	Market-1501			佐田斌	DukeMTMC-reID				
	作用或	mAP	Rank-1	Rank-5	Rank-10	TF用	mAP	Rank-1	Rank-5	Rank-10
STAR-DAC ^[8]	None	33.9	67.0	80.6	84.9	None	31.6	56.4	72.1	76.5
GSRA ^[14]	None	41.7	68.7	84.2	88.4	None	33.5	58.3	71.0	75.8
MV-ReID ^[15]	None	45.6	73.3	85.3	89.1	None	31.7	54.5	67.5	72.1
DLCL ^[16]	None	48.3	73.8	84.7	89.2	None	55.4	74.3	48.3	86.5
ACAN-GRL ^[17]	None	50.6	73.3	87.6	91.8	None	46.6	65.1	80.6	85.1
SDAAL ^[18]	None	56.7	82.6	91.7	94.7	None	52.3	72.8	82.5	86.1
CVSE ^[19]	Duke	63.2	84.1	92.8	95.0	Market	56.1	75.3	82.9	85.4
SFE-CR ^[10]	None	65.4	83.1	92.5	95.3	None	61.8	77.7	87.2	90. 7
MLC ^[20]	None	66.2	86.7	93.5	95.6	None	52.3	73.6	82.3	85.5
IICS ^[11]	None	72.9	89.5	95.2	97.0	None	64.4	80.0	89.0	91.6
SpCL ^[7]	None	73.1	88.1	95.1	97.0	None	_	_	_	_
GCL+(JVTC) ^[21]	Duke	74.4	89.7	95.5	96.7	Market	61.3	78.0	86.8	89.1
MPC ^[22]	None	77.4	90.9	96.4	97.6	None	63.6	79.2	87.4	90.3
Ours	None	77.5	91.8	97.0	97.9	None	67.5	82.1	89.9	92.3

从表 2 可以看出,在完全无监督设置下,本文方法在 Market-1501数据集上的mAP为77.5%,Rank-1为 91.8%。在DukeMTMC-reID数据集上的mAP为67.5%, Rank-1为82.1%,与其他方法相比,本文提出的方法均 取得了最好的结果。具体来说,与SpCL^[7],SFE-CR^[10]等 聚类方法相比,在Market-1501数据集中,其mAP分别提 高了4.4%、12.1%,Rank-1分别提高了3.7%、8.7%,充 分验证了本文提出的方法性能明显更加优越。与无监督 域自适应方法相比,如CVSE^[19],GCL+(JVTC)^[21],本文 使用完全无监督方法,不需要多余标记的源域数据集,且 实验效果相对较好,从而说明了本文提出的方法在现实 场景具有一定的鲁棒性。与完全无监督方法相比,如 GSRA^[14]、DLCL^[16]、MLC^[20]、IICS^[10]等,本文通过特征细 化信息融合模块细化学习关键区域特征,提高了模型提 取关键信息的能力,同时采用特征解耦学习方法减少无 关特征的影响,提升模型的识别精度。与最新方法相比, 如 DLCL^[16]、GCL+(JVTC)^[21]、MPC^[22],在 Market-1501 数据集中,其 mAP 分别提高了 28.2%、3.1%、0.1%, Rank-1分别提高了 18.0%、2.1%、0.9%,进一步验证了 本文所提方法的有效性和先进性。

2.4 实验结果与分析

1)平均精准度对比实验

本文通过设置平均精准度对比实验来验证所提方法 的有效性。在训练阶段保持同一实验设置的前提下,将 本文方法与基线方法在两大数据集中进行对比,其性能 对比结果如表3所示。从两个数据集中的性能对比数据 可以得出,相对于基线方法,本文所提出方法的 mAP 值 分别提升了4.6%、3.1%, Rank-1 值分别提升了2.3%、

(%)

