

文章编号:1007-2780(2022)10-1364-11

改进生成对抗网络的换衣行人再识别

张玉霞^{1,2}, 车进^{1,2*}, 贺愉婷^{1,2}

(1. 宁夏大学 物理与电子电气工程学院, 宁夏 银川 750021;

2. 宁夏沙漠信息智能感知重点实验室, 宁夏 银川 750021)

摘要:针对不同摄像机之间存在显著的类内差异,部分行人如换衣等发生变化均会导致识别下降的问题,提出一种将生成对抗网络(GAN)数据生成和行人再识别学习端对端结合起来的联合学习框架。首先,在 DGNet 网络的基础上,将生成器与鉴别器中的残差块 ResBlock 替换为 DenseBlock,加强特征传播并且避免梯度消失的问题。然后,生成模块通过切换外观和结构编码,生成高质量的合成图像。最后,在鉴别模块中加入新型注意力机制(NAM),抑制不太显著的权重,更关注所需的目标区域。生成模块生成的图像在线反馈给鉴别模块鉴别图像真假,同时将结果反馈给行人再识别模块进行分类识别(鉴别模块与行人再识别模块是共享的)。在 Market-1501、DukeMTMC-reID 数据集上的 rank-1/mAP 分别达到 95.7%/88.6% 和 87.1%/75.7%。

关键词:行人再识别;外观编码;结构编码;GAN;DenseBlock;NAM

中图分类号:TP391 **文献标识码:**A **doi:**10.37188/CJLCD.2022-0086

Improved generative adversarial network for pedestrian re-identification in clothes

ZHANG Yu-xia^{1,2}, CHE Jin^{1,2*}, HE Yu-ting^{1,2}

(1. School of Physics and Electronic-Electrical Engineering, Ningxia University, Yinchuan 750021, China;

2. Ningxia Key Laboratory of Intelligent Sensing for Desert Information, Yinchuan 750021, China)

Abstract: In view of the significant intra-class differences between different cameras, and the change of the clothes of some pedestrians, which leads to the degradation of recognition, an end-to-end combination of generative adversarial network (GAN) data generation and person re-identification learning is proposed, which makes up the federated learning framework. Firstly, on the basis of the DGNet network, the residual block (ResBlock) in the generator and discriminator is replaced with DenseBlock, which strengthens feature propagation and avoids the problem of gradient disappearance. Then, the generative module generates high-quality synthetic images by switching appearance and structure encodings. Finally, a new normalization-based attention module (NAM) is added to the discriminative module to suppress less salient weights and pay more attention to the desired target regions. The images generated by the generation module are fed back online to the identification module to identify the true and fake images, and the results are fed back to the

收稿日期:2022-03-15;修订日期:2022-03-25.

基金项目:国家自然科学基金(No.61861037)

Supported by National Natural Science Foundation of China(No.61861037)

*通信联系人, E-mail: koalache@126.com

person re-identification module for classification and identification (the identification module and the pedestrian re-identification module are shared). The rank-1/mAP on Market-1501 and DukeMTMC-reID datasets reach 95.7%/88.6% and 87.1%/75.7%, respectively.

Key words: person re-identification; appearance coding; structure coding; GAN; DenseBlock; NAM

1 引言

行人再识别^[1](Person re-identification, ReID)是指给定一个目标监控行人图像,检索跨摄像头、跨时段拍摄的该行人图像。行人再识别通常被视为一个度量学习问题^[2]。由于在安全和监控方面有重要的应用,因此行人再识别受到学术界和工业界的广泛关注。随着深度学习方法的广泛应用,不同算法的ReID性能迅速提高,但由于拍摄的行人图像会受到背景、光照、人体姿态变化、遮挡等引起的显著类内差异的影响,并且行人再识别研究^[3-6]都假设同一个人穿着相同的衣服,如文献[7]中提到的一名女嫌犯为躲避追捕,故意将黑色外套换成白色外套,行人再识别算法很容易将穿着不同、ID相同的行人识别错误。因此,行人再识别仍是一项艰巨的任务。

