

基于双风格增强的域泛化行人重识别

张铠镛, 谢楚雯, 张嘉慧, 何乐为

(华南师范大学人工智能学院, 广东 广州 510631)

摘要: 行人重识别是在不同、非重叠的摄像头视图中匹配并识别同一人员。现有的行人重识别方法大多基于有监督学习, 对于单域表现出色, 但在跨域应用时性能下降明显。为此, 提出一种基于双风格增强的多源域泛化行人重识别方法, 以提升模型在不同域之间的泛化能力。具体而言, 首先, 引入风格不确定性增强, 通过随机噪声模拟多样化的风格变化以增强数据。其次, 为避免单一风格不确定性增强可能导致原始图像风格特征丢失, 设计一种风格记忆模块, 用于存储图像的原始风格特征, 并基于存储的风格特征增强数据, 从而在数据增强的同时保持图像完整性。最后, 提出一种双风格损失函数, 结合原始风格和增强风格的特征, 进一步优化模型泛化性能。在多个行人重识别数据集上开展实验, 结果显示, 所提方法在未知域上的泛化能力显著提升, 其 mAP 较当前已有方法提高了 0.6%~1.9%。

关键词: 行人重识别; 域泛化; 风格增强; 元学习; 实例归一化

DOI: 10.11907/rjdk.241662

扫描二维码阅读全文:



中图分类号: TP391.41; TP18

文献标识码: A

文章编号: 1672-7800(2026)002-0001-07

Domain Generalization for Pedestrian Re-identification with Dual-Style Augmentation

ZHANG Kaiyong, XIE Chuwen, ZHANG Jiahui, HE Lewei

(School of Artificial Intelligence, South China Normal University, Guangzhou 510631, China)

Abstract: Pedestrian re-identification aims to match and identify the same person across different, non-overlapping camera views. Existing pedestrian re-identification methods are mostly based on supervised learning, which performs well in a single domain but suffers a significant performance drop when applied across domains. To address this issue, this paper proposes a multi-source domain generalization person re-identification method based on dual-style augmentation to enhance the model's generalization capability across different domains. Specifically, we first introduce style uncertainty augmentation, which uses random noise to simulate diverse style variations for data augmentation. Meanwhile, to prevent the loss of original image style features caused by single-style uncertainty augmentation, we design a style memory module to store the original style features of the images and enhance the data based on the stored style features, thereby maintaining the integrity of the images during augmentation. Finally, a dual-style loss function is proposed, which combines the features of the original and augmented styles to further optimize the model's generalization performance. Experimental results on multiple pedestrian re-identification datasets demonstrate that the proposed method significantly improves the generalization capability in unseen domains, with the mAP increasing by 0.6% to 1.9% compared with the existing methods.

Key Words: pedestrian re-identification; domain generalization; style augmentation; meta learning; instance normalization

0 引言

行人重识别也称行人再识别, 是一项计算机视觉任

务。其主要目标是确定一个特定的人是否出现在由不同摄像机拍摄到的视频中, 查询线索可以是图像、视频序列, 甚至是文本描述, 旨在弥补固定摄像头的视觉局限, 并可与行人检测、行人跟踪技术相结合^[1]。早期研究将行人重

收稿日期: 2024-12-23

基金项目: 国家自然科学基金青年科学基金项目(52308250); 广东省高等教育杰出青年人才基金项目(2022kQNCX017); 广东省基础与应用基础研究项目(2021a1515011171)

作者简介: 张铠镛(1999-), 男, 华南师范大学人工智能学院硕士研究生, 研究方向为行人重识别; 谢楚雯(2003-), 女, 华南师范大学人工智能学院学生, 研究方向为计算机视觉; 张嘉慧(2004-), 女, 华南师范大学人工智能学院学生, 研究方向为计算机视觉; 何乐为(1993-), 男, 博士, 华南师范大学人工智能学院副研究员、硕士生导师, 研究方向为基于计算机视觉的 3D 打印实时优化系统、面向工业应用的人工智能服务平台。本文通讯作者: 何乐为。

识别视作多分类问题,主要依赖于全局特征。随后,为了解决遮挡问题,研究者提出结合全局特征和局部特征的方法,以提取更具辨识度的特征表示。随着生成对抗网络和注意力机制的引入,行人重识别的图像提取方法得到了显著改进,进一步提升了模型鲁棒性和泛化能力。同时,特征度量学习算法的改进也在持续进行中,旨在通过更优秀的度量学习损失函数使算法能够更准确地识别相同的人并区分不同的人。

然而,在现实世界中,一个地点(如购物中心、机场)的摄像头系统与另一个地点的系统往往存在显著差异。当模型从一个域训练好并期望在另一个完全不同的域中部署时,性能通常会大幅下降。无监督领域自适应(Unsupervised Domain Adaptation, UDA)的方法试图解决这一问题,它们通常包含两个主要步骤:域对齐和特征学习。域对齐旨在通过各种统计技术减少源域和目标域数据分布之间的差异,而特征学习则关注于提取两个域中都有效且鲁棒的特征表示。

然而,现有的UDA方法无法处理更具挑战性但实际的应用场景,即目标数据集在训练过程中看不到,因为有时无法收集到足够的目标域训练数据,而且在这些未标记样本上调整模型非常耗时。鉴于此,本文提出了一个更关键的问题,即如何将现有数据上学习到的模型推广到看不见的领域。