2.1%。表明在无监督行人重识别网络中,本文所设计的 特征细化信息融合模块和特征解耦学习方法有效地提高 了网络模型的鲁棒性。

表 3 在不同数据集中的精准度对比实验结果

		Table 3 A	ccuracy compa	rison experin	nental results i	n different da	tasets	((%)
方法 ——		Marke	et-1501		DukeMTMC-reID				
	Rank-1	Rank-5	Rank-10	mAP	Rank-1	Rank-5	Rank-10	mAP	
基线	89.5	95.2	97.0	72.9	80.0	89.0	91.6	64.4	
木文	01.8	97.0	07 0	77 5	82 1	80.0	02.3	67.5	

2) 消融实验

为了验证特征细化信息融合模块与对抗性解纠缠损 失函数对模型性能的影响,通过设定4种不同的方式在 Market-1501数据集下开展消融实验,每组的实验设置均 与基线保持相同。实验结果如表4所示。

首先,对基线网络进行测试,如表4索引1所示,其 mAP 达到 72.9%, Rank-1 为 89.5%。其次, 本文对基线 网络分别添加特征细化信息融合模块与对抗性解纠缠损 失函数进行实验,实验结果如索引2和3所示。在仅添 加特征细化信息融合模块的情况下,mAP 达到 75.1%, Rank-1 达到 90.9%,相比于基线网络分别提升了 2.2%、 1.4%。这表明,特征细化信息融合模块增强了模型对融 合多语义信息的捕获能力,丰富了细腻特征描述,充分验 证了所提改进之处的有效性。在仅添加对抗性解纠缠损 失函数的情况下,其首位击中率和平均精度均值都有所 提升, mAP 达到 74.6%, Rank-1 达到 90.5%。上述提升 表明,对抗性解纠缠损失函数可以有效减少训练过程中 由于相机之间的差异导致网络生成的冗杂信息。最后, 为进一步分析所提出方法联合作用的有效性,本文对同 时添加特征细化信息融合模块与对抗性解纠缠损失函数 进行一次实验,如索引4所示。其mAP达到了77.5%, 与基线网络相比提升了 4.6%; Rank-1 达到了 91.8%, 在 基线网络基础上提升了 2.3%。实验结果体现了本文所

提出改进的有效性。且索引4数据相较于索引2和3都 有所提升,充分验证了将特征细化信息融合模块和对抗 性解纠缠损失函数联合作用的鲁棒性,有效提升了无监 督行人重识别方法的精度。

表 4 消融实验结果										
	Та	able 4	Results	riments	(%)					
Index	Base	FRIFM	ADL	Rank-1	Rank-5	Rank-10	mAP			
1	\checkmark			89.5	95.2	97.0	72.9			
2	\checkmark			90.9	96.9	97.7	75.1			
3	\checkmark		\checkmark	90.5	96.6	97.5	74.6			
4	\checkmark	\checkmark	\checkmark	91.8	97.0	97.9	77.5			

3)参数 λ 分析

为探究本文整体损失函数中参数 λ 的取值对模型性能的影响,选取了 $\lambda = 0.1$ 、0.3、0.5、0.7、0.9 这 5 个不同的数值进行实验,实验所使用的数据集为 Market-1501 和 DuleMTMC-reID 数据集,实验结果如表 5 所示。随着 λ 数值的增加, Market-1501 和 DuleMTMC-reID 数据集的 mAP 和 Rank-1 均呈现先升高后降低的趋势,当 λ 取值 为 0.5 时,两个数据集的训练结果均呈现最佳效果,这表明在两个数据集上,参数 λ 对模型性能的影响是一致的。因此本文中的 λ 取值 0.5。

表 5 参数 λ 在不同数据集中的实验结果

λ –		Marke	et-1501		DukeMTMC-reID				
	Rank-1	Rank-5	Rank-10	mAP	Rank-1	Rank-5	Rank-10	mAP	
0.1	91.2	96.4	96. 9	76.4	81.6	89.1	91.6	66.8	
0.3	91.4	97.0	97.4	77.0	82.0	89.8	92.1	67.1	
0.5	91.8	97.0	97.9	77.5	82.1	89. 9	92. 3	67.5	
0.7	91.5	96.5	97.4	77.1	81.7	89.6	92.0	67.3	
0.9	91.2	96.5	97.3	76.9	81.3	89.3	91.9	67.0	