为进一步减少类内变化的影响,文献[8]寻找最优的仿射变换用来更好地区分不同的身份ID,提出PAN网络,在不用额外的标注情况下同时进行行人对齐和学习判别性的特征描述。文献[9]提出一个基于身体部位对齐的双流网络模型,可以将人体姿势表示为部分特征图,并将它们直接与外观整体特征图结合以计算对齐的行人部位表示。张良等人^[10]提出将局部特征、全局特征和多粒度特征融合,提高行人再识别精度。文献[11]首次将生成对抗网络(Generative Adversarial Network, GAN)引入到行人再识别中,利用生成的数据扩充数据集来辅助训练,提升行人再识别精度。PTGAN^[12]的提出主要是用来弥合不同数据集之间的领域差距。文献[13]提出一种利用姿态以及外观特征混合编码的行人再识别算法,以解决摄像机视角造成的类内差异。PNGAN^[14]将每个行人图像生成8个标准姿态的假图,最后将生成图和原数据集图像一起输入识别网络,在一定程度上消除姿态变化的影响。文献[15]采用一个具有双分支结构的孪生神经网络,只学习与行人身份相关的特征。杨海伦等人^[16]因行人存

在遮挡问题提出基于可变形卷积的CNN模型,并用CycleGAN解决相机风格问题。

以上方法主要是将行人进行对齐或者用GAN生成不同相机风格、不同姿态数据来扩充训练集,再进行分类识别。不同于以上方法,本文在DGNet网络^[17]的基础上提出改进生成对抗网络的换衣行人再识别算法,不仅将GAN数据生成和行人再识别学习端对端结合起来,而且不需要任何辅助信息就能够生成高质量图像,对ReID进行图像增强。生成模块中的编码器将每个行人分解为外观编码(衣服+鞋子+手机+包包+身份信息等)与结构编码(姿态+头发+脸型+背景等)。通过切换外观编码和结构编码,生成模块能够生成高质量的合成图像,然后在线反馈给鉴别器进行辨别真假,并用于行人再识别模块提高行人识别精度。其中,鉴别模块与行人再识别学习是共享的。通过在鉴别模块中加入通道间的注意力机制模块(Normalization-based Attention Module, NAM)^[18],突出网络显著特征,更好地辨别图像真假,提高行人分类识别的精度。用DenseBlock^[19]替换DG-Net网络中生成器与鉴别器的ResBlock,加强特征传播,避免梯度消失问题,再一次提高生成图质量与行人再识别精度。通过实验可知,本文算法在Market-1501、DukeMTMC-reID两大公开数据集上的rank-1/mAP分别达到95.7%/88.6%和87.1%/75.7%,证明本文算法并具有一定的优势。

2 网络架构

本文网络架构如图1所示。图像生成模块与行人再识别模块端对端结合起来,外观编码器和鉴别器是共用的,这也是行人再识别学习的主干。生成器通过自生成网络和互生成网络两种映射合成假图像,将数据集图像与生成图像一起输入到鉴别器中判别真假,并且对其进行分类识别,提高行人再识别的识别精度。鉴别模块包含

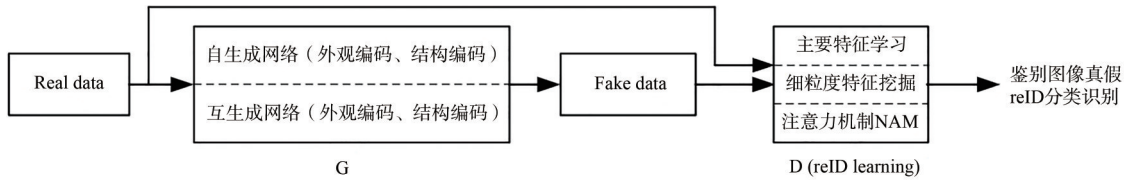


图1 网络架构

Fig. 1 Network structure

主要特征学习、细粒度特征挖掘以及注意力机制NAM。通过不同角度进行学习,网络不仅能够得到行人衣服等外观的信息,而且学习到外貌等更细致的结构特征,从而更好地鉴别图像真假,提高识别精度。

