

多因素引导的行人重识别数据增广方法研究



刘志刚^{1,2*}, 张国辉¹, 高月¹, 刘苗苗¹

(1. 东北石油大学 计算机与信息技术学院, 大庆 163318; 2. 黑龙江省石油大数据与智能分析重点实验室, 大庆 163318)

摘要 为解决行人重识别研究领域中标注图像获取困难的问题, 提出一种多因素引导的行人数据增广方法。首先, 在生成器网络中设计了一种局部多尺度引导机制, 通过特征融合抑制生成图像的局部伪影; 其次, 提出了长距离相关性引导机制, 通过外注意力引导生成图像的长距离依赖, 提高生成行人图像的整体视觉质量; 最后, 提出一种抗博弈判别网络, 通过嵌入到生成对抗网络, 从而构建一种三网络稳定博弈架构模型, 增加生成对抗网络训练的稳定性。通过 VIPeR、Market-1501、DukeMTMC-reID 这 3 种不同规模数据集的仿真实验, 结果表明该方法与目前主流方法相比, mAP 与 Rank-1 精度上均有不同程度的提升, 在小规模数据集上的提升较为显著。

关键词 行人重识别; 生成对抗网络; 数据增广; 局部多尺度; 注意力机制

中图分类号 TP39 文献标志码 A DOI 10.12178/1001-0548.2023056

Research on Pedestrian Re-Identification Data Augmentation Method Based on Multi-Factor Guidance

LIU Zhigang^{1,2*}, ZHANG Guohui¹, GAO Yue¹, and LIU Miaomiao¹

(1. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China;

2. Heilongjiang Petroleum Big Data and Intelligent Analysis Key Laboratory, Daqing 163318, China)

Abstract To solve the difficulty in obtaining annotated pedestrian images in the field of pedestrian re-identification research, a novel data augmentation method guided by multi-factor is proposed in this paper. Firstly, a local multi-scale guidance mechanism is designed in the generator network. It can suppress the local artifacts in generated images through feature fusion. Secondly, a long-distance correlation guidance mechanism is proposed to improve the overall visual quality of the generated pedestrian image by guiding the long-distance dependence of the generated image with external attention. Lastly, an adversarial discrimination network is designed and embed into original generative adversarial networks. The three network stability architecture model increases the stability of generative adversarial network training. The experiment are validated on the VIPeR, Market-1501 and DukeMTMC-reID benchmark datasets. The results demonstrate our method outperforms the state-of-the-art with the mAP and rank-1 scores, especially in small-scale datasets.

Key words person re-identification; generative adversarial network; data augmentation; local multi-scale; attention mechanism

行人重识别 (Person Re-Identification, Re-ID) 是在多个摄像头中进行特定目标行人图像的检索问题。近年来, 得益于大规模标注数据集和卷积神经网络的拟合能力, Re-ID 取得了较大进展。然而, 由于行人数据涉及隐私、标注代价等因素, 高质量的行人标注图像数据的获取极为困难, 目前有监督学习的 Re-ID 性能很大程度上受制于此。

随着生成对抗网络 (Generative Adversarial Networks, GAN) [1] 快速发展, 基于 GAN 的行人数据增广 [2-7] 方法引起了重视。文献 [8] 首次提出将 GAN 用于 Re-ID 的研究, 通过对 DCGAN [9] 生成行人图像, 实现对标注数据集的扩充。文献 [10-14] 使用改进的 CycleGAN [15] 进行域之间的行人风格迁移。文献 [16] 提出 AD-Cluster 模型, 根据源域与

收稿日期: 2023-02-28; 修回日期: 2023-07-02

基金项目: 国家自然科学基金 (51774090, 42002138); 黑龙江省自然科学基金 (LH2020F003); 河北省自然科学基金 (D2023107002); 黑龙江省属本科高校团队创新基金 (2022TSTD-03); 黑龙江省高等教育教学改革项目 (SJGY20210109)