域泛化(Domain Generalization, DG)是近年来为了应对该挑战而出现的一个研究方向,通常指在不知道目标域的分布的情况下,将通过源域的训练而得到的模型直接迁移部署到未知目标域,也能保证模型性能不会出现严重下降。领域自适应行人重识别的目的是利用现有的训练源域在给定的目标域上实现性能最大化。两者的区别在于,领域自适应行人重识别可以接触到目标域的数据,而域泛化行人重识别在训练过程中无法看到这些数据。这使得域泛化行人重识别比域适应行人重识别更具挑战性,但在实际应用中更有价值,也更符合现实应用的需要。然而,由于行人重识别数据集的特殊性,每个类别仅存在于一个域中,所有身份缺乏域多样性,从而引发了一个关键问题——不同域之间的身份标签并不互通,这使得DG Re-ID对于图像分类中部署的现有DG方法而言极为困难。

为了解决DG Re-ID所面临的这一挑战,最简单和最直接的方法是增强数据源域分布,以更好地利用源域数据。文献[2]提出了一种高效、保留内容的特征级增强方法,即风格不确定性增强(Style-uncertainty Augmentation, SuA)。该技术在每次迭代中将源数据的风格从容易到难随机化,以增强训练域的多样性。具体而言,每个实例的每个通道内的空间维数的平均值和方差被不同的高斯噪声扰动,以随机化其风格。

风格不确定技术虽然能在一定程度上提升模型对不同域的适应性,但它忽略了风格信息在跨域识别中的潜在

价值。这种方法通过在训练中引入噪声以提高泛化能力,却可能导致有用的风格特征丢失,从而影响到模型在实际应用中的表现力。因此,为了保留风格原有的特征信息,本文引入风格记忆增强方法(Style memory Augmentation, SmA),并且与风格不确定性增强相结合,提出了双风格增强模块(Dualstyle)。该模块旨在更有效地捕获和融合来自不同域的风格特征,而不是简单地通过添加噪声以实现随机化。它利用深度学习网络中间层的风格特征,并通过一种特别设计的融合策略,有选择地整合这些特征。双风格增强模块能够通过随机化训练样本的风格特征以增强模型泛化能力,同时保持了图像内容的完整性。为进一步增强数据,本文提出了专为该模块设计的双风格增强损失函数,以更好地优化模型。为了进一步加强双风格模块的实用性,本文引入了一种分阶段元学习训练策略,该策略逐步从简单阶段过渡到复杂阶段,从而在传统的单阶段元学习基础上有所拓展。

1 相关工作

行人重识别的早期研究主要集中在基于手工特征和距离度量的方法上,例如Zheng等^[3]提出概率相对距离比较方法,结合局部和全局特征,以提高识别准确性。随着深度学习技术的快速发展,研究重心逐渐转向神经网络方法。Zheng等^[4]将行人重识别视为一个多类别问题,提出判别性身份嵌入方法,将每个行人视为独立的类别。Sun等^[5]则通过统一分割策略,将行人图像划分为多个部分,以提取局部特征,从而更好地捕捉细粒度信息,提升识别效果。为了应对模型可迁移性和数据标注困难的问题,无监督学习(Unsupervised Learning, USL)和无监督域适应(Unsupervised Domain Adaptation, UDA)方法逐渐成为研究热点。Fan等^[6]提出一种包含聚类和微调的策略,以有效利用未标记数据,提升模型泛化能力。Chen等^[7]则通过实例引导的上下文渲染实现了域适应,成功应对了跨域性能下降的问题,从而显著提升了模型在新域中的表现。这些研究为行人重识别在跨域泛化和无监督学习方面提供了有力支持。

尽管在无监督域自适应(UDA)和无监督学习(USL)方法上取得了一些进展,但在实际应用中,尤其是在训练期间无法获取目标数据集的情况下,这些方法仍面临诸多挑战。这主要是因为难以收集到足够数量的目标域数据,且利用这些未标记样本对模型进行调整需要耗费大量时间。这一问题引发了一个关键的研究问题,即如何将从现有数据中学到的模型推广到未见的域。

域泛化(DG)的出现解决了该问题。域泛化旨在通过从多个源域中学习带标记的数据以应对跨域挑战,从而能够直接应用于任何未见的域。以往的域泛化方法大多假设源域和目标域共享相同的标签空间,即同质环境^[8-9]。

然而, 针对行人重识别的域泛化(DG-ReID)解决了一个更复杂的问题, 即域之间没有身份重叠。源域标签集的多样性易构成重大挑战, 因为每个身份在其域中都是独特的。这种复杂性使得行人重识别的域泛化特别困难, 尤其是在应用现有为图像分类设计的域泛化方法时。

