风格迁移和标签细化的域自适应行人重识别

董柔汐[1]， 朱 朦[2,3]， 任洪娥[1,3]， 赵 龙[1,4]， 陶 锐[1,5]
（1. 东北林业大学 计算机与控制工程学院，哈尔滨 150040；2. 哈尔滨学院 人工智能学院，哈尔滨 150086；
3. 黑龙江林业智能装备工程研究中心，哈尔滨 150040；4. 黑龙江东方学院 信息工程学院，哈尔滨 150066；
5. 呼伦贝尔学院 计算机学院，内蒙古 呼伦贝尔 021008）

摘 要：针对单独采用生成对抗网络难以充分获取行人特征且聚类生成伪标签存在噪声的问题，提出了生成对抗网络进行风格迁移，融合空间和通道注意力机制，进行全局和局部特征标签细化的域自适应行人重识别（style transfer and label refinement，STALR）。该方法利用生成对抗网络和聚类的双重域自适应网络，融入注意力机制对行人的全局和局部特征进行提取，利用局部和全局特征的相互关系进行标签细化，获取更准确的标签训练网络。实验结果表明：模型以 DukeMTMC-reID/Market1501 作为源域，Market1501/DukeMTMC-reID 作为目标域，mAP 分别取得 82.5% 和 71.8%，Rank-1 分别取得 93.8% 和 83.7%，识别效果优于现有域自适应行人重识别算法。

关键词：计算机视觉；行人重识别；域自适应；风格迁移；伪标签细化
DOI：10.15938/j.jhust.2026.03.003
中图分类号：TP751.1 **文献标志码**：A **文章编号**：1007-2683(2026)03-0024-11

Domain-Adaptive Person Re-Identification Based on Style Transfer and Pseudo-Label Refinement

DONG Rouxi[1]， ZHU Meng[2,3]， REN Honge[1,3]， ZHAO Long[1,4]， TAO Rui[1,5]
（1. College of Computer and Control Engineering，Northeast Forestry University，Harbin 150040，China；
2. School of Artificial Intelligence，Harbin University，Harbin 150086，China；
3. Heilongjiang Forestry Intelligent Equipment Engineering Research Center，Harbin 150040，China；
4. School of Information Engineering，East University of Heilongjiang，Harbin 150066，China；
5. School of Computer Science，Hulunbuir University，Hulunbuir 021008，China）

Abstract：To address the issues of insufficient person feature acquisition when solely using generative adversarial networks and the noise in pseudo-labels generated by clustering，a domain adaptive person re-identification model（Style Transfer and Label Refinement，STALR）is proposed，which performs style transfer via a generative adversarial network and integrates spatial and channel attention mechanisms to refine global and local feature labels. The method utilizes a dual-domain adaptive network based on generative adversarial networks and clustering，incorporating attention mechanisms to extract global and local features. It leverages the correlation between local and global features for label refinement，thereby obtaining more accurate labels for network training. The experimental results show that when using DukeMTMC-reID/Market1501 as the source domain and Market1501/DukeMTMC-reID as the target domain，mAP reaches 82.5% and 71.8%，and Rank-1 reaches 93.8% and 83.7%，respectively. The recognition performance outperforms existing domain-adaptive person re-identification algorithms.

Keywords：computer vision；person re-identification；domain adaption；style transfer；pseudo-label refinement

收稿日期：2024-08-19
基金项目：黑龙江省自然科学基金（LH2020F040）；中央高校基本科研业务费专项资金（2572017PZ10）.
作者简介：董柔汐（1999—），女，硕士研究生；
任洪娥（1962—），女，博士，教授，博士研究生导师.
通信作者：朱 朦（1989—），女，博士，讲师，E-mail：zhumeng913@hrbu.edu.cn.

0　引　言

行人重识别(person re-identification, Re-ID)[1-2]旨在跨相机检索同一目标行人,是智能视频监控领域的一项关键技术。行人重识别在失踪人口搜索、罪犯追踪和特定区域行人检索[3]等方面有广泛应用,同时重识别还可应用于动物中[4]。有监督行人重识别[5]已经取得了较好的性能,但是却忽略了在实际场景中对新领域的适应性,并且依赖大量带标签的行人数据。而获得带标签的行人数据需要进行标注,耗费大量的人力物力。监督方法在无样本情况下性能较差,不足以落地到实际应用,故本文将采用无监督自适应进行行人重识别。

无监督自适应行人重识别利用来自源域的带注释的数据来调整模型,使其适应目标域。现有工作可归纳为两类:基于生成对抗网络[6]的方法和基于聚类生成伪标签[7]的方法。前者通过进行源域和目标域之间的特征分布对齐或知识的迁移[8],缓解域间差异。一些研究人员通常采用生成对抗网络执行图像的风格迁移,将源域的风格向目标域转换。LI 等[9]利用矢量量化空间中的离散矢量,解决了原始生成器产生噪声伪图像的问题,从而生成质量更高的风格转换图像。还有学者通过生成对抗网络抑制背景的偏差,消除不同领域之间的特征分布差异[10]。但是由于生成对抗网络的生成器只会生成少数几种特定模式的行人图像,无法涵盖行人的多样性,在实际场景中适应能力不稳定,近年来不单独采用生成对抗网络进行域自适应行人重识别。

后者利用聚类技术在无监督的情况下将源域学习到的知识迁移到目标域,以实现目标域中行人的准确重识别。基本方法分为 2 种:第 1 种是进行伪标签的生成[11],通过主流聚类方式对未标注的行人数据图像进行聚类操作,将每个聚类结果视为 1 个类别,为聚类中的样本分配伪标签作为临时标注信息进行后续的模型训练。第 2 种是通过聚类进行特征学习和对齐,让模型学习到的特征在不同行人域保持一致的语义表征和判别能力,使得源域目标域的特征分布尽可能接近。基于聚类的方法不仅能挖掘数据内在的结构,还可与其他方法结合,进一步提升行人重识别的能力。但是聚类生成的伪标签可能存在噪声,尤其是数据存在复杂的分布变化时,伪标签的准确性难以保证。