作者简介: 刘志刚, 博士, 教授, 主要从事机器学习、计算机视觉方面的研究。

*通信作者 E-mail: dqpilzg@163.com

目标域图像的语义一致性, 增强跨域 Re-ID 模型的特征表达能力。文献 [17] 利用 FFGAN 实现基于增广判别聚类的数据迁移。文献 [18] 使用 UnityGAN 学习不同摄像机之间的背景风格差异, 生成基于这些差异的平均风格图像, 提升 Re-ID 模型的泛化能力。此外, 基于 GAN 的行人数据增广方法可解决行人姿态与外观发生变化而引起的精度降低问题。文献 [19] 提出一种基于姿态引导的生成对抗网络 (Pose Guided Person Generation Network, PG2), 根据给定的行人图像和目标姿态合成任意姿态。文献 [20] 将变分推理和 GAN 相结合, 提出一种生成人物衣服模型 VariGAN。文献 [21] 提出 Cloned Person 方法, 将真实世界的人物图像中的服装克隆到虚拟的三维人物中, 进而实现数据增广。

这些基于 GAN 的数据增广方法, 在一定程度上提升了 Re-ID 模型的性能、降低了手工标注样本的代价。但是, 它们更多关注图像风格变换, 而忽略生成图像质量给 Re-ID 模型带来的噪声影响, 生成的行人图像有时会面临局部细粒度特征表达较弱、图像整体视感质量降低的问题。虽然此类图像可以在一定程度上提升模型的鲁棒性, 但过多低质量图像会在 Re-ID 模型提取到的特征中融入过多噪声, 干扰模型训练的稳定性。

本文经过研究分析, 传统 GAN 生成图像质量较差的原因主要包括: 1) 受到卷积核尺寸的影响, 生成行人图像的局部信息表达、长距离相关性

依赖均受到较大限制, 进而导致生成图像局部伪影严重、图像整体视感质量欠佳; 2) 网络训练方法难以保证稳定性, 直接影响生成行人的图像质量。在此基础上, 本文提出一种基于多因素引导行人图像增广方法 (Multi-factor Guidance Data Augmentation Method, MG-DAM)。首先, 设计了一种多尺度引导机制 (Local Multi-scale Guidance, LMG), 通过在生成器网络中使用局部多尺度引导机制, 对行人图像进行特征提取与特征融合, 从而抑制生成图像的局部伪影, 增强生成图像细粒度特征表达能力; 其次, 提出一种长距离相关性引导机制 (Long-distance Correlation Guidance, LCG), 突破卷积核映射关系的限制, 增加生成行人图像的长距离依赖; 为提升网络训练的稳定性, 设计对抗博弈判别网络 (Adversarial Discrimination Network, AD), 进而更改网络整体训练方式。最后, 通过仿真实验证明本文所提算法的有效性。

1 多因素引导方法 MG-DAM

MG-DAM 包括 3 部分: 1) 局部多尺度引导机制 LMG: 侧重解决生成图像局部信息表征能力不足而导致的局部伪影严重问题; 2) 长距离相关性引导机制 LCG: 侧重于解决生成图像整体视感质量较差的问题; 3) 对抗博弈判别网络 AD: 侧重于解决生成图像质量不稳定问题。图 1 为本文所提 MG-DAM 的结构图。