Table 5 Experimental results of parameter λ in different datasets

2.5 可视化实验结果分析

为直观分析所提出方法的重识别效果,本文在 Market-1501数据集上随机选取3位行人进行可视化实 验,实验结果如图3所示。第1列为query图像,代表随 机挑选的待查询行人图像,右侧标有数字1~10为行人 重识别的检索结果,分别表示为 Rank-1~Rank-10。每张 图像上方的数字颜色代表该检索结果是否正确,数字中 带框的表示错误,不带框的表示正确。

如图 3(a)和(b)组行人图像所示,行人特征较为单一,在 baseline 方法中出现 8 个错误的识别结果。而本文

方法只出现两个错误的识别结果,充分验证了本文方法 对潜在特征的提取能力以及损失约束的效果。同时相比 于 baseline,本文方法检索到正确行人的位置更加靠前, 表明了本文方法的有效性。当行人特征较为复杂时,如 图 3(c)、(d)组和(e)、(f)组实验结果图所示,行人存在 手持雨伞和挎包状态。在图 3(c)组行人图像中,baseline 方法中错误识别了其他未手持雨伞的行人,出现了第 3、 4、5、7 位行人的错误识别结果,而本文方法只出现一个

错误的识别结果,且出现在第8位。在图3(e)组行人图像中,baseline 方法错误地识别了第7、9位相似挎包行人,而本文方法能够在复杂的行人表征中获取具有鉴别性的语义信息,并通过对抗性解纠缠损失函数减少识别错误,检索出正确的10位行人。上述实验结果证明了所提方法的鲁棒性,并表明其可以显著提高无监督行人重识别的准确率。



图 3 在 Market-1501 数据集上的可视化实验结果

Fig. 3 The visual experimental result diagram on the Market-1501 dataset

3 结 论

通过分析现有无监督行人重识别方法存在骨干网络 提取行人表征能力不足,训练过程中存在冗杂信息导致 行人识别准确率较低的问题,本文提出了一种细化特征 引导对抗性解纠缠学习的无监督行人重识别方法,该方 法采用多任务联合学习,从而提高模型的重识别精度。 此外,本文利用特征细化信息融合模块强化网络对细化 特征的激活响应,提升了网络定位关键信息的能力。同 时利用对抗性解纠缠损失和相机损失联合监督网络学 习,有效减少了训练过程中的冗杂信息,提高了行人识别 准确率。最后在两个公共数据集上对本文所提方法进行 训练和评估,实验结果显示识别精度有效提升,充分验证 了本文方法的有效性和可靠性。然而,在复杂多变的现 实监控场景中,当出现着装或姿态等相似的不同行人时, 本文的特征细化信息融合模块在辅助网络削弱无标记数 据间特征差异的同时,可能也会削弱相似行人图像中的 判别性特征。因此,下一步工作将考虑在本文的基础上 构建全局和局部特征融合网络,捕获具有多粒度的特征 描述,以此丰富行人判别特征的层次性,进一步增强无监 督模型鲁棒性。

参考文献

 [1] 张勃兴,马敬奇,张寿明,等.利用全局与局部关联 特征的行人重识别方法[J].电子测量与仪器学报, 2022,36(6):205-212.

> ZHANG B X, MA J Q, ZHANG SH M, et al. Person reidentification method based on global and local relation features [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(6): 205-212.

 [2] 钱亚萍,王凤随,熊磊.基于局部细化多分支与全局 特征共享的无监督行人重识别方法[J].电子测量与 仪器学报,2023,37(1):106-115.

> QIAN Y P, WANG F S, XIONG L. Unsupervised person re-identification method based on local refinement multibranch and global feature sharing [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 106-115.

- [3] LIU L, LI X, LEI X M. A brief survey for person reidentification based on deep learning [J]. International Journal of Computer Applications in Technology, 2022, 69(2): 101-111.
- [4] NING M N, ZENG K W, GUO Y, et al. Deviation based clustering for unsupervised person re-identification [J].
 Pattern Recognition Letters, 2020, 135: 237-243.
- [5] LISH, YUANMK, CHENJ, et al. AdaDC: Adaptive deep clustering for unsupervised domain adaptation in

person re-identification [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(6): 3825-3838.