Choi等^[10]引入了一种新颖的行人重识别框架, 称为元批实例归一化(MetaBIN), 该框架通过元学习优化归一化层, 模拟各种泛化场景以增强模型泛化能力。Zhou等^[11]提出一种方法, 使用带标记的合成数据集和未标记的真实世界数据集训练一个通用模型, 消除了手动标注的需要, 并适用于多样化的真实世界数据集。Li等^[12]开发了渐进域扩展网络(PDEN), 通过域扩展和表示学习子网络的联合学习以模拟未知域的变化, 从而增强泛化能力。Dai等^[13]引入了一种相关性感知的专家混合(RaMoE)模型, 该模型利用投票机制和去相关损失函数动态集成多样化特征, 从而提高了泛化能力。这些研究旨在开发在未知新环境中保持鲁棒性的行人重识别系统, 并在训练阶段需要较强的泛化能力。Liao等^[14]开发了QAConv, 使用查询自适应内核进行可解释的特征匹配, 并在位置或视角变化下增强泛化能力。Liao等^[15]还提出了一种图采样(GS)方法以用于深度度量学习, 通过在每个epoch开始时构建最近邻图以形成一个类及其最近邻的minibatch, 用于学习具有挑战性的样本。Jin等^[16]引入了SNR, 使用实例归一化去除风格变化, 并通过重新整合身份相关特征以保留判别信息, 从而提高识别能力。并且, 利用双重因果损失以分离身份相关和不相关特征。Song等^[17]提出了DIMN, 即一种学习人类图像与其身份分类器之间映射的方法, 使用一次性生成分类器, 通过在不同的训练任务之间共享一个内存库以保持可扩展性和识别能力, 并匹配目标域中任意数量的身份。

2 方法细节

2.1 域泛化行人重识别问题设置

在DG-Reid中, 给定源域 $D = \{D_1, D_2, \dots, D_S\}$, 将 $D_1, D_2, \dots, D_{S-1}$ 作为训练域, D_S 作为测试域。并且, 每一个域都有自己的标签域 $D_k = \{(x_i^k, y_i^k)\}_{i=1}^{N_k}$, N_k 是域 D_k 内实例的数量。每一个样本 $x_i^k \in X_k$ 都被标记为一个身份 $y_i^k \in Y_k = \{1, 2, \dots, H_k\}$, H_k 为域 D_k 内身份的数量, 即类别数。在训练阶段, 模型需要在训练集上进行训练, 在测试阶段, 模型会在看不见的域 D_S 上进行测试。

2.2 实例归一化

在风格迁移增强技术方面, 实例归一化(Instance Normalization, IN)操作通过对具有特定均值和标准差的实例进行归一化, 已被广泛用于去除模型中的图像样式^[18]。现给定一个特征图 $x \in R^{C \times H \times W}$, 其中 C, H, W 分别表示通道、高度和宽度的值。IN的公式如式(1)所示。

$$\text{IN}(x) = \gamma \frac{x - \mu(x)}{\sigma(x)} + \beta \quad (1)$$

其中, $\gamma, \beta \in R^C$ 是可以学习的参数, $\mu(x), \sigma(x) \in R^C$ 表示单独空间通道计算的平均值和标准差, 如式(2)和式(3)所示。

$$\mu(x)_c = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W x_{c,h,w} \quad (2)$$

$$\sigma(x)_c = \sqrt{\frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W (x_{c,h,w} - \mu(x)_c)^2} \quad (3)$$

其中, $c \in C$ 。许多研究已经证明, $\mu(x), \sigma(x)$ 可以编码特征的样式信息。

2.3 风格不确定性增强

实例归一化已被证实对于改进模型泛化性有效。在此基础上, 文献[2]提出一种随机高斯噪声以增强数据的方法, 如图1所示(右侧图)。该方法在训练过程中随机化数据实例的风格信息, 即在标准化输入张量后, 用高斯噪声扰动其在空间维度上的均值和方差。其公式如式(4)所示。

$$x' = (\sigma(x) + \omega_2) \frac{x - \mu(x)}{\sigma(x)} + (\mu(x) + \omega_1) \quad (4)$$

其中, $\omega_1 \sim N(\mu_1, \sigma_1), \omega_2 \sim N(\mu_2, \sigma_2)$ 采样自正态分布。这种即插即用的风格不确定性增强(Style-uncertainty Augmentation, SuA)模块, 通过增加源域的多样性, 帮助模型学习如何处理不同域中的数据, 从而在新的、未知的域上实现更好的性能。然而, 在增加源域多样性的同时, SuA通过扰动标准化特征的方式, 导致部分有助于区分不同域的特定风格信息丢失。这种丢失可能削弱模型对域特异性的感知能力, 并削弱目标域的表达能力, 从而产生有限的泛化性能。

2.4 风格记忆增强

为了弥补这一不足, 本文提出了双风格增强模块(Dualstyle Augmentation), 它将SuA方法与风格记忆增强方法相结合。风格记忆增强模块(Style memory Augmentation, SmA)旨在记住并应用来自训练数据的风格模式, 使模型在数据增强时尽可能保留域的特异性。通过将SuA与风格记忆整合, 双风格增强模块利用了两种方法的优势: SuA通过随机性增强鲁棒性, 而风格记忆则具有保留和利用多样风格特征的能力。

如图1所示(左侧图), 先用瓦氏距离(Wasserstein Distance)估计输入图像 x_k 和第 k 个风格记忆(M_k^μ, M_k^σ)的样式分布差异, 公式如式(5)所示。

$$d_k = (\mu(x)_c - M_k^\mu)^2 + \sigma(x)_c^2 + M_k^{\sigma^2} - 2\sigma(x)_c M_k^\sigma \quad (5)$$