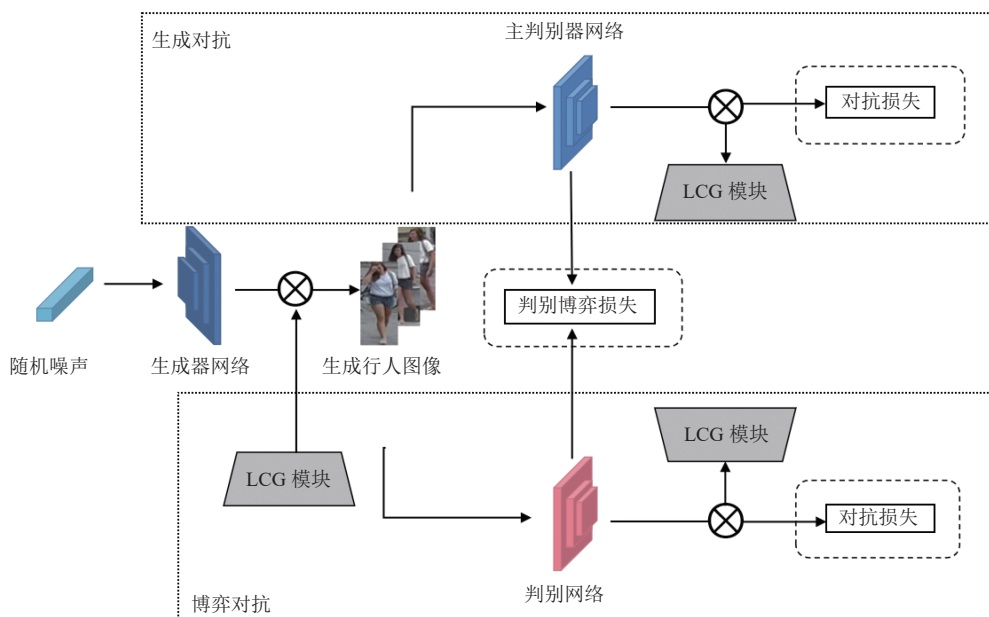


图 1 MG-DAM 模型结构图

1.1 局部多尺度引导机制 LMG

有效解决生成行人图像局部伪影的关键是增强生成图像局部特征的多层次表达能力。文献 [22] 提出一种单阶段可控的 SSCGAN, 在半监督中进行条件细粒度图像合成。文献 [23] 提出风格检索 RIS, 通过 StyleGAN 潜在空间的解耦特性, 完成局部特征迁移。文献 [24] 使用图像以及图像边缘一致性, 设计一种并行网络提取边缘细节特征。

与这些方法不同, 本文直接通过融合多尺度特征引导图像的局部信息表征, 进而建立局部多尺度引导机制 LMG。所谓多尺度, 就是对图像不同粒度的采样, 一般来说, 粒度更小、更密集的采样可

得到更多的细节; 粒度更大、更稀疏的采样可得到整体的趋势。如图 2 第一部分, LMG 是对相同卷积层使用不同尺寸卷积核提取到的特征进行融合。在训练过程中, 根据反馈的多尺度融合信息, 生成器网络调整生成图像的局部信息表达, 进而抑制局部伪影的出现。因为卷积神经网络通过逐层抽象的方式来提取目标的特征, 因此卷积核尺寸的设计尤为重要。如果卷积核过小, 则只能抽取到局部的特征, 反之则获得更多无效的信息。同时, 由于行人本身具有非刚性的特点, 因此在生成行人图像时, 不仅需要小卷积核获取局部细粒度特征信息, 还需要大卷积核获取局部区域信息。