- [6] LI M K, LI C G, GUO J. Cluster-guided asymmetric contrastive learning for unsupervised person reidentification [J]. IEEE Transactions on Image Processing, 2022, 31: 3606-3617.
- GE Y X, ZHU F, CHEN D P, et al. Self-paced contrastive learning with hybrid memory for domain adaptive object re-id [C]. Proceedings of the 34th Conference on Neural Information Processing Systems, 2020: 11309-11321.
- [8] SRIDHAR R S, PRASAD M V N K, RAMADOSS B. Spatio-temporal association rule based deep annotationfree clustering (STAR-DAC) for unsupervised person reidentification [J]. Pattern Recognition, 2022, 122: 108287.
- [9] CHEN Y Y, FAN Z Y, CHEN S N. Consistent camerainvariant and noise-tolerant learning for unsupervised person re-identification [J]. Image and Vision Computing, 2022, 123: 104462.
- [10] CAO G, JO K. Semantic foreground feature extraction and camera-aware re-allocation clustering for unsupervised person re-identification [C]. International Conference on Control, Automation and Systems (ICCAS), 2022: 168-173.
- 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), 2021: 11921-11930.
- [12] CHEN H, WANG Y H, BENOIT L, et al. Joint generative and contrastive learning for unsupervised person re-identification [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 2004-2013.
- [13] HUI Y, WU X, HU X H, et al. Unsupervised crossdomain person re-Identification method based on attention block and refined clustering [J]. IEEE Access, 2022, 10: 105930-105941.
- [14] SRIDHAR R S, PRASAD M V N K, RAMADOSS B. Generative segment-pose representation based augmentation (GSRA) for unsupervised person reidentification [J]. Image and Vision Computing, 2023, 131: 104632.
- [15] YIN Q Z, WANG G A, DING G D, et al. Multi-view label prediction for unsupervised learning person reidentification[J]. IEEE Signal Processing Letters, 2021, 28: 1390-1394.

- [16] ZHAO Y, SHU Q Y, SHI X. Dual-level contrastive learning for unsupervised person re-identification [J]. Image and Vision Computing, 2023, 129: 104607.
- [17] QI L, WANG L, HUO J, et al. Adversarial camera alignment network for unsupervised cross-camera person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32 (5): 2921-2936.
- [18] XU S M, LUO L K, HU J L, et al. Semantic driven attention network with attribute learning for unsupervised person re-identification [J]. Knowledge-Based Systems, 2022, 252: 109354.
- [19] ZHOU S R, WANG Y, ZHANG F, et al. Cross-view similarity exploration for unsupervised cross-domain person re-identification [J]. Neural Computing and Applications, 2021, 33(9): 4001-4011.
- [20] LI Q, PENG X J, QIAO Y, et al. Unsupervised person re-identification with multi-label learning guided selfpaced clustering [J]. Pattern Recognition, 2022, 125: 108521.
- [21] CHEN H, WANG Y H, BENOIT L, 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.

[22] LI X B, LI Q Y, LIANG F J, et al. Multi-granularity pseudo-label collaboration for unsupervised person reidentification [J]. Computer Vision and Image Understanding, 2023, 227: 103616.

作者简介





陈元妹,硕士研究生,主要研究方向为 计算机视觉、行人重识别。 E-mail: 1026394852@gq.com

Chen Yuanmei is a M. Sc. candidate. Her main research interests include computer vision and person re-identification.

王凤随(通信作者),博士,教授,主要 研究方向为图像与视频信息处理、视觉计算 与智能分析。

E-mail: fswang@ahpu.edu.cn

Wang Fengsui (Corresponding author) is a Ph. D., professor. His main research

interests include image and video information processing, visual computing and intelligent analysis.



王路遥,硕士研究生,主要研究方向为 计算机视觉、行人重识别。 E-mail:578651059@qq.com

Wang Luyao is a M. Sc. candidate. Her main research interests include computer vision

and person re-identification.