

软伪标签和多尺度特征融合的行人重识别

陈昊¹, 张宝华^{1,3*}, 吕晓琪^{2,3}, 谷宇^{1,3}, 王月明^{1,3}, 刘新^{1,3}, 任彦¹, 李建军^{1,3}, 张明^{1,3}

¹内蒙古科技大学信息工程学院, 内蒙古自治区 包头 014010;

²内蒙古工业大学信息工程学院, 内蒙古自治区 呼和浩特 010051;

³内蒙古自治区模式识别与智能图像处理重点实验室, 内蒙古自治区 包头 014010

摘要 传统无监督域自适应行人重识别算法,抑制伪标签噪声效果差、域间泛化能力弱。针对这些问题,提出了一种基于软伪标签和多尺度特征融合的无监督域自适应行人重识别算法。为抑制伪标签噪声,利用并行网络的预测值作为软伪标签,通过交叉校对方式对伪标签噪声进行纠偏,为无监督域自适应任务提供更鲁棒的软伪标签。为增强域间泛化能力,利用多尺度特征重构和哈达玛积特征融合方法对深浅特征层信息进行处理,实现源域数据到目标域的风格转换,并结合实例和批量归一化网络解决残差网络域自适应性差的问题,增强网络对源域和目标域的泛化能力。实验结果表明,所提算法在 Market-to-Duke 和 Duke-to-Market 无监督域自适应任务中都取得了较好的性能,明显优于相关算法。

关键词 光计算; 软伪标签; 多尺度特征重构; 哈达玛积特征融合; 实例和批量归一化网络; 行人重识别

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2420001

Soft Pseudo-Label and Multi-Scale Feature Fusion for Person Re-Identification

Chen Hao¹, Zhang Baohua^{1,3*}, Lü Xiaoqi^{2,3}, Gu Yu^{1,3}, Wang Yueming^{1,3}, Liu Xin^{1,3}, Ren Yan¹,
Li Jianjun^{1,3}, Zhang Ming^{1,3}

¹School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010,
Inner Mongolia, China;

²School of Information Engineering, Mongolia Industrial University, Huhehaote 010051, Inner Mongolia, China;

³Inner Mongolia Key Laboratory of Pattern Recognition and Intelligent Image Processing, Baotou 014010,
Inner Mongolia, China

Abstract The traditional unsupervised domain adaptive person re-identification algorithm suppressed the noise of pseudo-label poorly and lack inter-domain generalization ability. For the above problems, an unsupervised domain adaptive person re-identification algorithm was proposed which based on soft pseudo-label and multi-scale feature reconstruction. In order to suppress pseudo-label noise, the predicted value of the parallel network is used as the soft tag, and pseudo-label noise is corrected by cross-proofreading methods, which provides a more robust soft false tag for unsupervised domain adaptive tasks. In order to enhance the generalization ability between domains, multi-scale feature reconstruction and Hadamard product feature fusion methods are used to process the deep and shallow feature layer information, realize the style conversion from source domain data to target domain, and solve the problem of poor adaptability of residual network domain with instance normalization and batch normalization network, so as to enhance the generalization ability of the network to source domain and target domain. Experimental results show that the proposed algorithm has achieved good performance in both Market to Duke and Duke to Market unsupervised domain adaptive tasks, which is significantly better than the related algorithms.

Key words optics in computing; soft pseudo-label; multi-scale feature reconstruction; Hadamard product feature fusion; instance normalization and batch normalization net; person re-identification

收稿日期: 2021-09-17; 修回日期: 2021-10-18; 录用日期: 2021-10-29

基金项目: 国家自然科学基金(61962046, 62001255, 61841204)、内蒙古杰青培育项目(2018JQ02)、内蒙古科技计划项目(2020GG0315, 2021GG0082)、中央引导地方科技发展资金项目(2021ZY0004)、内蒙古草原英才、内蒙古自治区自然科学基金(2019MS06003, 2018MS06018)、教育部“春晖计划”合作科研项目(教外司留 1383号)、内蒙古自治区高等学校科学技术研究项目(NJZY145)

通信作者: *zbh_wj2004@imust.cn

1 引言

行人重识别(ReID)的主要任务是在跨摄像头场景下跟踪特定行人,即在摄像头拍摄的图像中搜索目标行人并在其他摄像头拍摄的图像中搜索该行人,此技术能极大提高工作效率,被广泛运用于智能安保、智能监控和失踪人群搜寻等领域^[1]。

基于深度学习的行人重识别算法主要分为有监督学习和无监督学习。有监督算法存在高相似困难样本对和跨区域视角变化等问题,导致样本标注精度降低,训练模型的可拓展性差,限制了其应用^[2]。而无监督重识别算法很好解决了这一问题,但无监督学习中的数据无标注,跨摄像头场景中图像与标签无映射关系,导致模型识别精度低。且在实际应用中训练后的无监督行人重识别模型直接应用于新场景时,域间差异会导致算法精度明显下降。此后,无监督域自适应任务逐渐成为研究热点,更贴近应用需求,但也更具挑战性^[3]。