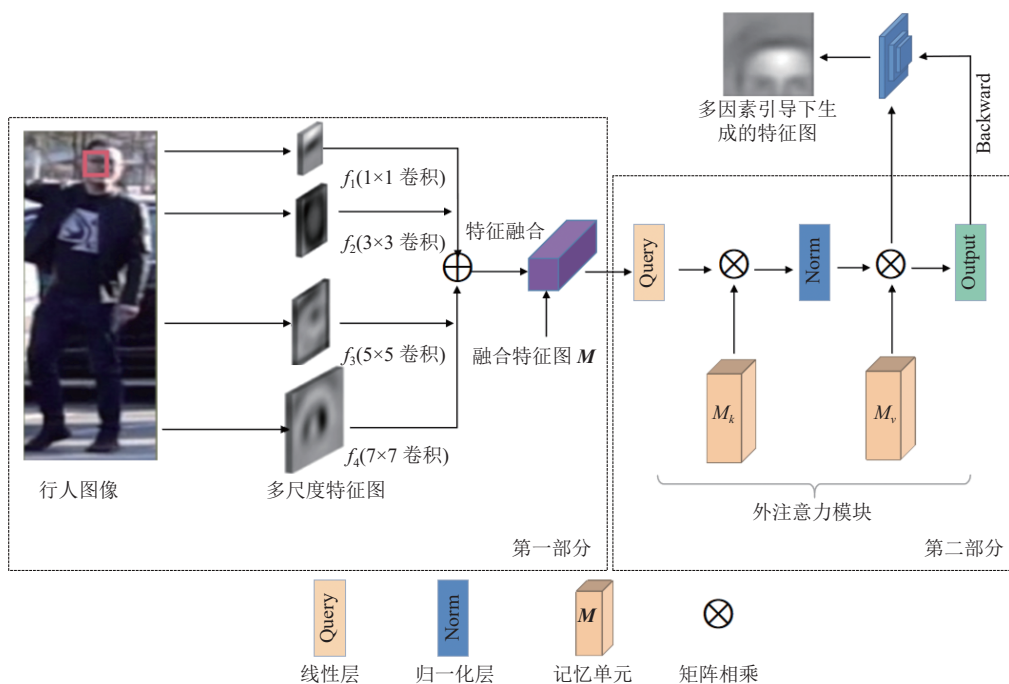


图 2 LMG 与 LCG 引导机制

1.2 长距离相关性引导机制 LCG

在图像生成领域, 生成图像细节与全局平衡依然是一个重要问题。文献 [25] 提出分层自适应对角空间注意方法 DAT, 以分层方式分别处理样式中的空间内容; 文献 [26] 提出了一个可微的全局流局部注意框架 GLFA, 用于在特征级的重组输入。这些方法均使用到了注意力机制, 表明注意力可解决因卷积核尺寸受限而导致的难以获得图像长距离依赖关系的问题。考虑到计算复杂度以及样本间的联系, 本文使用外注意力机制引导图像的长距离相关性生成, 该方法相较基于 Transformer 的注意力机制, 外注意力更加的轻便, 训练周期缩短。如图 2 第二部分, 外注意力模块由两个线性层、两

个归一化层和两个记忆单元组成, 具体为:

$$A = (\alpha)_{i,j} = \text{Norm}(FM^T) \quad (1)$$

$$F_{\text{out}} = AM \quad (2)$$

式中, $F \in \mathbf{R}^{N \times d}$ (N 为像素个数, d 为特征维度) 表示输入特征图; $(\alpha)_{i,j}$ 表示第 i 个元素和 M 的第 j 行之间的相似性; 记忆单元 $M \in \mathbf{R}^{S \times d}$ 为整个训练数据集的记忆。

1.3 对抗博弈判别网络 AD

由于 GAN 模型存在收敛不稳定性, 往往需要通过设置较多的参数去平衡判别器网络与生成器网络, 进而达到纳什均衡状态。不收敛问题主要有两种原因。1) 梯度消失。原始 GAN 使用分类误差

作为真实分布与生成分布的相似度量。然而,当真实分布与生成分布的重叠区域可忽略时,生成器获得的梯度接近零,此时生成器网络得不到来自判别器网络的反馈,无法完成自我优化。2) 模式崩塌。当使用梯度下降算法对 GAN 优化时,实际上不区分 min-max 和 max-min,这导致生成器会过多地生成一些重复样本。

为此,本文通过在原始 GAN 模型的基础上添加对抗博弈判别网络 AD 的方式,构建一种三网络稳定博弈架构模型。与常见的三网络模型不同,AD 不仅与生成器网络进行判别对抗,还会与主判别器网络进行博弈对抗,以校正最终的整体损失函数。AD 详细介绍如下。

令 $\{x_i\}_{i=0}^n \subseteq X$ 表示从未知的行人数据分布 P_{data} 中提取的训练数据集, $\{z_i\}_{i=0}^n \subseteq Z$ 表示从先验分布噪声 P_z 中采样得到的样本。在传统 GAN 网络模型中,将随机采样的噪声输入到生成器网络中,通过判别器网络与生成器网络进行最大最小博弈产生新的行人图像 $\{\tilde{x}\}_{i=0}^n = G(z)$, 其服从新的生成分布 P_G , 博弈过程如下:

$$\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}} [\log D(x)] + E_{z \sim P_z} [\log(1 - D(G(z)))] \quad (3)$$

针对传统 GAN 网络训练困难的问题,本文在上述 GAN 模型结构的基础上融入了新的对抗再判别网络,同时使用了新的数据分布形式:

$$P_{\text{Game}} (\forall x \subseteq P_{\text{data}}, P_G : P_{\text{Game}} = [P_{\text{data}} + P_G] / 2)$$

新的判别网络与生成网络间的最大最小博弈过程为:

$$\min_G \max_{D_1, D_2} V(D_1, D_2, G) = E_{x \sim P_{\text{data}}} [\log D_1(x)] + E_{z \sim P_z} [\log(1 - D_1(G(z)))] + E_{x \sim P_{\text{data}}} [\log D_2(x)] + E_{z \sim P_z} [\log(1 - D_2(G(z)))] + \lambda D_{s_Game} \quad (4)$$