其中, d_k 表示当前图像 x_k 和第 k 个源域之间的分布距离, 然后再用 d_k 的倒数表示 x_k 和第 k 个风格记忆之间的相似性, 公式如式(6)所示。

$$\lambda_k = \frac{\exp(1/(1 + d_k))}{\sum_{k=1}^K \exp(1/(1 + d_k))} \quad (6)$$

其中, 利用Softmax使得 $\lambda = \{\lambda_k | k = 1, 2, \dots, K\}$ 的和等于

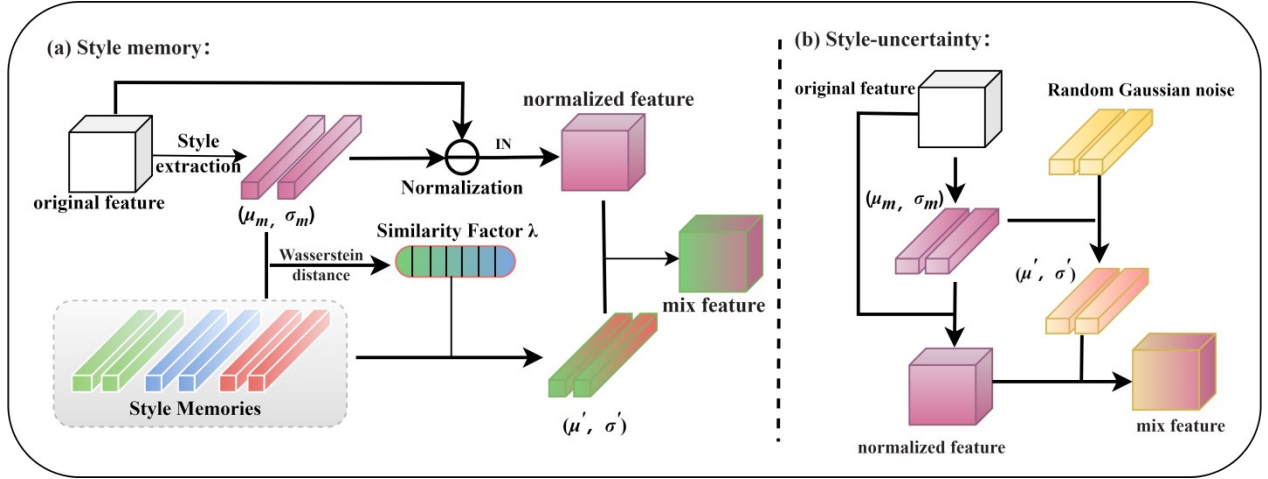


Fig. 1 The frame work of style memory augmentation and style-uncertainty augmentation

图1 风格记忆增强和风格不确定性增强架构

1, 通过得到的风格相似性 λ 即可算出映射后的样式信息, 公式如式(7)所示。

$$\mu(x)_k^* = \sum_{k=1}^K \lambda_k M_k^\mu, \sigma(x)_k^* = \sum_{k=1}^K \lambda_k M_k^\sigma \quad (7)$$

最后用式(1)计算出最终增强的特征, 如式(8)所示。

$$x' = \sigma(x)_k^* \frac{x - \mu(x)}{\sigma(x)} + \mu(x)_k^* \quad (8)$$

在训练阶段, 利用动量更新策略, 实现风格记忆更新, 公式如式(9)所示。

$$M_k^\mu = \alpha M_k^\mu + (1 - \alpha) \mu(x)_c \quad (9)$$

$$M_k^\sigma = \alpha M_k^\sigma + (1 - \alpha) \sigma(x)_c \quad (10)$$

其中, $\alpha \in [0, 1]$ 表示动量参数。在开始训练时, 对风格记忆进行随机初始化。其中, M_k^μ 和 M_k^σ 分别用零均值分布和单均值分布进行初始化。该模块能够提取源域的风格统计量, 并将其有效地存储为风格记忆。经过风格记忆增强后, 将增强的特征 x' 输入到网络的下一层。然而, 随着层次的深入, 风格信息越来越松散, 而语义信息起到更重要的作用。因此, 风格增强只用于网络的前两层, 以获得最佳的泛化性能。

2.5 双风格增强损失

在训练过程中, 双模块使用风格不确定性增强(SuA)和风格记忆增强(SmA)以增强特征。SuA引入随机高斯噪声以增加风格多样性, 而SmA基于风格记忆动态调整特征, 这种组合提供了广泛的风格覆盖范围和改进的泛化能力。SuA作为一个轻量级的模块, 可以轻松集成到任何网络中, 几乎不增加训练负担; SmA通过动态更新风格原型, 进一步增强了模型对风格多样性的适应性。这种方法不仅提高了跨域泛化效果, 还增强了对风格变化的鲁棒性, 使模型能够更好地处理来自未见过域的数据。