如果能解决噪声伪标签问题,获得更干净的伪标签进行训练,可大幅提高行人重识别的准确性。因此,目前的工作集中于解决标签噪声的影响,一部分通过聚类本身的改进实现[12],大部分通过伪标签的细化完成[13]。在伪标签细化方面,GE 等[14]提出的网络采取了协同网络学习,通过离线优化的硬伪标签和在线优化的软伪标签来替代训练的方式,从目标域学习更好的特征,从而优化目标域的伪标签。TENG 等[15]提出了邻域引导的通用标签细化模块,该模块用来自相邻样本的软标签替换了同质和异构空间中的硬伪标签,减少标签噪声。以上方法通过软标签和硬标签交替训练产生了很好的效果,但是只涉及通过全局特征的伪标签细化,没有注意到行人的局部特征也可辅助全局特征标签的细化。也有一些方法关注于聚类被丢弃的离群值和损失函数的设计来解决伪标签的噪声,GE 等[16]引入对比学习解决聚类产生的离群值无法被充分挖掘的问题。ZHENG 等[17]通过加权身份损失的设计,利用置信度加权策略降低模型对错误标签的依赖。吴蕾等[18]提出的多尺度局部特征融合,关注到局部特征关联性信息缺失,提供了新思路。本文通过全局特征和局部特征的关系进行标签的细化,不仅可以关注到行人判别性的局部特征,也可以利用局部特征对全局特征进行细化。

针对上述问题,本文提出双重域自适应框架,创新性地融合生成对抗网络的风格迁移与聚类生成伪标签的方式,解决源域和目标域之间存在的较大视觉差异。此外,设计了利用局部特征与全局特征相互关系进行细化的模块,为行人的局部特征和全局特征分配相同的伪标签,并计算相关性得分,依据此分数进行两部分的标签细化。第 1 步是进行局部特征的细化,局部特征存在一部分有助于判断行人信息的特征,也有一部分背景或模糊信息。调整 3 块局部特征的优化权重,权重低时将局部标签视为无关信息。第 2 步进行全局特征的细化,根据不同权重的局部特征与预测结果进行聚合,获得更可靠的全局特征。与现有方法相比,本文显著提升了模型在类似实际场景的目标域泛化能力。

1　研究技术与模型架构

1.1　总体架构

本文提出了解决现实中领域间差异较大及伪标签存在噪声的方法。首先,使用 SPGAN(similarity preserving cycle consistent generative adversarial net-

work)网络将源域的图像转换为目标域风格的图像进行预训练。然后,经过引入了空间及通道注意力机制的 Resnet50[19] 网络寻找行人重识别需要的重要信息。最后,关注了行人局部特征的辅助作用,对行人的全局特征和局部特征进行聚类并分配伪标签,计算相关性得分。提取匹配度更高的局部特征对全局特征进行细化,更新更准确的伪标签来减少噪声对模型效果的影响。总体框架如图 1 所示,CA 模块为通道注意力模块,SA 模块为空间注意力模块。

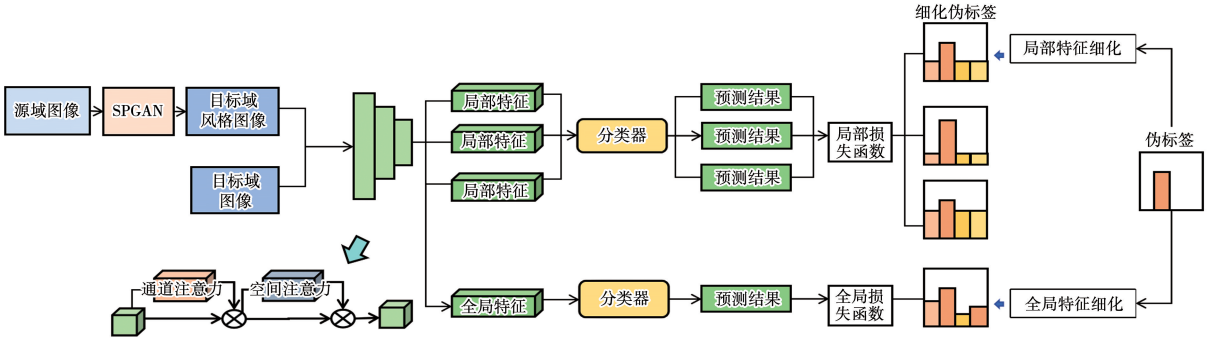


图 1 总体框架图

Fig. 1 General frame diagram

1.2 域转换

无监督聚类行人重识别模型先提取行人特征并完成聚类,根据聚类结果为特征分配伪标签,使用标签进行模型的预训练。聚类方法借助特征相似性约束来降低源域和目标域差异过大而造成的影响。为了进一步降低域间差异,将风格迁移模型 SPGAN 引入无监督聚类方法中,通过 SPGAN 模型将源域数据转换为目标域数据的背景风格。使用域转换后的源域图像对以 ResNet50[17] 为骨干网络的模型进行预训练,获得更倾向于目标域图像风格特征的模型。然后,对无标签数据进行特征提取,使用基于密度的聚类算法(density-based spatial clustering of applications with noise,DBSCAN)进行聚类分配伪标签,并进行标签细化。行人重识别的基础骨干网络如图 2 所示。

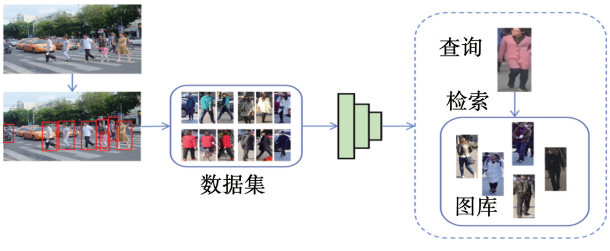


图 2 行人重识别框架

Fig. 2 Person re-identification framework

图 3 为 SPGAN 的网络结构,共分为 2 部分,第 1 部分是 SiaNet(siamese network)孪生网络。通过 SiaNet 控制样本对的远近,将迁移后的图像与目标域中的对应图像聚集到一起,远离目源域的所有图像,更好地学习目标域的风格。第 2 部分是 CycleGAN(cycle generative adversarial network)网络。CycleGAN 包含 2 个生成器和 2 个判别器,其中 1 个生成器完成源域图像到目标域图像的风格迁移,另 1 个进行目标域图像到源域图像的风格迁移。CycleGAN 通过对抗损失、循环一致性损失、ID 不变损失联合进行优化。