其中,判别器网络之间的相互博弈过程如下:

$$\begin{aligned} D_{s_Game} = & E_{x \sim P_{\text{game}}} [f(D_1(x), l(D_2(x) > \frac{1}{2})) - \\ & \mu f(D_1(x_{p_g1}), l(D_2(x_{p_g2}) > \frac{1}{2}))] + \\ & E_{x \sim P_{\text{game}}} [f(D_2(x), l(D_1(x) > \frac{1}{2})) - \\ & \mu f(D_2(x_{p_g1}), l(D_1(x_{p_g2}) > \frac{1}{2}))] \end{aligned} \quad (5)$$

式中, x, x_{p_g1}, x_{p_g2} 相互独立,且 x_{p_g1}, x_{p_g2} 随机取样于 P_{Game} ; $l(\cdot)$ 为指示函数; $\mu, \lambda \in [0, 1]$ 为调节权重的超参; f 为评估函数,具体如下:

$$f(D_i(x), y) = \begin{cases} \log(D_i(x)) \rightarrow \text{if}(y = 1) \\ \log(1 - D_i(x)) \rightarrow \text{if}(y = 0) \end{cases} \quad (6)$$

式(4)中, D_1 与 D_2 拥有相同的网络结构与优化方式,并与生成器网络 G 进行对抗训练。式(5)表示判别器网络之间的对抗训练过程,针对梯度弥散问题,式(5)的(a)与(c)约束 D_1 与 D_2 具有相同的判别结果,最终使两个判别器达到彼此收敛;针对模式崩塌问题,本文在目标函数中引入(b)与(d)来惩罚 D_1 与 D_2 的判别结果。在对(b)与(d)的采样过程中, x_{p_g1}, x_{p_g2} 是相互独立样本,由于独立性的作用,使得判别器网络之间的判别结果不会过度的一致,从而避免两个判别器网络的判别结果过度一致而导致梯度消失。

2 实验结果与分析

为验证本文方法的有效性,仿真实验数据集为 VIPeR^[27]、Market-1501^[28] 以及 DukeMTMC-reID。这 3 个数据集规模与风格差异较大,可较好地验证本方法的有效性。实验环境: Intel Xeon (R) E5-2640, 32 GB, GTX2070super。模型性能评估采用标准的评价指标 Rank-1 和平均准确率 (mean average precision, mAP), mAP 反映模型的总体性能。

2.1 MG-DAM 与其他主流方法的比较

将 MG-DAM 与多种主流 GAN 模型在 3 个不同规模的数据集上进行数据增广,并使用相同的 Re-ID 方法,结果如表 1。

1) 在使用本文所提方法增广过的数据集上进行 Re-ID,实验效果一般优于所比较的方法。其中,在小规模数据集 VIPeR 上,本文方法明显优于其他方法,Rank-1 为 67.8%,mAP 为 51.6%。

2) 在 Market-1501 数据集上,本文方法略低于 DGNet 方法 0.3%,DGNet 是关注行人换装的网络,一定程度上提升了 Re-ID 抗干扰能力,本文方法更加关注生成行人图像的质量;在 DukeMTMC-reID 数据集上,本文方法识别精度超过 DGNet,原因是 DGNet 的跨域能力与实际换装能力有限,降低了生成图像的质量和模型的识别精度。

表 1 与其他主流方法对比实验数据表

| 方法 | VIPeR | | Market-1501 | | DukeMTMC-reID | | % |
|------------------------------|-------------|-------------|-------------|-------------|---------------|-------------|---|
| | Rank-1 | mAP | Rank-1 | mAP | Rank-1 | mAP | |
| DeformGAN ^[29] | — | — | 80.6 | 61.3 | — | — | |
| PTGAN ^[30] | 62.1 | 49.5 | 87.7 | 75.9 | 71.6 | 46.6 | |
| AD-Cluster ^[16] | — | — | 90.3 | 80.9 | 75.6 | 48.9 | |
| FFGAN ^[17] | — | — | 89.4 | 77.6 | 76.2 | 52.0 | |
| UnityGAN ^[18] | 65.3 | 48.6 | 91.3 | 78.3 | 74.3 | 49.7 | |
| PG2 ^[19] | 64.3 | 45.2 | 89.4 | 75.7 | 72.2 | 50.6 | |
| VariGAN ^[20] | 66.2 | 47.7 | 89.3 | 78.8 | 72.2 | 51.3 | |
| DGNet ^[31] | — | — | 91.7 | 84.0 | 77.2 | 52.3 | |
| ClonedPerson ^[32] | — | — | 84.5 | 59.9 | — | — | |
| 本文 | 67.8 | 51.6 | 91.4 | 83.2 | 78.6 | 52.7 | |