无监督域自适应重识别领域中,主流方法有基于聚类的伪标签法和域转换法等。前者通过聚类算法(例如K-means算法)对无标签的目标域图像特征聚类生成伪标签,再用该伪标签监督网络在目标域中训练,以上过程不断循环直至模型收敛^[4]。Ge等^[5]提出了一种新的行人重识别框架,该框架可以动态生成源域类级、目标域类级和非聚类实例级的监督信号用于学习有效特征,利用多种形式的类别原型提供混合监督,以充分挖掘所有训练数据内在分布。但是挖掘目标域数据的方法,并不能从根本上解决域间差异的问题。

目前,大量文献已经证明了基于聚类的伪标签算法的有效性,但模型的训练被伪标签噪声所干扰,也无法达到近似最优解。鉴于伪标签噪声主要源于目标域类别数的不确定性和聚类算法本身的局限性等,如何降低伪标签噪声对网络性能的影响成为目前亟需解决的问题。Deng等^[6]提出了一种基于风格迁移模型(SPGAN)的域转换法,该算法可以实现不同数据集图

像风格的迁移,由循环生成对抗网络(CycleGAN)和孪生网络(SiaNet)组成,前者引入目标域身份损失函数辅助生成器生成目标域风格的行人图像,后者引入一致性损失函数,保证图像转换前后的身份一致性^[7]。但是SPGAN模型的生成器提取的特征尺度单一,转换过程会导致部分特征信息的丢失,并不能很好地迁移目标域的风格特征。

针对上述算法中由于域间差异导致模型迁移效果差、伪标签噪声影响模型训练、生成器特征提取尺度单一导致部分信息丢失的问题,本文提出了一种基于软伪标签和多尺度特征融合的无监督域自适应行人重识别方法。首先,采用软伪标签纠偏法降低伪标签噪声对网络训练性能的影响,解决聚类算法中硬量化损失问题;其次,进行多尺度特征重构和哈达玛积(HP)特征融合改进,融合深浅特征层多语义信息,解决特征尺度转换单一导致部分特征信息丢失的问题;最后,将源域数据转换为目标域风格,并与实例和批量归一化网络(IBN-Net)进行融合,从域间间隔和网络本身提升模型的域间泛化能力。

2 基于软伪标签和多尺度特征融合的行人重识别算法

2.1 所提算法模型

所提基于改进SPGAN的软伪标签纠偏法模型如图1所示,模型由改进SPGAN模型和两组师生网络组成(用数字1,2区分)。首先将有标签数据源域图像通过改进SPGAN模型转换为目标域风格输入网络进行预训练,得到更适合目标域的模型。随后将无标签目标域图像传入两组网络进行正式训练,并对输入的样本进行不同的随机数据增强。两组学生网络通过聚类算法(DBSCAN)生成含有噪声的伪标签,利用带有伪标签的特征训练模型,协同训练两组教师网络预测值和的1/2作为软伪标签,对方学生网络产生的伪标签噪声进行纠偏,随着网络不断更新,预测值误差不断减小,形成正反馈效应,直至模型收敛^[8]。