2.1 生成网络

生成网络框架图如图2所示。本文数据集中的真实图像与身份标签用 $X = \{x_i\}_{i=1}^N$ 和 $Y = \{y_i\}_{i=1}^N$ 表示,其中 N 表示图像的数量, $y_i \in [1, K]$, K 表示数据集中类别的数量。给定训练集中两幅真实图像,通过生成模块交换两幅图像的外观与结构代码生成几幅新图像。生成网络包括自生成网络和互生成网络,生成模块由一个外观编码器 $E_a: x_i \rightarrow a_i$ 、一个结构编码器 $E_s: x_j \rightarrow s_j$ 、一个生成解码器 $G: (a_i, s_j) \rightarrow x_j^i$ 以及一个区分生成图像与原始图像的鉴别器 D 组成。

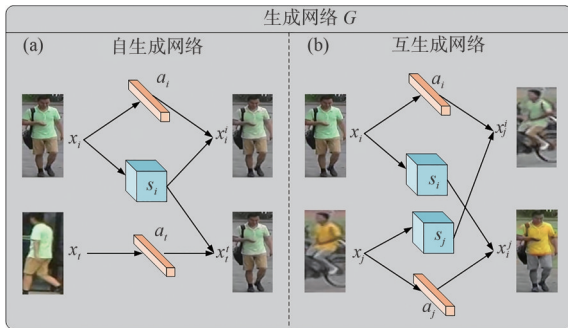


图2 生成网络框架图。(a)自生成网络;(b)互生成网络。

Fig. 2 Generate network frame diagram. (a) Self generation network; (b) Mutual generation network.

2.1.1 自生成网络

自生成网络框架如图2(a)所示。给定一对行人图像 $x_i, x_j \in X$ (身份相同),首先生成模块学习如何重建图像,自生成网络首先提取图像 x_i 的外观特征 a_i 以及其结构特征 s_i 生成与原始图像外观、结构均一致的假图像 x_i^i ,使用像素级损失

函数 $L_{\text{recon1}}^{\text{img1}}$:

$$L_{\text{recon1}}^{\text{img1}} = \mathbb{E} \left[\|x_i - G(a_i, s_i)\|_1 \right], \quad (1)$$

自生成网络同时会生成同一身份、不同结构的图像,因此采用图像 x_i 的结构特征 s_i 与 x_i 的外观特征 a_i 生成图像 x_i^i ,同样使用像素级损失函数 $L_{\text{recon2}}^{\text{img2}}$:

$$L_{\text{recon2}}^{\text{img2}} = \mathbb{E} \left[\|x_i - G(a_i, s_i)\|_1 \right], \quad (2)$$

这种跨图像(相同身份但不同结构)重建用外观编码器将相同身份的外观联系在一起,减少类内变化。对于不同身份的图像来说,使用身份损失来区分不同身份。

$$L_{\text{id}}^s = \mathbb{E} \left[-\log(p(y_i | x_i)) \right], \quad (3)$$

其中, $p(y_i | x_i)$ 是基于外观编码预测 x_i 属于实际类别 y_i (ID) 的概率。

2.1.2 互生成网络

互生成网络框架如图2(b)所示。互生成网络的输入图像为 $x_i, x_j \in X$ (身份不同,即 $y_i \neq y_j$),主要是生成不同身份的图像。该网络提取图像 x_i 的外观特征 a_i 、结构特征 s_i 和 x_j 的外观特征 a_j 和结构特征 s_j ,通过交叉组合成新图像 x_j^i 和 x_i^j 。其中 x_j^i 生成图像与 x_i 外观一致、 x_j 结构一致; x_i^j 生成图像与 x_j 外观一致、 x_i 结构一致。这里,使用外观和结构编码控制图像生成,损失函数如式(4)和式(5)所示:

$$L_{\text{recon}}^{\text{code1}} = \mathbb{E} \left[\|a_i - E_a(G(a_i, s_j))\|_1 \right], \quad (4)$$

$$L_{\text{recon}}^{\text{code2}} = \mathbb{E} \left[\|s_j - E_s(G(a_i, s_j))\|_1 \right]. \quad (5)$$