2.2 LMG 有效性验证

为验证局部多尺度引导机制 LMG 对提高生成图像的局部信息表达能力的有效性, 设计以下 3 组实验: 第一组使用 ResNet 作为判别器网络的主干网络结构; 第二组实验在第一组实验的基础上增加 LMG, 不同组实验所使用的 Re-ID 方法完全一致。从表 2 可以看出: 在第一组实验的基础上, 加

入了 LMG 的第二组实验在各个数据集上的表现都超越了原始模型, 在小规模数据集 VIPeR 上尤为明显, 最高的 Rank-1 达到了 53.8%, mAP 达到了 36.2%。因此, 局部多尺度引导机制 LMG 可以增强生成图像的局部特征表达能力, 抑制生成图像出现局部伪影, 同时有效提升了 Re-ID 方法的识别精度。

表 2 局部多尺度特征引导机制 LMG 验证数据表

| 实验组号 | 方法 | VIPeR | | Market-1501 | | DukeMTMC-reID | | % |
|------|--------------------------|-------------|-------------|-------------|------|---------------|------|---|
| | | Rank-1 | mAP | Rank-1 | mAP | Rank-1 | mAP | |
| 第一组 | GAN ^[1] | 41.6 | — | 57.6 | 49.5 | 53.6 | 34.2 | |
| | WGAN ^[30] | 45.5 | 30.3 | 68.5 | 49.9 | 61.3 | 37.2 | |
| | CycleGAN ^[15] | 48.3 | 31.2 | 70.6 | 52.1 | 62.1 | 38.1 | |
| | DualGAN ^[33] | 44.6 | 30.0 | 72.3 | 51.2 | 60.0 | 37.1 | |
| 第二组 | GAN + LMG | 47.6 | — | 57.6 | 57.6 | 61.5 | 36.1 | |
| | WGAN + LMG | 53.1 | 35.4 | 72.7 | 58.6 | 64.3 | 40.3 | |
| | CycleGAN + LMG | 53.8 | 36.2 | 74.5 | 59.0 | 65.8 | 40.6 | |
| | DualGAN + LMG | 50.2 | 34.8 | 71.3 | 58.6 | 62.7 | 40.0 | |

2.3 LCG 的有效性验证

为验证全局相关性依赖机制 LCG 对提高生成行人图像整体视感质量的有效性, 设计以下 3 组实验: 第一组实验选用自注意力 (Self-Attention, SA) 作为长距离依赖机制与相应的 GAN 模型结合; 第二组实验选用其他较先进的基于长距离依赖的 GAN 方法进行数据增广; 第三组实验在第一组实验的基础上, 选用 LCG 替代自注意力。实验结果如表 3。对比第一、三组实验, 由于 LCG 可以考虑到样本间的差异, 其在各个数据集上的表现均优于自注意

力机制, 最高的 Rank-1 达到了 83.9%、mAP 达到了 73.9%; 相较于第二组实验中目前较为先进的方法, 虽然本文所提方法并未达到与其相同的实验效果, 但是 Rank-1 与 mAP 之差基本浮动在 3% 左右, 最小的 Rank-1 精度仅差 2.7%, mAP 精度仅差 1.7%。

结合本模块的实验效果图 (图 3②) 可以看到, 生成行人图像的整体视感质量正在逐步提升。因此, 不难看出, LCG 可以解决因为卷积核尺寸受限导致的生成图像整体质感质量的问题, 切实提升 ReID 方法的识别精度。