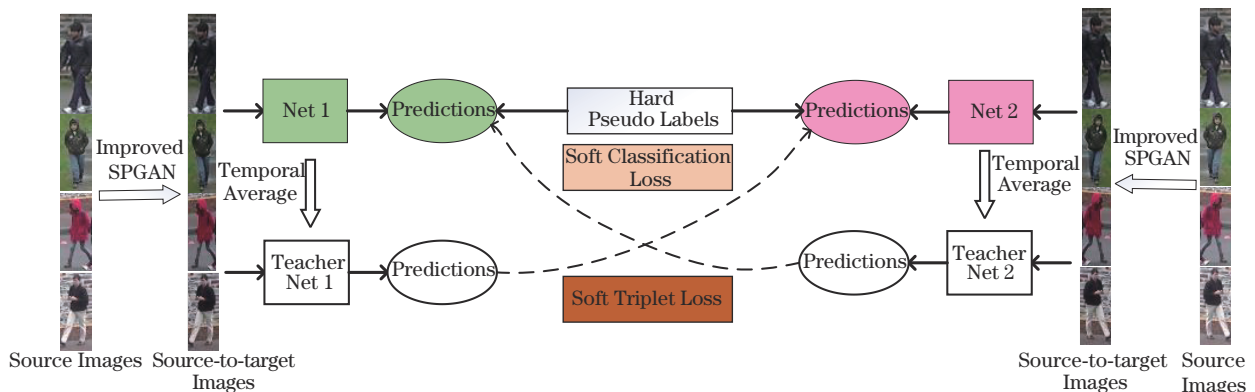


图1 基于改进SPGAN的软伪标签纠偏法模型

Fig. 1 Model of soft pseudo-label correction based on improved SPGAN

2.1.1 分类损失函数

为提高聚类算法精度,使用分类损失与三元损失来进行联合训练,其中分类损失作用于分类器的预测值,三元损失作用于图像特征。硬伪标签和软伪标签的分类损失函数分别为

$$\theta'_{id}(\theta) = \frac{1}{N_t} \sum_{i=1}^{N_t} \ell_{ce} \left\{ C' \left[F(\chi'_i | \theta) \right], \tilde{y}'_i \right\}, \quad (1)$$

$$\zeta'_{sid}(\theta_1 | \theta_2) = -\frac{1}{N_t} \sum_{i=1}^{N_t} \left\{ C'_2 \left[F(\chi'_i | E^{(T)}[\theta_2]) \right] \cdot \log C'_1 \left[F(\chi'_i | \theta_1) \right] \right\}, \quad (2)$$

$$\zeta'_{sid}(\theta_2 | \theta_1) = -\frac{1}{N_t} \sum_{i=1}^{N_t} \left\{ C'_1 \left[F(\chi'_i | E^{(T)}[\theta_1]) \right] \cdot \log C'_2 \left[F(\chi'_i | \theta_2) \right] \right\}, \quad (3)$$

式中: $F(\cdot | \theta)$ 为编码器; C 为分类器,用角标 s, t 来区分源域和目标域,角标 $1, 2$ 区分 Net 1 和 Net 2; $D_s =$

$\left\{ (\chi'_i, y'_i) \mid i=1 \right\}^{N_t}$ 代表源域图像及其标签; $D_t = \left\{ \chi'_i \mid i=1 \right\}^{N_t}$ 代表目标域的无标签图像; \tilde{y}'_i 为目标域图像

χ'_i 的硬伪标签; χ 和 χ' 代表经过不同数据增强的同一张图片; $C' \left[F(\chi'_i | E^{(T)}[\theta]) \right]$ 为软伪标签的分类预测值; ℓ_{ce} 为一般多分类交叉熵损失函数; $-q \log p$ 为软交叉熵损失函数。

2.1.2 三元损失函数

硬伪标签三元损失函数为

$$\zeta'_{in}(\theta) = \frac{1}{N_t} \sum_{i=1}^{N_t} \ell_{bce} [\tau_i(\theta), 1]. \quad (4)$$

由于伪标签存在噪声,不能完全区分正负样本,为更好支持软伪标签训练。使用 softmax-triplet 来表示三元组内特征间的关系:

$$\tau_i(\theta) = \frac{\exp \left[\left\| F(\chi'_i | \theta) - F(\chi'_{i,n} | \theta) \right\| \right]}{\exp \left[\left\| F(\chi'_i | \theta) - F(\chi'_{i,p} | \theta) \right\| \right] + \exp \left[\left\| F(\chi'_i | \theta) - F(\chi'_{i,n} | \theta) \right\| \right]} \quad (5)$$

式中: $\|\cdot\|$ 代表欧氏距离; $\chi'_{i,p}$ 和 $\chi'_{i,n}$ 分别代表 χ'_i 的正样本和负样本。则软伪标签三元损失函数可表示为

$$\zeta'_{stri}(\theta_1 | \theta_2) = \frac{1}{N_t} \sum_{i=1}^{N_t} \ell_{bce} [\tau_i(\theta_1), \tau_i(E^T[\theta_2])], \quad (6)$$

$$\zeta'_{stri}(\theta_2 | \theta_1) = \frac{1}{N_t} \sum_{i=1}^{N_t} \ell_{bce} [\tau_i(\theta_2), \tau_i(E^T[\theta_1])], \quad (7)$$

式中: $\tau_i(E^T[\theta])$ 表示用模型输出的特征距离比来代替硬标签 1。软化后的标签取值范围在 $[0, 1)$ 之间,损失函数让 Net 1 输出的 softmax-triplet 逼近 Mean-Net 2 的 softmax-triplet 预测值,从而实现三元组的软监督。

整体损失函数的表达式为

$$\begin{aligned} \zeta(\theta_1, \theta_2) = & (1 - \lambda'_{id}) [\zeta'_{sid}(\theta_1) + \zeta'_{sid}(\theta_2)] + \\ & \lambda'_{id} [\zeta'_{sid}(\theta_1 | \theta_2) + \zeta'_{sid}(\theta_2 | \theta_1)] + \\ & (1 - \lambda'_{in}) [\zeta'_{in}(\theta_1) + \zeta'_{in}(\theta_2)] + \\ & \lambda'_{in} [\zeta'_{stri}(\theta_1 | \theta_2) + \zeta'_{stri}(\theta_2 | \theta_1)], \end{aligned} \quad (8)$$