模块计算原始特征与增强特征之间的距离, 原始特征集 $F = \{f_1, f_2, f_3, \dots, f_B\}$ 与风格不确定性增强(SuA)特征 $F^{SuA} = \{f_1^{SuA}, f_2^{SuA}, f_3^{SuA}, \dots, f_B^{SuA}\}$ 以及风格记忆增强(SmA)特征 $F^{SmA} = \{f_1^{SmA}, f_2^{SmA}, f_3^{SmA}, \dots, f_B^{SmA}\}$ 之间的距离通过指定的

损失函数进行量化。公式如式(11)所示。

$$L_{DS} = \frac{1}{B^2} \|d(F, F^{SuA}) - d(F, F^{SmA})\|_2 \quad (11)$$

其中, B 表示特征批次数量, $d(\cdot, \cdot)$ 表示原始特征和增强特征之间的欧几里得距离, 公式如式(12)所示。

$$d(F, F') = \sum_{i=1}^B \sum_{j=1}^B \sqrt{(f_i - f'_j)^2} \quad (12)$$

双风格损失评估模型用于对其在保留原始特征的同时整合新风格信息的能力进行评估, 从而确保模型在各种风格下都能实现有效适应。

2.6 训练设置

行人重识别可以被定义为分类任务或检索任务。分类任务是学习一个人的特征表征, 从而对一个人的身份进行分类, 其更加关注独特的身份信息。同时, 检索任务试图减小正样本对的距离, 增加负样本对的距离, 以匹配同一身份的人。本文通过识别损失和三元组损失以实现上述两种互补任务, 配合双风格增强损失, 组成网络总损失, 计算方法如式(13)所示。

$$L_{total} = L_{id} + L_{tri} + L_{DS} \quad (13)$$

其中, L_{id} 、 L_{tri} 、 L_{DS} 分别表示识别损失、三元组损失和双风格增强损失。为了更好地适配双风格增强模块的使用, 采用一种渐进式的元学习策略, 通过多阶段的训练过程模拟人类的学习机制, 逐步提高对未知域的泛化能力, 具体流程如图2所示。直接汇总所有增强数据训练模型可能会导致泛化能力不足, 原因是缺乏多样化域的协同效应。虽然可以基于元学习, 采用训练—测试流程以解决该问题, 但其一阶段的训练策略可能容易忽略那些简单的域或因为困难的域而导致训练不稳定。

2.6.1 识别损失

通过识别损失监督训练过程, 以学习具有强判别力的行人特征表示。考虑到当多个域用于训练时, 传统分类器参数的数量会急剧增加, 因而采用了一种更高效的无参数记忆分类器^[19]。

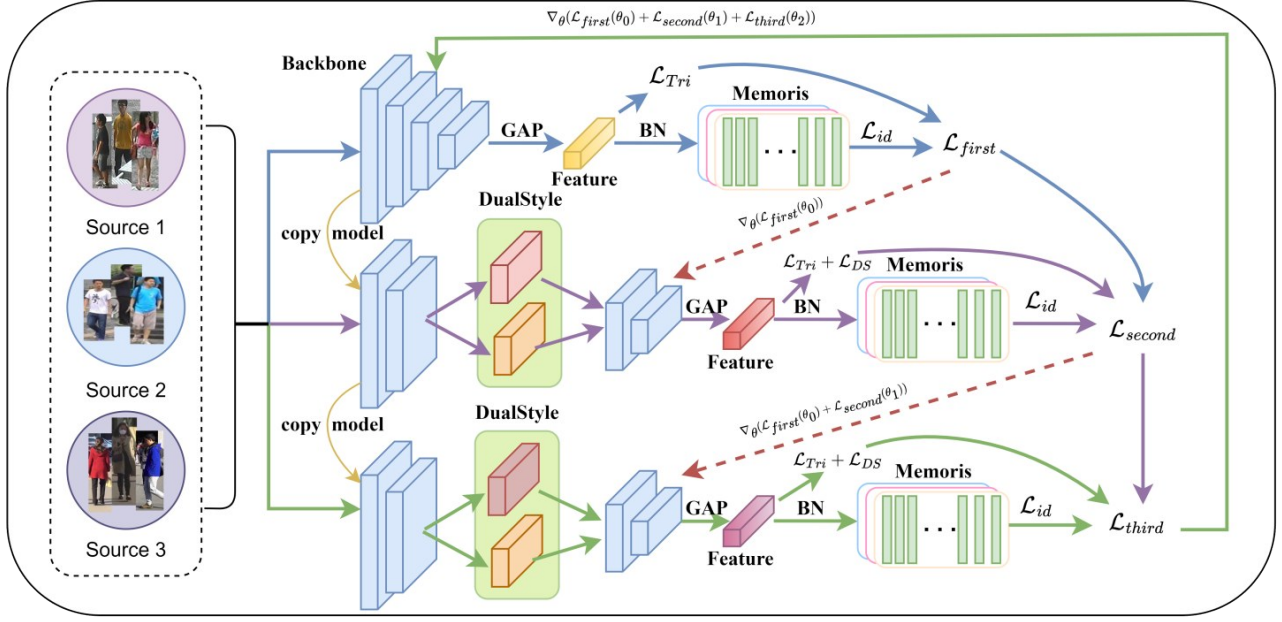


Fig. 2 The training flow

图 2 训练流程

具体而言,给定特征,通过计算特征与每个身份质心的相似性以实现无参数识别损失。具体公式如式(14)所示。

$$L_{id} = -\log\left(\frac{\exp(M[i]_T f(x_i)/\tau)}{\sum_{k=1}^{n_T} \exp(M[k]_T f(x_i)/\tau)}\right) \quad (14)$$