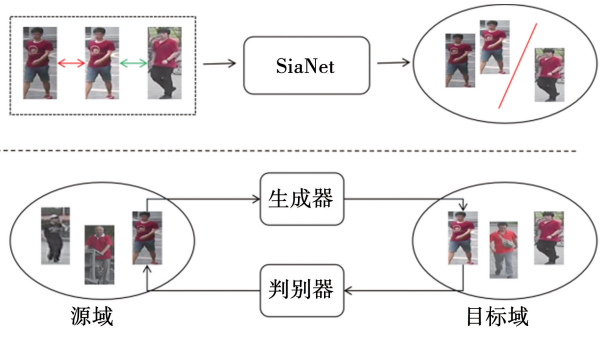


图 3 SPGAN 网络结构

Fig. 3 SPGAN network structure

SiaNet 利用一致性损失函数监督:

$$\mathcal{L}_{con}(i,\boldsymbol{x}_1,\boldsymbol{x}_2)=(1-i)\{\max(0,m-d)\}^2+id^2 \tag{1}$$

式中:$\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 为 1 对输入图像向量;d 为 2 个输入向量的欧式距离;i 为样本配对标签,如果 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 为同一行人的图像,则 $i=1$ 为正对,否则 $i=0$ 为负对;m 为 margin 参数,$m\in[0,2]$。当 $m=0$ 时,负对计算的损失不反向传播;$m>0$ 时,表示正对和负对都参与计算,m 越大表示负对损失在反向传播时的

权重越大。使用转换风格后的图片对加载好的模型进行预训练,可获得对目标域效果更好的模型,提高后续行人重识别任务的效果。

1.3　注意力机制

本文将空间注意力机制和通道注意力机制融合到骨干网络,使行人在特征图中分配较高的权重。此模块能解决行人图像模糊或行人部位被遮挡、图像中无关背景占比较大的问题。通道信息代表着图像的不同特征,因此,对通道进行选择可使网络注意到对检测有利的行人特征。空间注意力可以寻找网络中需要关注的行人信息进行处理。本文提出的注意力模块如图 4 所示。

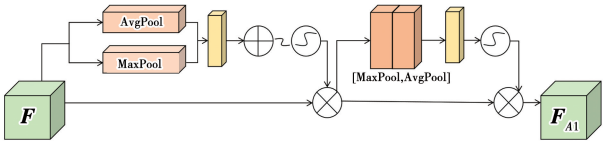


图 4　空间通道注意力模块网络结构

Fig. 4　Spatial channel attention module network structure

通道注意力用于建模不同权重的特征通道。针对行人重识别任务,该机制对有效特征通道进行增强,同时抑制无关冗余通道。本文设计了通道注意力模块,输入行人图像特征图 $\boldsymbol{F}\in\mathbb{R}^{H\times W\times C}$,其中 $H\times W$ 为空间维度,C 为通道数。对输入特征图进行平均池化和最大池化,获取每个通道的平均值和通道空间最大值,得到 2 组 $1\times1\times C$ 的特征向量后,进行特征向量拼接,经由全连接层将特征图映射到样本标记空间,再经过 Sigmoid 激活函数归一化,生成最终的通道注意力特征 $A_C(\boldsymbol{F})$。最后,将注意力权重与输入的特征图 $\boldsymbol{F}$ 相乘,生成通道注意力优化后的特征 $\boldsymbol{F}_A$。通道注意力公式如下:

$$\left.\begin{aligned}&\boldsymbol{F}_A=A_c(\boldsymbol{F})\otimes\boldsymbol{F}\\&A_c(\boldsymbol{F})=\sigma(C_2(\mathrm{ReLU}(C_1\boldsymbol{F}_{\mathrm{Avg}}))+C_2(\mathrm{ReLU}(C_1\boldsymbol{F}_{\mathrm{Max}})))\end{aligned}\right\}\tag{2}$$

式中:$A_c(\boldsymbol{F})$为通道注意力权重计算函数;σ 为 Sigmoid 归一化激活函数;C_1,C_2 为卷积层;$\boldsymbol{F}_{\mathrm{Avg}}$、$\boldsymbol{F}_{\mathrm{Max}}$分别为特征在通道维度平均池化和最大池化得到的聚合特征。

空间注意力机制在空间中进行转换并获取需要的重要特征,将通道注意力机制的输出特征通过最大池化和平均池化,得到 2 个 $H\times W\times1$ 特征图。然后,对 2 个特征图做通道拼接,经过 1 个 7×7 卷积将特征图降维成 1 个通道,通过 Sigmoid 激活生成空间注意力权重图 $A_S(\boldsymbol{F})$。最后,与输入的特征图 $\boldsymbol{F}_A$ 相乘,变回原来的大小,生成空间注意力所需要的特征。空间注意力公式如下:

$$\left.\begin{aligned}&\boldsymbol{F}_{A1}=A_S(\boldsymbol{F}_A)\otimes\boldsymbol{F}_A\\&A_S(\boldsymbol{F})=\sigma(f_{7\times7}([F_{\mathrm{Avg},S};F_{\mathrm{Max},S}]))\end{aligned}\right\}\tag{3}$$

式中:$\boldsymbol{F}_{A1}$ 为空间注意力增强后的输出特征图;$A_S(\boldsymbol{F})$为空间注意力权重计算函数;$f_{7\times7}$为 7×7 卷积;$F_{\mathrm{Avg},S}$与 $F_{\mathrm{Max},S}$为空间维度平均池化和最大池化得到的聚合特征。