表 3 长距离相关性引导机制 LCG 验证数据表

| 实验组号 | 方法 | VIPeR | | Market-1501 | | DukeMTMC-Re-ID | |
|------|--------------------------------|-------------|-------------|-------------|-------------|----------------|-------------|
| | | Rank-1 | mAP | Rank-1 | mAP | Rank-1 | mAP |
| 第一组 | GAN ^[1] + SA | 44.6 | — | — | — | 59.3 | 36.3 |
| | WGAN ^[30] + SA | 49.6 | 31.6 | 78.6 | 63.7 | 63.1 | 39.0 |
| | CycleGAN ^[15] + SA | 50.1 | 32.3 | 79.3 | 68.5 | 63.5 | 40.0 |
| | DualGAN ^[33] + SA | 48.7 | 30.2 | 79.6 | 69.3 | 61.0 | 39.6 |
| 第二组 | DAT ^[25] | — | — | 85.7 | 74.6 | 73.2 | 73.6 |
| | GLFA ^[26] | — | — | 86.6 | 75.8 | 74.2 | 74.6 |
| 第三组 | GAN + LCG | 50.3 | — | 72.6 | 59.4 | 62.7 | 38.3 |
| | WGAN + LCG | 56.2 | 38.2 | 80.6 | 70.8 | 66.3 | 41.7 |
| | CycleGAN ^[15] + LCG | 56.6 | 39.5 | 83.9 | 73.9 | 67.3 | 43.8 |
| | DualGAN + LCG | 53.7 | 37.0 | 82.4 | 73.3 | 65.8 | 42.5 |



图 3 模型实验效果对比图

2.4 AD 的有效性验证

为验证对抗博弈判别网络 AD 能够提升网络训练的稳定性, 设计了基于自拟数据集进行聚类的对比仿真实验, 验证各类模型生成样本种类的多样性。模型训练稳定性的验证结果如图 4 所示, 生成样本种类已经涵盖所有类别, 而其他方法达到同样效果至少需要 20 000 个 epoch。且在训练过程中并无出现因梯度弥散而导致聚类失败的状况。因此, AD 可以更加快速的推进网络生产多样性样本, 增加了 Re-ID 模型的鲁棒性。虽然该方法在 GAN 的基础上添加了新的判别网络模型, 但仅仅引入了很少的计算量。相对于当前的 GPU 计算效率, 引入计算量可忽略不计。同时, 相较于较为主流的 GAN 模型, 本文使用的判别网络模型均为无标签

学习, 这进一步提升了网络的计算效率。

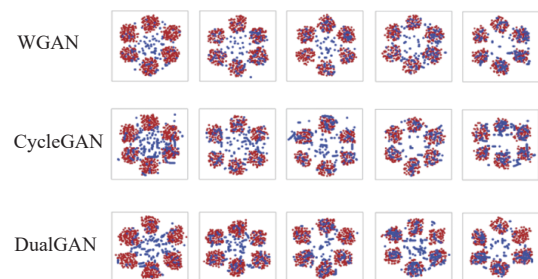


图 4 模型多样性样本匹配图

3 结束语

本文提出了一种基于多因素引导的行人数据增广方法, 通过局部多尺度引导机制 LMG 增强生成行人图像的局部特征表达能力, 抑制生成图像局部

伪影的产生; 通过长距离相关性引导机制 LCG 突破了卷积核映射关系的限制, 提升生成行人图像的整体视感质量; 通过对抗博弈判别网络 AD 的使用, 改变了 GAN 模型的对抗机制, 提升了网络训练的稳定性。在数据集 VIPeR、Market-1501 和 DukeMTMC-reID 上的实验结果表明, 本文所提方法可以提升 Re-ID 模型的性能, 并且在一定程度上优于多数较为流行的 GAN 模型, 为有效提升高质量的行人图像生成与数据增广提供了研究思路。