式中: λ'_{id} 、 λ'_{in} 分别是软分类损失函数和软三元损失函数的权重超参数。

所提软伪标签纠偏法与 Bhardwaj 等^[9]提出的 Teacher-Student Networks 有着本质区别。Teacher-Student Networks 使用一组协同训练的师生网络,两者直接相互监督,相互逼近。此情况下,两组网络输出预测值相似且容易失去本身输出独立性,从而导致训练误差放大模型崩溃。所提模型为保持每个网络的输出独立性,采取以下措施:将整个模型分类预测值的 1/2 作为软标签采用交叉纠偏方式抑制噪声;对输入

图像采取随机数据增强;对两组网络设置不同初始化参数。

2.1.3 IBN-a-ResNet-50

He 等^[10]提出的 ResNet-50 通过残差跳跃式的结构设计,解决了梯度消失的问题,提升了模型的识别精度。但针对无监督域自适应任务,残差网络作为骨干网络的识别精度不佳。本实验组将其与 IBN-Net 融合,IBN-Net 具有极好兼容性,能够当前主流网络架构进行有效结合(如 DenseNet、ResNet 等),在不影响计算效率的前提下提升模型域自适应性。

IBN-Net 由 IBN-a 和 IBN-b 两个模块构成:其中 IN 主要学习视觉变化的相关性,例如颜色、风格等;BN 主要学习内容相关的信息,同时可以加速训练和学习到更有区分性的特征。在残差块浅层使用 IN+BN 的组合,底层只使用 BN。ResNet-50 网络结构的实验结果已经证明恒等映射路径的有效性,所以在残差路径上添加了 IN。同时为了得到与恒等映射路径对齐的特征,避免特征不对齐的问题出现,IN 被加在残差模块中第 1 个卷积层之后,可以学习到更有效的特征,从而增强模型的域间泛化性能^[11]。

2.2.1 基于多尺度特征融合的生成器模型

SPGAN 模型的生成器由编码器、转换器、解码器组成,如图 3 所示:其中编码器包括 3 个步长为 2 卷积层,作用是提取输入图像的特征;转换器由若干个残差块组成,作用是将源域图像特征转换为接近目标域图像风格的特征;解码器由 3 个步长为 1/2 反卷积层组

成,作用是复原转换后的特征向量。但生成器的编码器特征提取时存在尺度和位置不变性的问题,并且只有最底层的特征信息才能进行转换,单一尺度的特征信息经过转换和复原后并不能很好反应目标域的风格

特征。针对上述问题,本实验组对 SPGAN 模型生成器的编码器进行多尺度特征融合改进,利用深浅特征层的多语义信息,缓解特征信息传递过程中丢失的问题,提升 SPGAN 模型的风格迁移效果。

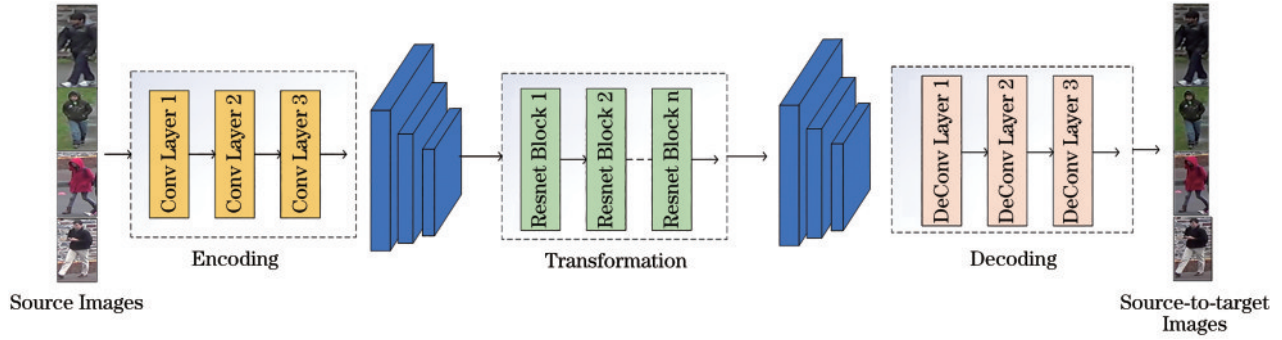


图 2 SPGAN 模型生成器结构图

Fig. 2 Generator structure diagram for SPGAN

基于多尺度特征融合的生成器模型改进分为 3 步。首先,由于 SPGAN 生成器网络的卷积过程是由下向上的,提取的特征也逐步加深。因此,在初步特征提取后,对特征图进行自上向下的 1×1 卷积降维操作,使得相邻特征层通道数相匹配。再对降维后的特征进行尺度为 2 的上采样操作,这样可以最大程度地