1.4　标签细化

本文利用水平切片获得的局部特征降低全局特征聚类的标签噪声。首先,从数据集获取的行人图片经水平三等分得到行人的 3 个精细局部特征,同时,将行人图片输入神经网络获取全局特征。然后,对局部特征和全局特征进行密度聚类,分配相同的伪标签。最后,利用局部特征和全局特征的关系进行相互细化,减少噪声对模型效果的影响。

仅通过全局特征进行聚类生成的标签训练网络,可能将并不相同的 2 张图片划分为同一行人。因为 2 张不同行人被关注的部分可能非常相似,缺失一部分细节特征,导致后续进行监督训练时,把错误的行人信息传递给模型,导致对行人重识别的准确率下降。同样,仅通过局部特征进行聚类获得的伪标签,不利于快速获得图像间的差异,无法获得关键性的特征进行不同行人的判断。全局特征和局部特征相互配合可发现相似行人细节上的差异及判别性的细节,通过互相细化获得更干净的标签,使后续模型训练获得更好的结果。先对图像 m_i 聚类生成的伪标签 $\boldsymbol{x}_i$ 的局部特征进行细化处理。定义细化公式如下:

$$\boldsymbol{x}_{i,pn}=\lambda_{i,pn}\boldsymbol{x}_i+(1-\lambda_{i,pn})\boldsymbol{u}\tag{4}$$

式中:$\boldsymbol{x}_{i,pn}$为经过局部特征细化的伪标签;n 为分块值;$\boldsymbol{u}$ 为均匀分布向量;$\lambda_{i,pn}$为通过 k 近邻算法进行相似度排名并计算杰卡德系数获得的伪标签平滑的权重。

使用局部特征的预测结果,可令全局特征生成的标签更准确。全局特征经过聚类生成的伪标签存在噪声,一方面是因为聚类算法自身存在噪声信息,另一方面是因为算法忽略了丰富的局部特征信息,进而导致标签匹配错误。行人的局部特征含有较多细粒度信息,通过一部分特别的衣着配饰可判断出行人,但也有一部分局部信息对判别行人没有作用。将不同权重的局部特征信息进行综合利用,细化全局特征,更新生成的标签。对全局特征进行细化的处理公式如下:

$$\boldsymbol{x}_{ig}=\alpha\boldsymbol{x}_i+(1-\alpha)\sum_{n=1}^{3}\left(\frac{\exp\boldsymbol{\lambda}_{i,pn}}{\sum_{K}\exp\boldsymbol{\lambda}_{i,pn}}\boldsymbol{y}_{i,pn}\right) \tag{5}$$

式中:$\boldsymbol{x}_{ig}$为经过全局特征细化的伪标签;$\boldsymbol{y}_{i,pn}$为局部特征的预测向量;$\alpha\in(0,1)$控制伪标签和预测集合的比例。

伪标签细化模块通过互相优化的方法细化了全局特征和局部特征的标签。针对局部特征的细化,训练使用的交叉熵损失函数如下:

$$L_{\mathrm{p}}=-\sum_{i=1}^{D}\sum_{n=1}^{3}(\lambda_{i,pn}H(\boldsymbol{x}_i,\boldsymbol{y}_{i,pn})+(1-\lambda_{i,pn})D_{\mathrm{KL}}(\boldsymbol{u}\parallel\boldsymbol{y}_{i,pn})) \tag{6}$$

式中:D为图片数量;$H(\ )$为交叉熵函数;$D_{\mathrm{KL}}(\ )$为KL散度。

针对全局特征的细化,训练使用的交叉熵损失函数如下:

$$L_{\mathrm{g}}=-\sum_{i=1}^{D}\boldsymbol{x}_{i,g}\log(\boldsymbol{y}_{i,g}) \tag{7}$$

此模块的损失函数定义为上述2个损失函数之和,即

$$L=L_{\mathrm{g}}+L_{\mathrm{p}} \tag{8}$$

2 实验与结果分析

2.1 数据集设置和评估方法

本文在3个大型的行人重识别基准数据集,即Market-1501[20]数据集、DukeMTMC-ReID[21]数据集和MSMT17[22]数据集上对本文方法进行了性能评估。利用平均精度均值(mAP)和Rank-k(R_k)作为模型的评估指标。

mAP是对所有类别计算的平均精确率,可以反映检索结果中正确行人图像的排列顺序,能准确地获得多类别场景下行人重识别的性能。Rank-k(R_k)是目标图像检测返回的排序列表中,前k位查找到目标行人的概率值。Rank-1(R_1)是搜索结果中置信度最高的1张图是目标行人的概率。

Market-1501[20]数据集是在清华大学通过5个高清摄像头和1个低清摄像头采集到的1 501个行人的32 668张图像。每个行人在同一摄像头中会被多次采集,并且在跨2个摄像头及以上的场景下可以匹配到同一行人对应的图像。训练集有751人,包含12 936张图片;测试集有750人,包含19 732张图像。数据集集中在校园,光照较为稳定,但摄像头视角差异明显。Market-1501作为学术研究的基准数据集,常用于验证行人重识别算法的基本性能。

DukeMTMC-ReID[21]数据集是在Duke校园通过8个静态摄像头采集到的1 812个不同行人的36 441张图像,其中训练图像有16 552张,查询图像有2 228张,验证图像有17 661张。数据集中包含很多室外场景,行人背景多样性较高,适合进行在复杂背景下的模型鲁棒性测试。因此,在使用此数据集进行模型训练时要注重行人关键性特征的提取和跨场景行人的匹配准确性。

MSMT17[22]数据集是在校园内通过12个户外摄像头和3个室内摄像头采集到的4 101个不同行人的126 441张图片。该数据集有大量的行人数据,更接近真实场景。训练集有1 041人,包含32 621张图片;测试集有3 060人,包含93 820张图片。MSMT17是目前最大规模的行人重识别数据集,覆盖室内外多个场景。采用真实复杂环境下的行人数据集验证模型在多变环境下的泛化能力,未来可实现安防监控等领域的落地应用。