其中, T 和 n_T 分别表示源域 T 和该域中身份总数量, τ 是控制分布尺度的温度因子。 M 表示身份中心的记忆库, 训练开始时通过初始图像特征进行初始化, 并通过动量更新公式更新。公式如式(15)所示。

$$M[k] \leftarrow \varphi \cdot M[k] + (1 - \varphi) \cdot \frac{1}{|v_k|} \sum_{x_i \in v_k} f(x_i) \quad (15)$$

其中, v_k 和 $|v_k|$ 表示第 K 个身份的样本及其样本数量, $\varphi \in [0, 1]$ 控制更新速率。

2.6.2 三元组损失

在实验中,基于批处理的三元组损失结合正样本对和负样本对,对特征进行相似性学习。给定皮革身份和每个身份 Q 张的批次,三元组损失可表示如式(16)所示。

$$L_{tri} = \sum_{i=1}^P \sum_{j=1}^Q [d_{i,j}^{pos} - d_{i,j}^{neg} + \delta]_+ \quad (16)$$

其中, $d_{i,j}^{pos}$ 和 $d_{i,j}^{neg}$ 分别表示正样本对和负样本对之间的欧几里得距离, $[\cdot]_+ = \max(0, \cdot)$ 表示取正值, δ 为边际值。

3 实验与结果分析

如表 1 所示,实验包括 4 个主要数据集: Market-1501 (M)^[20]、DukeMTMC-reID (D)^[21]、CUHK03 (C)^[22] 和 MSMT17 (MS)^[23]。所有数据集以 3:1 的比例被分成训练集和目标集。

Table 1 Statistic of the datasets

表 1 数据集

数据集	训练集		测试集	
	图像/张	身份/个	图像/张	身份/个
Market-1501	12 936	751	19 732	750
DukeMTMC-reID	16 522	702	17 661	1 110
CUHK03	7 365	767	5 332	700
MSMT17	32 621	1041	82 161	3 060

3.1 实验细节

实验中,在 CUDA11.7 环境下利用 PyTorch 对模型进行训练和测试,显卡为 NVIDIA GeForce RTX 3090 24 G。使用 ResNet50^[24] 作为特征提取器的主干,图像尺寸设置为 256×128,批处理大小为 64,数据预处理使用数据增强技术,如随机翻转和裁剪。在损失函数方面,三元组损失的边际 δ 等于 0.3。对于双风格样式设置,使用高斯噪声作为随机噪声,均值为 0,方差在第二阶段训练和第三阶段训练分别为 0.1 和 0.15。在训练阶段,采用 Adam 优化器,权重衰减为 0.000 5,初始学习率设置为 0.000 35。

3.2 评价指标

使用平均精度均值 (mAP) 和 Rank-1 精度作为模型性能评价指标。其中, mAP 是目标检测和行人重识别等任务中的常用指标,衡量模型在预测不同类别目标时精确性与召回率之间的平衡。Rank-1 精度则表示在检索结果中排名第一的匹配样本是否与查询样本属于同一类别,衡量模型在检索任务中首个候选样本的准确性,是评估识别系统性能的重要指标。具体计算公式分别如式(17)一式(19)所示。

$$AP = \int_0^1 p(r) dr \quad (17)$$

$$mAP = \frac{\sum_{i=1}^C AP_i}{C} \quad (18)$$

$$Rank - 1 = \frac{N_{Top1}}{|Q|} \quad (19)$$

式(17)中, $p(r)$ 表示为精确率与召回率的平滑曲线。式(18)中, C 表示类别总数, AP_i 表示类别 i 的平均精度。式(19)中, N_{Top1} 表示在所有查询中, 排名第一的匹配样本与查询样本属于同一类别的数量, $|Q|$ 表示查询样本总数。

3.3 对比实验结果

表2展示了双风格增强方法与当前最先进方法(SOTA)比较结果, 这些方法包括 QAConv^[14]、CBN^[25]、SNR^[16]、DAML^[26]、CamStyle^[27]、M3L^[19]、MetaBIN^[10]、SuA-SpML^[2]、DSU^[28]、QAconv-GS^[15]、MixNorm^[29] 和 MSF-Net^[30]。可以看出, 双风格增强方法在 Rank-1 精度和平均精度均值(mAP)方面普遍优于现有的多域泛化方法。

在 Market-1501 数据集中, 双风格增强方法在 mAP 上仅比当前最高结果提高了 0.9%, 在 Rank-1 上也只提升了 0.4%, 表现相对一般。同样, 在 DukeMTMC-reID 数据集中, 双风格增强方法在 mAP 上领先 0.6%, Rank-1 提升了 0.8%, 虽然有提升, 但不算显著。

在 MSMT17 数据集中, 双风格增强方法效果非常显著。该方法在 mAP 上提升了 1.9%, 而在 Rank-1 上提升了 1%, 显示出其在处理更大规模和更复杂数据集时的强大性能。