保留语义信息^[12]。

然后进行特征重构,过程如图 3 所示,对 Conv Layer 1~Conv Layer 3 等 3 个阶段的激活特征进行重构,共两阶段分别对应 S1 和 S2,其中“ $\times 2$ ”为尺度为 2 的上采样操作,“+”为特征融合。

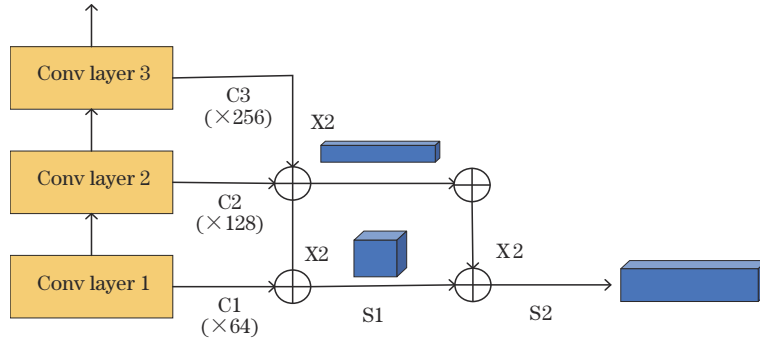


图 3 多尺度特征重构模型

Fig. 3 Multiple scale feature reconstruction model

最后,融合前考虑到深浅层特征的差异性,直接相加或拼接操作会增加训练的梯度计算,降低特征融合效果,因此采用哈达玛积特征融合法解决这个问题:

$$\begin{cases} \gamma_\epsilon = \gamma_h \otimes \gamma_l = [Q_{ij}]_{H \times w}, \\ Q_{ij} = a_{ij} * b_{ij} \end{cases} \quad (9)$$

$$\gamma_c = \gamma_\epsilon \oplus \gamma_h, \quad (10)$$

式中: γ_ϵ 为提取的特征; $\gamma_h = [a_{ij}]_{H \times w}$ 、 $\gamma_l = [b_{ij}]_{H \times w}$ 分别代表不同层的特征向量; γ_c 为融合后的特征; \otimes 为矩阵向量对应位置相乘; \oplus 为对应位置向量相加。卷积过程中,不同元素对应的卷积核不同,不同卷积核关注的图像区域也不同,特征向量值较大的为卷积层的关注区域^[13]。根据 Hadamard product 公式,特征差异较小相乘后结果会很大,而差异较大的结果则会很小,通

过此原理,保留了不同特征的共性特征信息并弱化特征间的显著性差异,提高了深浅特征的融合效果,最终生成强化版的特征信息,恢复后增强了模型的风格迁移效果。

3 实验结果与分析

3.1 数据集与评价指标

采用无监督行人重识别领域常用的 Market1501 和 Duke-MTMC-reID 作为实验数据集,其中 Market-1501 和 DukeMTMC-reID 的训练图像分别为 12936 和 16522 张^[14]。测试图像分别为 19732 张和 2228 张。两个数据集互为源域和目标域数据进行训练和测试,并采用行人重识别领域常用的 Rank 1、Rank 5、Rank 10 和平均精度均值(mAP)来作为评价标准^[15]。

3.2 实验环境和参数设置

实验环境基于 Linux 环境下的 Pytorch 1.0 框架, 采用 4 块 Tesla V100GPU 进行训练。输入图像大小为 256×256 , 动量因子为 0.9, 并以 0.5 的概率进行随机数据增强。实验中: 预训练模型训练次数 epoch 为 80, 每次训练样本数为 64; 正式训练模型训练次数 epoch 为 40, 每次训练样本数为 64。使用 ResNet-50 模型作为预训练的骨干网络, 使用风格迁移后的源域数据进行预训练, 改进后的 IBN-a-ResNet-50 作为正式训练的骨干网络, 使用目标域数据进行正式训练。

3.3 算法有效性验证实验

首先, 对所提软伪标签纠偏法在数据集 Market1501 和 Duke-MTMC-reID 上进行有效性实验验证

证, 两个数据集互为源域和目标域数据进行训练和测试, 实验结果如表 1 所示。从表中看出, 对原始聚类算法进行软伪标签纠偏法改进后, 在 Market1501→Duke-MTMC-reID 和 Duke-MTMC-reID→Market1501 任务中, mAP 达到 63.5% 和 71.4%, Rank 1 达到 77.5% 和 88.3%, 模型识别精度明显提升, mAP 提高了 9.3 个百分点~10.6 个百分点, 其他指标均有提升。实验结果表明, 软伪标签纠偏法利用并行网络的预测值作为软标签对含有噪声的伪标签纠偏, 有效降低了噪声带来的硬量化损失, 减小了伪标签噪声对网络训练性能的影响, 显著提高了模型的识别精度。由此可见, 对传统聚类算法进行软伪标签纠偏法的改进是必要的。