以上3个数据集反映了行人重识别研究从实验室到真实场景的需求,可推动算法在复杂条件下的实用性提升。由于MSMT17和实际场景比较类似,不作为源域进行模型的训练。除此之外,3个数据集分别作为源域和目标域来提升行人重识别模型的跨域能力,也有助于未来在实际场景中的部署。针对数据集背景复杂的问题,可通过注意力机制更好地关注行人的判别特征,同时利用全局特征和局部特征的关系对标签进行细化,减少聚类噪声标签对模型的影响。

2.2 实验设置

本实验的软件和硬件环境相同,保证实验的准确性。本文利用在ImageNet[23]预训练的ResNet50[19]作为主干网络,其中行人图片尺寸统一为256×128像素,采取随机擦除和自动增强的方法对源域和目标域进行数据增强。训练轮数设置为60轮,采用Adam进行模型的优化,权重设置为5×10^{-4}。在20轮之前,学习率初始化数值为0.1,20轮到40轮为0.01,最后20轮设置为0.001,在标签细化模块中,α值取0.5。

2.3 与现有方法比较

本文在上述3个数据集进行无监督域自适应行人重识别实验,并将实验结果与相关方向文章的实验结果进行对比分析。本文方法实验结果如表1~2所示。

表1为Duke和Market-1501分别作为源域和目标域进行域自适应的实验结果。其中Duke作为源域时,α_{mAP}达到了82.5%,R_1为93.8%;Market1501

作为源域时,α_{mAP}达到了 71.8%,R_1 为 83.7%。与表中和本文思路类似的 2 个工作进行对比,SSG 和本文一样考虑到全局特征和局部特征对行人识别效果的影响,P2LR 和本文采取不同的方式对伪标签进行平滑细化,本文的 2 个主要评价指标 α_{mAP}和 R_1 都优于以上 2 种方法。相较于 GPMMT 方法,本文方法在两组跨域任务中,α_{mAP}分别提高 1.2 和 1.9 个百分点,R_1 分别提高 1.4 和 0.6 个百分点。

表 2 为 Duke 和 Market1501 作为源域,MSMT17 作为目标域进行域自适应的实验结果。由于 MSMT17 和实际场景具有很大的相似性,故 MSMT17 不作为源域进行实验。MSMT17 作为目标域时,评价指标可展示本文域自适应的效果。其中 Duke 作为源域时,α_{mAP} 达到了 32.9%,R_1 为 64.1%;Market1501 作为源域时,α_{mAP} 达到了 33.4%,R_1 为 65.8%。相较于 P2LR 方法,本文方法在两组跨域任务中的 α_{mAP}分别提高 3 和 4.4 个百分点,R_1 分别提高 3.2 和 7 个百分点。通过表 1~2 可知,本文方法和主流方法相比,重识别效果有所提升。

表 1　Duke 和 Market-1501 数据集上的实验结果对比(%)

Tab. 1　Comparison of experimental results on Duke and Market-1501 datasets

方法	Duke to Market-1501				Market-1501 to Duke			
	α_{mAP}	R_1	R_5	R_{10}	α_{mAP}	R_1	R_5	R_{10}
PUL[24]	20.5	45.5	60.7	66.7	16.4	30.0	43.4	48.5
ECN[25]	43.0	75.1	87.6	91.6	40.4	63.3	75.8	80.4
SSG[26]	58.3	80.0	90.0	92.4	53.4	73.0	80.6	83.2
ADCluster[27]	68.3	86.7	94.4	96.5	54.1	72.6	82.5	85.5
MMT[14]	71.2	87.7	94.9	96.9	65.1	78.0	88.8	92.5
SPCL[16]	76.7	90.3	96.1	97.6	67.7	82.1	90.1	92.5
MEBNet[15]	76.0	89.9	96.0	97.5	66.1	79.6	88.3	92.2
Dual-Refinement[28]	78.0	90.9	96.4	97.7	67.7	82.1	90.1	92.5
UNRN[29]	78.1	91.9	96.1	97.8	69.1	82.0	90.7	93.5
P2LR[30]	81.0	92.6	97.4	98.3	70.8	82.6	90.8	93.7
HQP[31]	80.3	92.3	96.9	97.9	68.0	82.6	90.2	92.4
GCL +[32]	76.5	91.6	96.3	97.6	68.3	82.6	89.4	91.2
GPMMT[33]	81.3	92.4	97.6	98.4	69.9	83.1	92.5	93.9
本文方法	82.5	93.8	97.6	98.3	71.8	83.7	92.0	93.7

表 2　MSMT17 数据集的算法实验结果对比

Tab. 2　Comparison of algorithm experimental results on MSMT17

方法	Duke to MSMT17				Market – 1501 to MSMT17			
	α_{mAP}	R_1	R_5	R_{10}	α_{mAP}	R_1	R_5	R_{10}
ECN[25]	8.5	23.5	36.3	42.1	10.2	30.2	41.5	46.8
SSG[26]	13.2	31.9	—	49.6	13.3	32.2	—	51.2
MMT[14]	22.9	49.2	63.1	68.8	23.3	50.1	63.9	69.8
SPCL[16]	26.5	53.1	65.8	70.5	26.8	53.7	65.0	69.8
Dual-Refinement[28]	26.9	55.0	68.4	73.2	25.1	53.3	66.1	71.5
UNRN[29]	26.2	54.9	67.3	70.6	25.3	52.4	64.7	69.7
P2LR[30]	29.9	60.9	73.1	77.9	29.0	58.8	71.2	76.0
UDA-RI[34]	29.7	59.1	70.3	74.2	29.1	58.4	68.8	72.9
本文方法	32.9	64.1	74.5	79.4	33.4	65.8	75.0	79.7

2.4 消融实验

为了验证本文方法的有效性，分别将 Duke、Market-1501 和 MSMT17 数据集作为源域和目标域测试各个模块对行人重识别效果的影响。本文在基础模型中加入 SPGAN、CASA 及 LR 模块，评估各个模块对行人重识别性能的影响。先分别进行测试，再进行模块的组合测试，得到的实验结果如表 3 ~4 所示。

表 3 不同模块在 Market-1501 和 Duke 的实验结果(%)