此外, 在 CUHK17 数据集中, 双风格增强方法在 mAP 上提升了 1.1%, Rank-1 也有相应提升。这进一步证明了该方法在不同环境和条件下的有效性和鲁棒。

Table 2 Comparison of the experimental results with state-of-the-art(SOTA) methods

表2 与SOTA方法对比实验结果 (%)								
Method	MS+D+C→M		MS+M+C→D		M+D+C→MS		MS+M+D→C	
	mAP	R1	mAP	R1	mAP	R1	mAP	R1
QAConv	39.5	68.6	43.4	64.9	10.0	29.9	19.2	22.9
CBN	47.3	74.7	50.1	70.0	15.4	37.0	25.7	25.2
SNR	48.5	75.2	48.3	66.7	13.8	35.1	29.0	29.1
DAML	49.3	75.1	47.6	66.5	12.6	32.2	29.3	29.8
CamStyle	48.7	75.1	47.1	66.0	11.7	31.1	28.6	29.3
M3L	51.7	76.5	48.2	67.1	13.1	32.0	30.9	31.9
MetaBIN	52.7	78.8	50.4	68.5	14.8	36.1	29.3	29.4
SuA-SpML	56.6	81.0	51.2	70.8	18.2	42.8	32.0	33.1
DSU	50.4	75.2	49.4	66.6	13.4	32.0	30.1	31.1
QAconv-GS	51.0	79.2	50.6	69.1	16.8	47.1	23.8	24.0
MixNorm	51.4	78.9	49.9	70.8	19.4	47.2	29.0	29.6
MSF-Net	50.4	75.4	49.5	68.4	11.7	30.6	32.0	36.7
DualStyle(ours)	57.4	81.4	51.8	71.6	21.3	48.2	33.1	33.8

总体而言, 虽然双风格增强方法在 Market-1501 和 DukeMTMC-reID 数据集中提升不大, 但在 MSMT17 和 CUHK17 数据集中表现突出, 提升显著, 超过了 1%。这些结果表明, 双风格增强方法不仅能够提升识别精度和准确性, 还具有较强的跨域泛化能力, 特别是在更复杂的数据集上展示出其优势, 是当前更为先进的多域泛化方法。

3.4 消融实验

如表3所示, 消融实验剖析了本文方法中每个部分的贡献, 特别强调了风格记忆模块和双风格增强损失的有效性。本文 Baseline 采用文献[35]的代码和文献[2]的复现。

Table 3 Ablation study result

表3 消融实验结果 (%)				
method	MS+D+M→M		M+D+C→MS	
	mAP	R1	mAP	R1
Baseline	55.5	80.3	17.2	41.3
Baseline + SmA	56.9	81.1	19.5	45.9
Baseline + SmA + L_{DS}	57.4	81.4	21.3	48.2

实验表明, 两个模块显著提升了 Baseline 的性能, 在多个数据集测试中取得了显著提升。具体而言, 在 Market-1501 数据集测试中, 加入风格记忆模块使得 mAP 从 55.5% 增加到 57.4%, Rank-1 准确率从 80.3% 提升到 81.1%。对于 MSMT17 数据集, 提升更为显著, mAP 从 17.2% 提升至 19.5%, Rank-1 准确率从 41.3% 提升至 45.9%。此外, 在 MSMT17 数据集中, 加入 L_{ds} 后, mAP 额外提升了 1.8%, Rank-1 准确率提升了 2.3%; 在 Market-1501 数据集中, mAP 额外提升了 0.5%, Rank-1 准确率提升了 0.3%。这些结果进一步证明了双风格模块在规范和实现特征风格多样化方面的能力, 对于提高模型应对域移位的抗性起到了至关重要的作用。

4 结语

本文提出了一种基于双风格增强的多元域泛化行人重识别方法, 旨在提高模型在未知域上的泛化能力。双风格增强方法结合了风格不确定性(SuA)和风格记忆增强(SmA)两种技术, 一方面通过在训练过程中引入随机高斯噪声以增强数据的复杂性和多样性, 另一方面引入动态记忆风格特征, 以保留源域的风格, 使模型能够更好地适应不同语种的风格变化。实验表明, 双风格增强方法在多个数据集上均有提升, 尤其在 MSMT17 和 CUHK03 上表现突出。

未来, 将进一步优化风格记忆模块设计, 并探索更复杂的数据增强策略, 以进一步提高模型泛化能力。此外, 可尝试将双风格模块应用于其他计算机任务, 如图像分类和目标检测, 以验证其通用性和有效性。

参考文献:

- [1] WU D, ZHENG S J, ZHANG X P, et al. Deep learning-based methods for person re-identification: a comprehensive review [J]. Neurocomputing, 2019, 337: 354-371.
- [2] ZHANG L, LIU Z P, ZHANG W S, et al. Style uncertainty based self-paced meta learning for generalizable person re-identification [J]. IEEE Transactions on Image Processing, 2023, 32: 2107-2119.
- [3] ZHENG W S, GONG S G, XIANG T, et al. Person re-identification by probabilistic relative distance comparison [C]//Colorado Springs: Confer-