与自生成网络一致,同样采用身份损失判断图像身份ID是否一致。

$$L_{\text{id}}^c = \mathbb{E} \left[-\log(p(y_i | x_j^i)) \right], \quad (6)$$

其中 $p(y_i | x_j^i)$ 是预测 x_j^i 属于真实类别 y_i 的概率,并且使用对抗损失使生成数据的分布更接近真

实数据的分布。如式(7)所示:

$$L_{adv} = \mathbb{E} \left[\log D(x_i) + \log \left(1 - D(G(a_i, s_j)) \right) \right]. \quad (7)$$

生成器的结构图如图3所示,本文将DGNet网络中的生成器模块内ResBlock模块替换为DenseBlock模块。结构图中输入图像 x_i, x_l, x_j ,首先经过外观编码器得到128个 64×32 的特征

图,经过4个DenseBlock模块之后,输出依旧为 $128 \times 64 \times 32$,特征图不变。其次,进行上采样操作得到128个 128×64 的特征图,第一次卷积是 5×5 ,卷积后得到64个 128×64 的特征图;同样进行上采样操作得到64个 256×128 的特征图,进行第二次卷积操作得到32个 256×128 的特征图;Conv3、Conv4、Conv5的卷积核分别为 $3 \times 3, 3 \times 3, 1 \times 1$,最终得到 $3 \times 256 \times 128$ 大小的生成图。

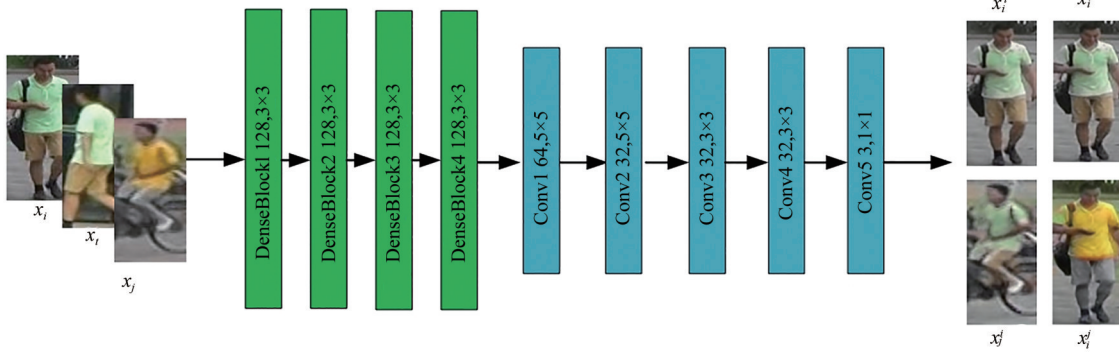


图3 生成器结构图

Fig. 3 Generator block diagram

DenseNet是Huang等人^[19]提出的一种稠密卷积网络,它主要从特征的角度提高网络性能,并且随着网络深度的增加,能够避免梯度消失问题。ResBlock主要含有一个跳跃连接,而DenseBlock中将所有的特征计算单元的结果以密连接的方式沿着某一维度进行压缩,不是进行简单的求和。DenseBlock结构如图4所示。DenseBlock结构图中除了输入端,还包含4个模块,每个模块均为4层。以 X_0 作为输入,通过 H_i 进行非线性变化,依次进行归一化、激活函数以及卷积操作,然后传递到后续网络。本文采用即插即用型模块DenseBlock代替原来的ResBlock模块,一方面能够提取更多特征信息,另一方面改善网络的特征提取能力,减少过拟合现象的出现。经过DenseBlock之后,特征图大小不变。DenseBlock结构的任何两层之间都是直接连接的,而每层的特征图也会直接传递给其后的所有层,传递方式为:

$$X_l = H_l([X_0, X_1, \dots, X_{l-1}]), \quad (8)$$

其中, $[X_0, X_1, \dots, X_{l-1}]$ 表示从第0层到第 l 层生成的特征图并且各个特征图大小一致,可在通道维度上连接。

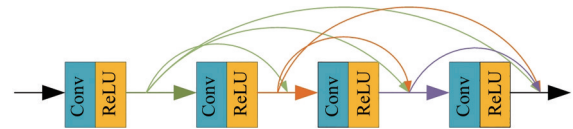


图4 DenseBlock结构图

Fig. 4 DenseBlock structure diagram

2.2 鉴别网络

鉴别网络框架图如图5所示,通过共享外观编码器 E_a 作为行人再识别学习的主干,将鉴别网络(同是识别模块)嵌入在生成模块中。生成网络通过外观编码和结构编码生成图像,鉴别网络

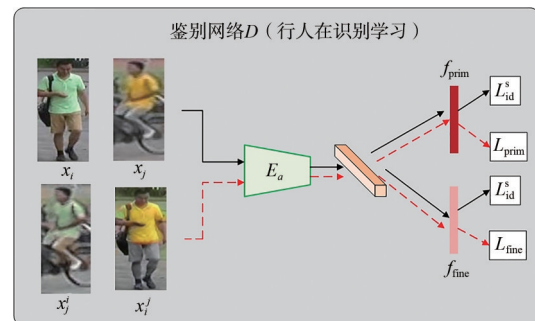


图5 鉴别网络框架图

Fig. 5 Discriminator network frame diagram

中提出特征学习与细粒度特征挖掘以及新加入的通道间的注意力机制 NAM 三者进行联合学习判断生成图像真假,以更好地利用生成图像。

(1) 主要特征学习。将真实图像 x_i, x_j 与生成图像 x_j^i, x_i^j 一起输入 E_a 中作为训练样本,由于生成图像的多样性,采用带有身份损失的卷积神经网络(即教师模型)、外观编码器学习主要特征,通过损失函数将预测概率分布与真实概率分布差异降到最低。

$$L_{\text{prim}} = \mathbb{E} \left[- \sum_{k=1}^K q(k|x_j^i) \log \left(\frac{p(k|x_j^i)}{q(k|x_j^i)} \right) \right], \quad (9)$$

其中, K 表示身份类别数量。

(2) 细粒度特征挖掘。当生成的图像用主要特征学习无法判断真假时,本文采用细粒度挖掘。当相同身份的人穿着不同的衣服或者不同身份人穿着同样的衣服时,鉴别模块采用细粒度特征挖掘,使网络被迫学习与衣服无关而与细粒度 ID 相关的特征(如头发、体型等),进一步提高

鉴别能力与行人再识别分类识别的精度。损失函数如式(9)所示:

$$L_{\text{fine}} = \mathbb{E} \left[- \log \left(p(y_i|x_j^i) \right) \right], \quad (10)$$

该损失为鉴别模块增加身份监督,不仅学习到细节纹理特征,而且有利于解决行人外观(如换衣)发生变化均导致识别下降的问题。

(3) 注意力机制 NAM。鉴别网络结构图如图 6 所示。输入原图像 x_i, x_j 与生成器生成的图像 x_j^i, x_i^j , 首先用 1×1 的卷积得到 32 个 256×128 的特征图,然后经过通道注意力 NAM,这里特征图大小不变。然后在经过 4 次卷积核为 3×3 的卷积运算,得到 64 个 64×32 的特征图,并经过 4 个 DenseBlock 模块。这里后 3 个 DenseBlock 模块前加入通道注意力 NAM,不改变通道以及特征图大小,依旧是 64 个 64×32 的特征图。最后,经过 1×1 的卷积得到一个 64×32 的特征图,给出一个 0~1 之间的数值判断图像真假,同时对图像进行分类识别。