Tab. 3 Experimental results of different modules on Market1501 and Duke(%)

方法	Duke to Market-1501		Market-1501 to Duke	
	α_{mAP}	R_1	α_{mAP}	R_1
Baseline	80.2	89.6	67.8	77.6
Baseline + SPGAN	81.2	93.0	71.0	82.9
Baseline + CASA	81.4	93.4	71.4	83.2
Baseline + LR	81.7	93.2	71.2	83.4
Baseline + SPGAN + CASA	81.8	93.4	71.5	83.3
Baseline + SPGAN + LR	82.1	93.5	71.5	83.6
Baseline + CASA + LR	82.2	93.4	71.5	83.6
ALL	82.5	93.8	71.8	83.7

表 4 不同模块在 MSMT17 的实验结果(%)

Tab. 4 Experimental results of different modules on MSMT17(%)

方法	Duke to MSMT17		Market-1501 to MSMT17	
	α_{mAP}	R_1	α_{mAP}	R_1
Baseline	28.9	59.9	28.1	57.6
Baseline + SPGAN	30.4	61.3	29.5	60.3
Baseline + CASA	30.3	61.5	29.6	60.5
Baseline + LR	30.9	62.2	30.4	62.6
Baseline + SPGAN + CASA	31.4	63.1	31.7	63.7
Baseline + SPGAN + LR	31.8	63.6	32.1	64.2
Baseline + CASA + LR	31.5	63.4	32.0	64.0
ALL	32.9	64.1	33.4	65.8

本文采用生成对抗网络实现源域到目标域的风格迁移，在对比多种生成对抗网络后选择 SPGAN 模型。实验表明：相较于基准模型，SPGAN 能更好地适配目标域的风格分布，验证了风格迁移策略的有效性，可有效降低实际场景的数据标注成本。引入 CASA 注意力模块可精准聚焦行人判别性特征。而 LR 模块可结合行人生成的局部特征和全局特征迭代优化伪标签，抑制标签噪声，生成更平滑的标签，为模型训练提供辨识度更高的特征。实验结果表明：在域自适应中融合上述 3 个模块可以显著提升行人重识别的准确率。

2.5 参数敏感性实验

为验证行人图片分块数量对提取局部特征的影响，本文在 Market-1501 和 Duke 互为源域和目标域的跨域任务下，设置不同分块数量开展多组对比实验，以此验证模型性能，实验结果如图 5 所示。

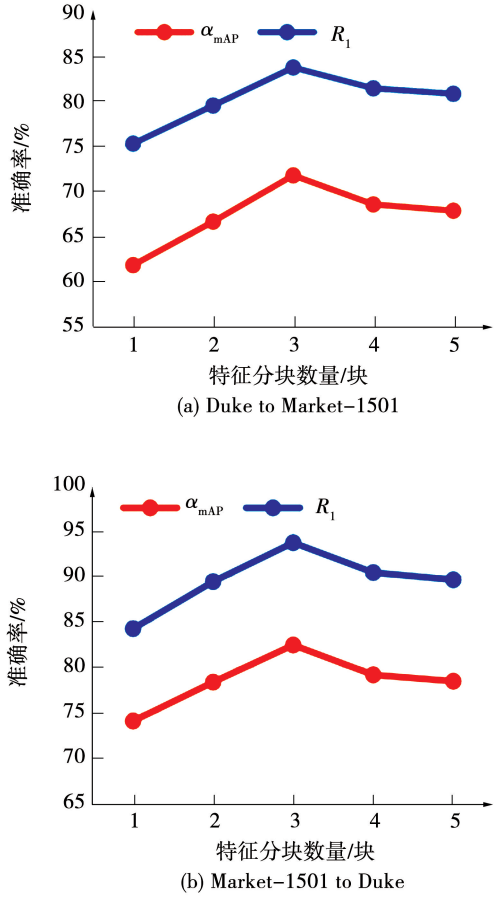
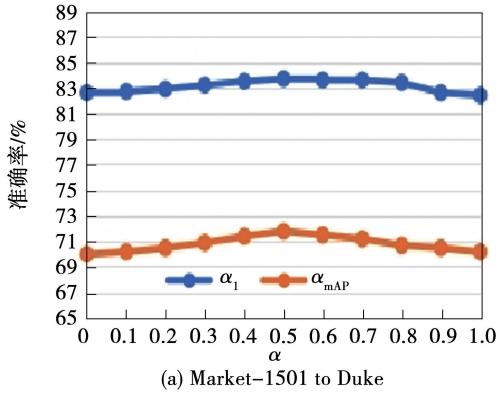


图 5 不同分块值对模型效果对比

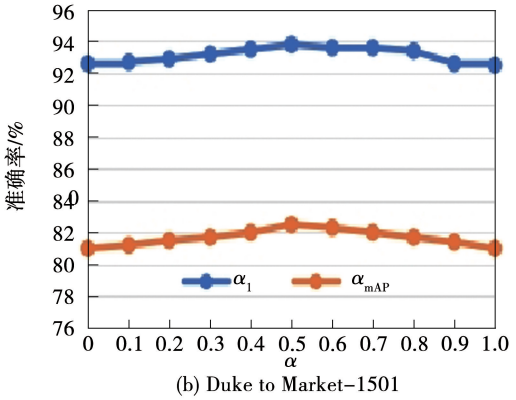
Fig. 5 Comparison of model effects with different block values

由图可见,如果不采取分块措施,模型对于含有判别性的细节信息关注度较低,导致对行人识别的效果降低。如果分块过多,会导致重要的局部特征被切分,起到相反的效果。实验结果表明:大多数图片分为 3 块时,模型可以更好地关注到行人的局部特征。因此,本文将行人图片分为 3 块进行局部特征的提取。

为了评判标签细化模块中加权参数 α 的变化对实验效果的影响,本节采用控制变量法,以 Market-1501 和 Duke 数据集分别作为源域和目标域进行多组实验。实验结果如图 6 所示。由式(5)可知,当 α 取 0 时,模型将忽略全局特征而只关注局部特征,实验结果也验证了这一点。当 α 取 1 时,结果和只关注全局特征的结果一致。经不同 α 值对比,本文将 α 值设为 0.5。