- ence on Computer Vision and Pattern Recognition, 2011: 649–656.
- [4] ZHENG L, ZHANG H H, SUN S Y, et al. Person re-identification in the wild[C]//Honolulu: Conference on Computer Vision and Pattern Recognition, 2017: 1367–1376.
- [5] SUN Y F, ZHENG L, YANG Y, et al. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline)[C]//Munich: European Conference on Computer Vision, 2018: 480–496.
- [6] 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 (TOMM), 2018, 14(4): 1–18.
- [7] CHEN Y B, ZHU X T, GONG S G. Instance-guided context rendering for cross-domain person re-identification[C]//Long Beach: Conference on Computer Vision and Pattern Recognition, 2019: 232–242.
- [8] CHATTOPADHYAY P, BALAJI Y, HOFFMAN J. Learning to balance specificity and invariance for in and out of domain generalization[C]//Glasgow: European Conference on Computer Vision, 2020: 301–318.
- [9] KHOSLA A, ZHOU T H, MALISIEWICZ T, et al. Undoing the damage of dataset bias[C]//Munich: European Conference on Computer Vision (ECCV), 2012: 158–171.
- [10] CHOI S K, KIM T Y, JEONG M K, et al. Meta batch-instance normalization for generalizable person re-identification[C]//Online: Conference on Computer Vision and Pattern Recognition, 2021: 3425–3435.
- [11] ZHOU K Y, YANG Y X, HOSPEDALES T, et al. Learning to generate novel domains for domain generalization[C]//Glasgow: European Conference on Computer Vision, 2020: 561–578.
- [12] LI L, GAO K, CAO J, et al. Progressive domain expansion network for single domain generalization[C]//Online: Conference on Computer Vision and Pattern Recognition, 2021: 224–233.
- [13] DAI Y X, LI X T, LIU J, et al. Generalizable person re-identification with relevance-aware mixture of experts[C]//Online: Conference on Computer Vision and Pattern Recognition, 2021: 16145–16154.
- [14] LIAO S C, SHAO L. Interpretable and generalizable person re-identification with query-adaptive convolution and temporal lifting[C]//Glasgow: European Conference on Computer Vision, 2020: 456–474.
- [15] LIAO S C, SHAO L. Graph sampling based deep metric learning for generalizable person re-identification[C]//New Orleans: Conference on Computer Vision and Pattern Recognition, 2022: 7359–7368.
- [16] JIN X, LAN C L, ZENG W J, et al. Style normalization and restitution for generalizable person re-identification[C]//Seattle: Conference on Computer Vision and Pattern Recognition, 2020: 3143–3152.
- [17] SONG J F, YANG Y X, SONG Y Z, et al. Generalizable person re-identification by domain-invariant mapping network[C]//Long Beach: Conference on Computer Vision and Pattern Recognition, 2019: 719–728.
- [18] ULYANOV D, ANDREA V, VICTOR L. Instance normalization: the missing ingredient for fast stylization[DB/OL]. <https://arxiv.org/abs/1607.08022>.
- [19] ZHAO Y Y, ZHONG Z, YANG F X, et al. Learning to generalize unseen domains via memory-based multi-source meta-learning for person re-identification[C]//Online: Conference on Computer Vision and Pattern Recognition, 2021: 6277–6286.
- [20] ZHENG L, SHEN L Y, TIAN L, et al. Scalable person re-identification: a benchmark[C]//Santiago: International Conference on Computer Vision, 2015: 1116–1124.
- [21] ZHENG Z D, ZHENG L, YANG Y. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro[C]//Venice: International Conference on Computer Vision, 2017: 3754–3762.
- [22] ZHONG Z, ZHENG L, CAO D L, et al. Re-ranking person re-identification with k-reciprocal encoding[C]//Honolulu: Conference on Computer Vision and Pattern Recognition, 2017: 1318–1327.
- [23] WEI L H, ZHANG S L, GAO W, et al. Person transfer GAN to bridge domain gap for person re-identification[C]//Salt Lake City: Conference on Computer Vision and Pattern Recognition, 2018: 79–88.
- [24] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Las Vegas: Conference on Computer Vision and Pattern Recognition, 2016: 770–778.
- [25] ZHUANG Z J, WEI L H, XIE L X, et al. Rethinking the distribution gap of person re-identification with camera-based batch normalization[C]//Glasgow: European Conference on Computer Vision, 2020: 140–157.
- [26] SHU Y, CAO Z J, WANG C Y, et al. Open domain generalization with domain-augmented meta-learning[C]//Online: Conference on Computer Vision and Pattern Recognition, 2021: 9624–9633.
- [27] ZHONG Z, ZHENG L, ZHENG Z D, et al. Camera style adaptation for person re-identification[C]//Salt Lake City: Conference on Computer Vision and Pattern Recognition, 2018: 5157–5166.
- [28] LI X T, DAI Y X, GE Y X, et al. Uncertainty modeling for out-of-distribution generalization[DB/OL]. <https://arxiv.org/abs/2202.03958>.
- [29] QI L, WANG L, SHI Y H, et al. A novel mix-normalization method for generalizable multi-source person re-identification[J]. IEEE Transactions on Multimedia, 2023, 25: 4856–4867.
- [30] DU H S, HE L B, GUO J T, et al. Meta separation-fusion for generalizable person re-identification[J]. Knowledge-Based Systems, 2024, 284: 111224.

(责任编辑: 孙 娟)