表 1 软伪标签纠偏法有效性实验验证结果
Table 1 Results of soft pseudo-label correction method verification experiment unit: %

Method	Market1501→ Duke-MTMC-reID				Duke-MTMC-reID→ Market1501			
	mAP	Rank 1	Rank 5	Rank 10	mAP	Rank 1	Rank 5	Rank 10
	ResNet-50-DBSCAN	54.2	71.1	82.5	86.0	60.8	75.0	84.7
Soft pseudo-label-ResNet-50-DBSCAN	63.5	77.5	87.6	93.1	71.4	88.3	95.6	97.5
Soft pseudo-label-IBN-a-ResNet-50-DBSCAN	67.2	80.2	89.9	92.7	76.6	89.7	95.4	97.7

然后, 改变训练骨干网络, 选择 IBN-a-ResNet-50 进行训练和测试, 对残差网络的改进——IBN-a-ResNet-50 进行有效性实验验证, 实验结果如表 2 所示。实验结果表明, 改变骨干网络后, mAP 达到 67.2% 和 76.6%, Rank 1 达到 80.2% 和 89.7%。较改进前, mAP 提高了 3.7 个百分点~5.2 个百分点, 其他指标均有提升。将 IBN-Net 和 ResNet-50 进行融合, 在保留残差结构优越性的前提下, 在浅层网络使用 BN 和 IN 联合解决外观差异带来的特征差异, 在深层网络使用 BN 来解决处于主导地位的内容差异。模型对不同域数据的泛化能力得到进一步提高, 在不增加

计算量的情况下, 提升了模型本身的域自适应性, 更加适合无监督域自适应任务。

最后, 对所提基于改进 SPGAN 模型的伪标签聚类算法进行实验, 实验结果如表 2 所示。从表中可以看出, 将 SPGAN 模型和伪标签算法进行结合后, mAP 达到 69.3% 和 79.2%, Rank 1 达到 81.6% 和 90.9%。较改进前, mAP 提高了 2.1 个百分点~2.6 个百分点, 识别精度较传统聚类算法有进一步提升, 仅在 Duke-MTMC-reID→Market1501 实验中, 仅有 Rank-10 的精度较改进前降低了 0.1 个百分点, 其他精度均有提升。

表 2 基于 SPGAN 模型的伪标签聚类算法有效性实验验证结果
Table 2 Results of pseudo-label clustering algorithm based on the SPGAN model verification experiment unit: %

Method	Market1501→Duke- MTMC-reID				Duke-MTMC-reID→ Market1501			
	mAP	Rank 1	Rank 5	Rank 10	mAP	Rank 1	Rank 5	Rank 10
	Soft pseudo-label-IBN-a-ResNet-50-DBSCAN	67.2	80.2	89.9	92.7	76.6	89.7	95.4
SPGAN-soft pseudo-label-IBN-a-ResNet-50-DBSCAN	69.3	81.6	90.4	93.5	79.2	90.9	94.5	97.6
Improved SPGAN-Soft pseudo-label-IBN-a-ResNet-50-DBSCAN	70.2	85.8	92.1	95.7	80.4	92.5	97.2	98.4

对 SPGAN 模型进行多尺度特征融合改进后, mAP 达到 70.2% 和 80.4%, Rank 1 达到 85.8% 和 92.5%。较改进前, mAP 提高了 0.9 个百分点~1.2 个百分点, 识别精度较传统 SPGAN 模型有进一步提升。实验结果表明, 引入改进 SPGAN 模型能够有效降低源域数据和目标域风格差异性的影响, 相较于传统聚类算法, mAP 提高了 3 个百分点~3.8 个百分点,

最高 mAP 达到 80.4%, Rank 1 达到 92.5%, 识别精度媲美有监督学习方法, 所提改进方法是有效的。

改进后生成器图像转换效果如图 4 所示, 图中为原数据集图像、SPGAN 输出图像和改进后 SPGAN 输出图像。从图中可以看出, 相较于原始 SPGAN 模型, 改进后的模型更接近目标域风格。实验结果表明, 改进后的生成器模型对风格迁移具有更好的视觉



图 4 转换后的抽样图像

Fig. 4 Visual examples of image-image translation

效果。

为进一步验证基于多尺度特征融合生成器模型的有效性,对相关卷积层特征进行可视化操作,实验结果如图 5 所示。从图中可以看出,在域自适应行人重识

别任务中,改进后 SPGAN 模型所提取的特征较改进前保留了更加丰富的细节特征,表明了所提改进方法的有效性,能够更好迁移目标域风格,对于降低域间差异对模型迁移精度的影响有重要意义。