(a) Market-1501 to Duke



(b) Duke to Market-1501

图 6　不同 α 值对模型效果对比

Fig. 6　Comparison of model effect with different α values

2.6　可视化

为了研究本文设计的注意力 CASA(channel and spatial attention)模块对特征提取的影响,采用 Grad-CAM 结合梯度信息和特征图生成热力图,凸显本文模块在预测时关注的重点区域。图 7 为 Resnet50 模型经 CASA 注意力模块增强后的行人特征可视化热力图。由图可见,CASA 模块具有较好的特征提取性能,可以根据行人图片的不同精准锁定判别行人的重点特征。此外,视觉上可以观察到,CASA 模块减少了对行人身后的无关元素和背景的注意,增强了模型对行人信息的聚焦效果。

为证明本文提出的方法对无监督域自适应行人重识别的有效性,本文在 Market-1501、MSMT17 数据集上分别进行了检索结果的可视化,结果如图 8~9 所示。在不同数据集上通过 3 种不同的方法进行了行人重识别匹配,绿色框代表匹配正确,红色框代表匹配错误。由可视化结果可知,识别错误的行人通常与目标行人具有相同的姿态或相同颜色的衣着。同时,背包及衣服上相似的装饰等细节也会影响重识别的效果。本文面向跨域风格差异问题,深度挖掘行人局部特征和全局特征的内在关联,依托风格迁移算法,结合特征关联性完成伪标签细化,提升行人识别准确率。

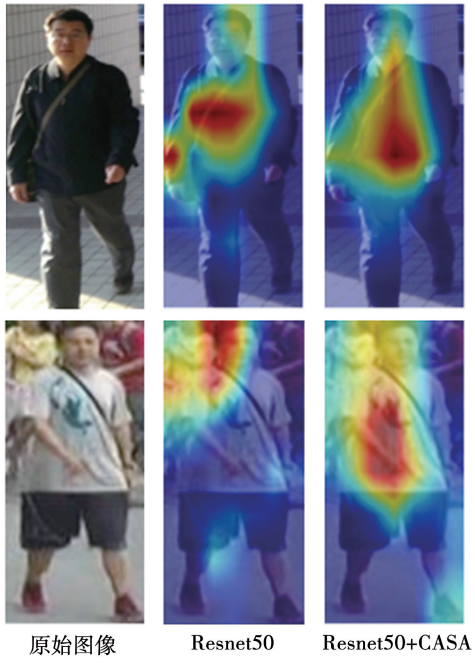


图 7　CASA 模块可视化结果

Fig. 7　CASA module visualization results



图 8 Market-1501 数据集排名前十结果可视化

Fig. 8 Visualization of top-10 retrieval results on Market-1501 dataset



图 9 MSMT17 数据集排名前十结果可视化

Fig. 9 Visualization of top-10 retrieval results on MSMT17 dataset

3 结 语

本文针对域间差异和标签噪声问题,提出了域风格转换和伪标签细化的域自适应行人重识别模型。该模型将域转换方法引入聚类域自适应网络,经实验验证,本文采用 SPGAN 网络完成行人图像风格迁移,有效缓解了真实场景域间分布差异造成的识别性能下降问题。针对聚类伪标签存在噪声的问题,本文引入融合通道与空间注意力的 CASA 模块,并构建分支网络,依托全局特征与局部特征关联实现伪标签细化。实验对比发现:通过局部特征和全局特征互相细化可获得行人的细节信息,并且减少噪声对重识别的影响。本文提出的方法在 Duke 和 Market1501 分别作为源域进行域自适应时,取得了较好的性能,证明了本文方法的有效性,与其他的域自适应行人重识别方法相比达到了更好的重识别效果。接下来的工作将探究更加有效的标签细化方法,获得噪声鲁棒的行人重识别模型。

参 考 文 献:

[1] 张国庆, 杨珊, 汪海蕊, 等. 基于深度学习的多模态行人重识别综述[J]. 南京信息工程大学学报, 2024, 16(4): 437.

ZHANG Guoqing, YANG Shan, WANG Hairui, et al. Multi-modal Person Re-identification Based on Deep Learning: A Review[J]. Journal of Nanjing University of Information Science & Technology, 2024, 16(4): 437.

[2] 陈昊, 张宝华, 日晓琪, 等. 软伪标签和多尺度特征融合的行人重识别[J]. 激光与光电子学进展, 2022, 59(24): 2420001.
CHEN Hao, ZHANG Baohua, RI Xiaoqi, et al. Soft Pseudo-label and Multi-scale Feature Fusion for Person Re-identification[J]. Laser & Optoelectronics Progress, 2022, 59(24): 2420001.

[3] 李琳琳, 孙海龙. 基于 YOLOv8 的林区行人目标检测研究[J]. 森林工程, 2025, 41(1): 138.
LI Linlin, SUN Hailong. Forest Pedestrian Detection Based on Improved YOLOv8[J]. Forest Engineering, 2025, 41(1): 138.

[4] 李晓楠,朱朦,任洪娥,等. 融合多分支与多粒度特征的东北虎重识别[J]. 哈尔滨理工大学学报,2023,28(6):95.
LI Xiaonan,ZHU Meng,REN Honge,et al. Re-Identification of Amur Tigers by Combining Multi-branch and Multi-granularity Features[J]. Journal of Harbin University of Science and Technology,2023,28(6):95.

[5] WANG H, SHEN J, LIU Y, et al. Nformer: Robust Person Re-identification with Neighbor Transformer[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 7297.

[6] SONG L, WANG C, ZHANG L, et al. Unsupervised Domain Adaptive Re-identification: Theory and Practice[J]. Pattern Recognition, 2020, 102: 107173.

[7] 王水涛,张思佳,孙旭菲,等. 融合伪标签生成与数据增强的多标签文本分类[J]. 哈尔滨理工大学学报, 2024,29(6):22.
WANG Shuitao,ZHANG Sijia,SUN Xufei,et al. Multi-label Text Classification by Fusing Pseudo-label Generation and Data Augmentation[J]. Journal of Harbin University of Science and Technology,2024,29(6):22.