参考文献

- [1] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. *Advances in Neural Information Processing Systems*, 2014(2): 2672-2680.
- [2] LI Y J, CHEN Y C, LIN Y Y, et al. , Recover and identify: A generative dual model for cross-resolution person re-identification[C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Seoul: IEEE, 2019: 8089-8098.
- [3] CHEN Y B, ZHU X T, GONG S G. Instance-guided context rendering for cross-domain person re-identification[C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Seoul: IEEE, 2019: 232-242.
- [4] LIANG W Q, WANG G C, LAI J H. Asymmetric cross-domain transfer learning of person re-identification based on the many-to-many generative adversarial network[J]. *Acta Automatica Sinica*, 2022, 48(1): 103-120.
- [5] WANG Z X, WANG Z, ZHENG Y Q, et al. Learning to reduce dual-level discrepancy for infrared-visible person re-identification[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2019: 618-626.
- [6] ZHENG Z D, YANG X D, YU Z D, et al. Joint discriminative and generative learning for person re-identification[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2019: 2133-2142.
- [7] ZHA Z J, LIU J W, CHEN D, et al. Adversarial attribute-text embedding for person search with natural language query[J]. *IEEE Transactions on Multimedia*, 2020, 22(7): 1836-1846.
- [8] ZHENG Z D, 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 (ICCV)*. Venice: IEEE, 2017: 3774-3782.
- [9] RADFORD A, METZ L, CHINTALA S, et al. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. [2020-12-22]. <https://arxiv.org/pdf/1511.06434.pdf>.
- [10] WEI L H, ZHANG S L, 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 (CVPR)*. Salt Lake City: IEEE, 2018: 79-88.
- [11] 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]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City: IEEE, 2018: 994-1003.
- [12] HUANG Y, XU J, WU Q, et al. Multi-pseudo regularized label for generated data in person re-identification[J]. *IEEE Transactions on Image Processing*, 2018, 28(3): 1391-1403.
- [13] QIAN X L, FU Y W, XIANG T, et al. Pose-normalized image generation for person re-identification[EB/OL]. [2021-06-22]. <https://arxiv.org/pdf/1712.02225.pdf>.
- [14] ZHONG Z, ZHENG L, ZHENG Z D, et al. Camera style adaptation for person re-identification[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City: IEEE, 2018: 5157-5166.
- [15] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, 2017: 2242-2251.
- [16] ZHAI Y P, LU S J, YE Q X, 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 (CVPR)*. Seattle: IEEE, 2020: 9018-9027.
- [17] ZHANG X W, LYU M Q, LI H. Cross-domain person re-identification based on partial semantic feature invariance[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(9): 1682-1690.
- [18] LIU C, CHANG X J, SHEN Y D. Unity style transfer for person re-identification[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020: 6886-6895.
- [19] MA L, JIA X, SUN Q, et al. Pose guided person image generation[C]//*Proceedings of the Conference on Neural Information Processing Systems (NIPS)*. Long Beach: NIPS, 2017: 405-415.
- [20] ZHAO B, WU X, CHENG Z Q, et al. Multi-view image generation from a single-view[EB/OL]. [2022-07-15]. <https://arxiv.org/pdf/1704.04886.pdf>.
- [21] WANG Y, LIANG X, LIAO S. Cloning outfits from real-world images to 3D characters for generalizable person re-identification[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 4900-4909.
- [22] CHEN T, LIU Y, ZHANG Y, et al. Semi-supervised single-stage controllable GANs for conditional fine-grained image generation[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Montreal: IEEE, 2021: 9244-9253.
- [23] CHONG M J, CHU W S, KUMAR A. Retrieve in style:

- Unsupervised facial feature transfer and retrieval [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Montreal: IEEE, 2021: 3867-3876.
- [24] 卢涛, 陈冲, 许若波, 等. 基于边缘增强生成对抗网络的人脸超分辨率重建[J]. 华中科技大学学报:自然科学版, 2020, 48(1): 87-92.
- LU T, CHEN C, XU R B, et al. Face hallucination based on edge enhanced generative adversarial network[J]. Journal of Huazhong University of Science and Technology, 2020, 48(1): 87-92.
- [25] KWON G, YE J C. Diagonal attention and style-based gan for content-style disentanglement in image generation and translation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Montreal: IEEE, 2021: 13960-13969.
- [26] REN Y, YU X, CHEN J, et al. Deep image spatial transformation for person image generation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 7687-7696.
- [27] GRAY D, HAI T. Viewpoint invariant pedestrian recognition with an ensemble of localized features [C]//Proceedings of European Conference on Computer Vision (ECCV). [S.l.]: IEEE, 2008: 3408-3416.
- [28] ZHENG L, SHEN L, LU T, et al. Scalable person re-identification: A benchmark[C]//Proceedings of the IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1116-1124.
- [29] ALIAKSANDR S, ENVER S, STEPHANE L, et al. Deformable GANs for pose-based human image generation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE, 2018: 3408-3416.
- [30] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[EB/OL]. [2022-09-15]. <https://arxiv.org/pdf/1704.00028.pdf>.
- [31] ZHENG Z, YANG X, YU Z, et al. Joint discriminative and generative learning for person re-identification [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2020: 2133-2142.
- [32] WEI L H, ZHANG S L, GAO W, et al. Person transfer GAN to bridge domain gap for person re-identification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 79-88.
- [33] YI Z, ZHANG H, TAN P, et al. DualGAN: Unsupervised dual learning for image-to-image translation[J]. IEEE Computer Society, 2017: 2849-2857.

编辑 税 红