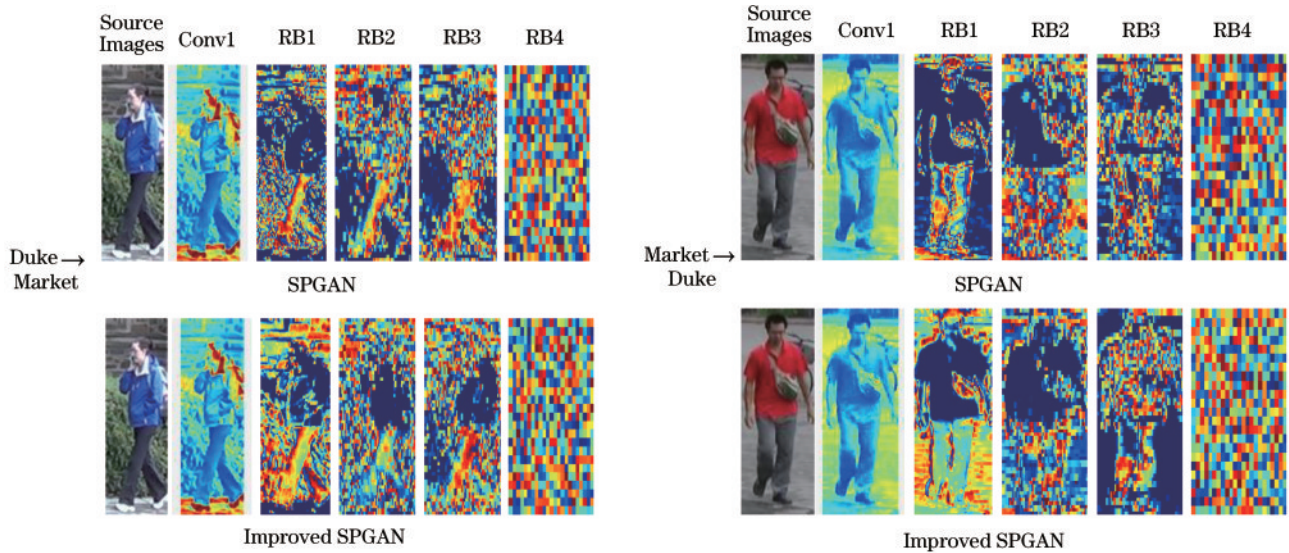


图 5 模型特征图可视化对比图

Fig. 5 Visual collation map of model characteristic pattern

3.3.5 所提方法与主流方法性能对比

将实验结果与主流无监督行人重识别方法进行对比,如表 3 所示,“-”表示对应数据并未实现。由表 3

可知,所提方法相较于其他主流方法,mAP 和其他指标均取得了最高精度,说明所提方法的有效性和优越性。

表 3 与相关方法无监督行人重识别精度对比

Table 3 Comparison of unsupervised person recognition accuracy with related methods

unit: %

Method	Market1501→Duke-MTMC-reID				Duke-MTMC-reID→Market1501			
	mAP	Rank 1	Rank 5	Rank 10	mAP	Rank 1	Rank 5	Rank 10
PAD-Net ^[17]	45.1	63.2	77.0	82.5	47.6	75.2	86.3	90.2
MMT-700/500 ^[20]	68.7	81.8	91.2	93.4	76.5	90.9	96.4	97.9
AE ^[22]	46.7	67.9	79.2	83.6	58.0	81.6	91.9	94.6
Co-teaching-500 ^[18]	61.7	77.6	88.0	90.7	71.7	87.8	95.0	96.5
ECN ^[21]	40.4	63.3	75.8	80.4	43.0	75.1	87.6	91.6
AD-Cluster ^[19]	54.1	72.6	82.5	85.5	68.3	86.7	94.4	96.5
PCB-PAST ^[23]	54.3	72.4	—	—	54.6	78.4	—	—
SSG ^[16]	53.4	73.0	80.6	83.2	58.3	80.0	90.0	92.4
Proposed method	70.2	85.8	92.1	95.7	80.4	92.5	97.2	98.4

4 结 论

提出了一种基于软伪标签和多尺度特征融合的无监督域自适应行人重识别方法。首先,通过并行网络结构使用软标签对伪标签噪声进行纠偏并通过数据增强、交叉纠偏等方式保持网络独立性。在保证聚类算法优越性的前提下,降低伪标签噪声对网络训练性能的影响。然后,通过所提基于多尺度特征融合的SPGAN模型实现源域数据到目标域风格的转换并融合实例归一化和批量归一化网络的方法,增强网络的域间泛化能力。在 Market-to-Duke 和 Duke-to-Market 任务中都取得了较好的效果,最高 mAP 达到 70.2% 和 80.4%,识别精度媲美有监督学习。但本研究并未对聚类算法进行深入研究,如何提升算法聚类类别的准确性是今后需要完善的工作。