[8] HUANG Y, WU Q, XU J S, et al. SBSGAN: Suppression of Inter-domain Background Shift for Person Re-identification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9527.

[9] 李开放,惠冠程,王汝涵,等. 基于生成对抗网络与自校准卷积的行人重识别[J]. 激光与光电子学进展, 2022,59(10):303.
LI Kaifang, HUI Guancheng, WANG Ruhan, et al. Person Re-Identification Based on Generative Adversarial Network and Self-Calibrated Convolution[J]. Laser Optoelectronics Progress, 2022,59(10):303.

[10] ZHANG Y, YAN Y, LU Y, et al. Adaptive Middle Modality Alignment Learning for Visible-infrared Person Re-identification[J]. International Journal of Computer Vision, 2025, 133(4): 2176.

[11] 李成严,郑企森,王昊. 多元化渐进域迁移弱监督实时目标检测[J]. 哈尔滨理工大学学报,2024,29(3):11.
LI Chengyan,ZHENG Qisen,WANG Hao. Domain Diversification Progressive Cross-Domain Weakly-Supervised Real-time Object Detection[J]. Journal of Harbin University of Science and Technology,2024,29(3):11.

[12] ZHANG M, ZHAO X, YAO J, et al. When Noisy Labels Meet Long Tail Dilemmas: A Representation Calibration Method[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 15890.

[13] YIN X, SHI J, ZHANG Y, et al. Robust Pseudo-label Learning with Neighbor Relation for Unsupervised Visible-infrared Person Re-identification[C]//Proceedings of the 32nd ACM International Conference on Multimedia, 2024: 2242.

[14] GE Y, CHEN D, LI H. Mutual Mean-teaching: Pseudo Label Refinery for Unsupervised Domain Adaptation on Person Re-identification[J]. arXiv Preprint arXiv:2001.01526, 2020.

[15] TENG X, LAN L, CHEN D, et al. Relieving Universal Label Noise for Unsupervised Visible-Infrared Person Re-Identification by Inferring from Neighbors[J]. arXiv Preprint arXiv:2412.12220, 2024.

[16] GE Y, ZHU F, CHEN D, et al. Self-paced Contrastive Learning with Hybrid Memory for Domain Adaptive Object Re-id[J]. Advances in Neural Information Processing Systems, 2020, 33: 11309.

[17] ZHENG X, ZHANG Y, LU Y, et al. Semi-supervised Visible-Infrared Person Re-identification via Modality Unification and Confidence Guidance[C]//Proceedings of the 32nd ACM International Conference on Multimedia, 2024: 5761.

[18] 吴蕾,王海瑞,朱贵富,等. 基于多尺度局部特征融合的行人重识别方法[J]. 计算机科学, 2024, 51(S1):506.
WU Lei, WANG Hairui, ZHU Guifu, et al. Person Re-identification Method Based on Multi-scale Local Feature Fusion[J]. Computer Science, 2024,51(S1):506.

[19] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770.

[20] ZHENG L, SHEN L, TIAN L, et al. Person Re-identification Meets Image Search[J]. arXiv Preprint arXiv:

1502.02171, 2015.

[21] ZHENG Z, ZHENG L, YANG Y. Unlabeled Samples Generated By Gan Improve the Person Re-identification Baseline in Vitro[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 3754.

[22] WEI L, ZHANG S, GAO W, et al. Person Transfer Gan to Bridge Domain Gap for Person Re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 79.

[23] 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, 2009: 248.

[24] FAN H, ZHENG L, YAN C, et al. Unsupervised Person Re-identification: Clustering and Fine-tuning[J]. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2018, 14(4): 1.

[25] ZHONG Z, ZHENG L, LUO Z, et al. Invariance Matters: Exemplar Memory for Domain Adaptive Person Re-identification[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 598.

[26] FU Y, WEI Y, WANG G, et al. Self-similarity Grouping: A Simple Unsupervised Cross Domain Adaptation Approach for Person Re-identification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6112.

[27] ZHAI Y, LU S, YE Q, et al. Ad-cluster: Augmented Discriminative Clustering for Domain Adaptive Person Re-identification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 9018.

[28] DAI Y, LIU J, BAI Y, et al. Dual-refinement: Joint Label and Feature Refinement for Unsupervised Domain Adaptive Person Re-identification[J]. IEEE Transactions on Image Processing, 2021, 30: 7815.

[29] ZHENG K, LAN C, ZENG W, et al. Exploiting Sample Uncertainty for Domain Adaptive Person Re-identification[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(4): 3538.

[30] LIN Y, DONG X, ZHENG L, et al. A Bottom-up Clustering Approach to Unsupervised Person Re-identification[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 8738.

[31] LI Y F, ZHU X D, SUN J, et al. Unsupervised Person Re-identification Based on High-quality Pseudo Labels[J]. Applied Intelligence, 2023, 53(12): 15112.

[32] CHEN H, WANG Y H, LAGADEC B, et al. Learning Invariance from the Generated Variance for Unsupervised Person Re-identification[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2023, 45(6), 7494.

[33] 李亚军,张敏,邓洋洋,等. 基于相关性感知和特征重学习的无监督域自适应行人重识别[J]. 激光与光电子学进展,2024,61(14):344.
LI Yajun, ZHANG Min, DENG Yangyang, et al. Unsupervised Domain Adaptive Person Reidentification Based on Relation Awareness and Feature Relearning[J]. Laser Optoelectronics Progress, 2024, 61(14): 344.

[34] 文锐,孔广黔,段迅. 基于可靠性集成的无监督域自适应行人重识别[J]. 计算机应用研究, 2024, 41(4):1228.
WEN Rui, KONG Guangqian, DUAN Xun. Unsupervised Domain Adaptive Person Re-identification Based on Reliability Integration[J]. Application Research of Computers, 2024, 41(4): 1228.

(编辑:温泽宇)