参 考 文 献

- [1] 郭迎春, 冯放, 阎刚, 等. 基于自适应融合网络的跨域行人重识别方法[J/OL]. 自动化学报: 1-13[2022-10-25]. DOI:10.16383/j.aas.c220083.
Guo Y C, Feng F, Yan G H, et al. Cross-domain pedestrian reidentification method based on Adaptive Fusion Network[J/OL]. Journal of Automation: 1-13 [2022-10-25]. DOI:10.16383/j.aas.c220083.
- [2] Mohanty A, Banerjee B, Velmurugan R. SSMTReID-Net: multi-target unsupervised domain adaptation for person re-identification[J]. Pattern Recognition Letters, 2022, 163: 40-46.
- [3] 杨萍, 吴晓红, 何小海, 等. 逐点特征匹配的跨域行人重识别方法[J]. 模式识别与人工智能, 2022, 35(6): 516-525.
Yang P, Wu X H, He X H, et al. Cross-domain pedestrian reidentification method for point-by-point feature matching[J]. Pattern Recognition and Artificial Intelligence, 2022, 35(6): 516-525.
- [4] Fan H H, Zheng L, Yan C G, et al. Unsupervised person re-identification: clustering and fine-tuning[J]. ACM Transactions on Multimedia Computing, Communications,

and Applications, 2018, 14(4): 83.

- [5] Ge Y X, Li Z W, Zhao H Y, et al. FD-GAN: pose-guided feature distilling GAN for robust person re-identification[EB/OL]. (2018-10-06) [2021-02-04]. <https://arxiv.org/abs/1810.02936>.
- [6] Deng W J, Zheng L, Ye Q X, et al. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 994-1003.
- [7] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.
- [8] Yang F X, Li K, Zhong Z, et al. Asymmetric Co-teaching for unsupervised cross-domain person re-identification[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12597-12604.
- [9] Bhardwaj S, Khapra M M. I have seen enough: a teacher student network for video classification using fewer frames [EB/OL]. (2018-05-12) [2021-02-04]. <https://arxiv.org/abs/1805.04668>.
- [10] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [11] Pan X G, Luo P, Shi J P, et al. Two at once: enhancing learning and generalization capacities via IBN-net[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11208: 484-500.
- [12] Li H F, Pang J, Tao D P, et al. Cross adversarial consistency self-prediction learning for unsupervised domain adaptation person re-identification[J]. Information Sciences, 2021, 559: 46-60.
- [13] Chong Y W, Peng C W, Zhang J J, et al. Style transfer for unsupervised domain-adaptive person re-identification

- [J]. *Neurocomputing*, 2021, 422: 314-321.
- [14] Chang X B, Yang Y X, Xiang T, et al. Disjoint label space transfer learning with common factorised space[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33: 3288-3295.
- [15] Zhong Z, Zheng L, Luo Z M, et al. Invariance matters: exemplar memory for domain adaptive person re-identification[EB/OL]. (2019-04-03)[2021-02-05]. <https://arxiv.org/abs/1904.01990>.
- [16] Fu Y, Wei Y C, Wang G S, et al. Self-similarity grouping: a simple unsupervised cross domain adaptation approach for person re-identification[EB/OL]. (2018-11-26)[2021-02-05]. <https://arxiv.org/abs/1811.10144>.
- [17] Song L C, Wang C, Zhang L F, et al. Unsupervised domain adaptive re-identification: theory and practice[J]. *Pattern Recognition*, 2020, 102: 107173.
- [18] Han B, Yao Q M, Yu X R, et al. Co-teaching: robust training of deep neural networks with extremely noisy labels[EB/OL]. (2018-04-18)[2021-02-05]. <https://arxiv.org/abs/1804.06872>.
- [19] Zhai Y P, Lu S J, Ye Q X, et al. AD-cluster: augmented discriminative clustering for domain adaptive person re-identification[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 9018-9027.
- [20] Ge Y X, Chen D P, Li H S. Mutual mean-teaching: pseudo label refinery for unsupervised domain adaptation on person re-identification[EB/OL]. (2020-01-06)[2021-02-05]. <https://arxiv.org/abs/2001.01526>.
- [21] Zhong Z, Zheng L, Luo Z M, et al. Invariance matters: exemplar memory for domain adaptive person re-identification [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 598-607.
- [22] Li Y J, Lin C S, Lin Y B, et al. Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 7918-7928.
- [23] Zhang X Y, Cao J W, Shen C H, et al. Self-training with progressive augmentation for unsupervised cross-domain person re-identification[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 8221-8230.