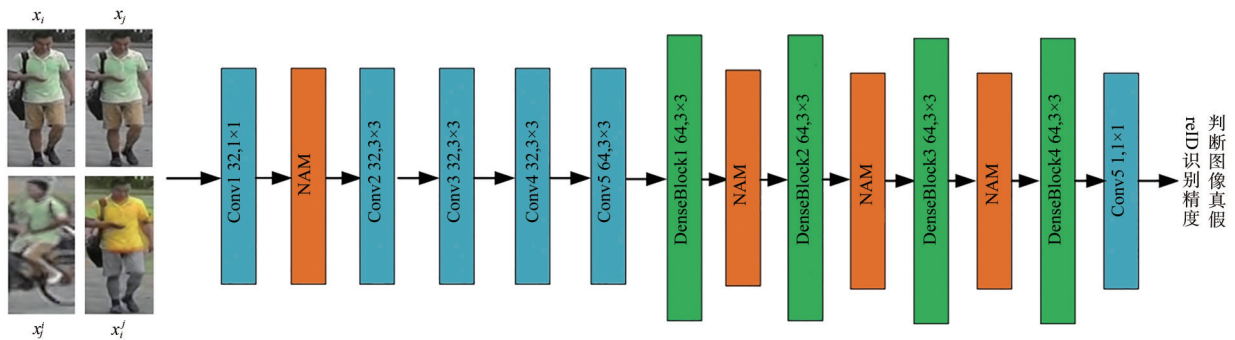


图 6 鉴别网络结构图

Fig. 6 Discriminator network structure diagram

除了主要特征学习与细粒度特征外,本文在鉴别器的第一次卷积之后与前 3 个 DensesBlock 模块之后加入了 2021 年提出的新型注意模块 NAM,主要是通过利用训练模型权重的方差度量来突出显著特征。以前的注意力机制模块大多数都是通过抑制不重要的权重来提高神经网络的性能,NAM 作为一种高效、轻量级的注意力机制,在保持相似性能的同时计算效率更高。采用 CBAM^[20] 的模块集成,并重新设计了通道和空间注意力模块,使用标准化 BN 中的比例因子,如式(11)所示,测量通道的方差并指示其重要性。

$$B_{\text{out}} = \text{BN}(B_{\text{in}}) = \gamma \frac{B_{\text{in}} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} + \beta, \quad (11)$$

式中, μ_B 和 σ_B 分别表示小批量 B 的平均值与标准偏差; γ 和 β 是可训练的仿射变换参数。通道注意力模块如图 7 所示。其中, F_1 表示输入特征, γ 是每个通道的比例因子,权重由公式 $W_\gamma =$

$$\frac{\gamma_i}{\sum_{j=0} \gamma_j} \text{ 获得, } M_c \text{ 表示输出特征, 如式(12)}$$

所示:

$$M_c = \text{sigmoid}(w_\gamma(\text{BN}(F_1))). \quad (12)$$

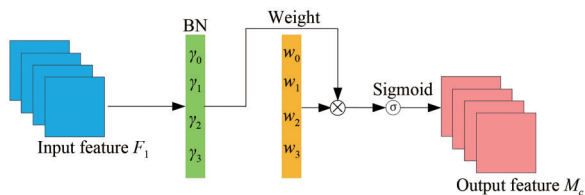


图 7 通道注意力机制 NAM

Fig. 7 Channel attention mechanism NAM

2.3 总体损失函数

由 2.1 节与 2.2 节介绍可得,总体损失函数为:

$$L_{\text{total}}(E_a, E_s, G, D) = \lambda_{\text{img}} L_{\text{recon}}^{\text{img}} + L_{\text{recon}}^{\text{code}} + L_{\text{id}}^s + \lambda_{\text{id}} L_{\text{id}}^c + L_{\text{adv}} + \lambda_{\text{prim}} L_{\text{prim}} + \lambda_{\text{fine}} L_{\text{fine}}, \quad (13)$$

其中, $L_{\text{recon}}^{\text{img}} = L_{\text{recon}}^{\text{img1}} + L_{\text{recon}}^{\text{img2}}$; $L_{\text{recon}}^{\text{code}} = L_{\text{recon}}^{\text{code1}} + L_{\text{recon}}^{\text{code2}}$ 。使用 $\lambda_{\text{img}} = 5$ 计算图像重建损失、 $\lambda_{\text{id}} = 0.5$ 计算图像身份损失,主要特征学习损失与细粒度挖掘损失之间设置权重为 $\lambda_{\text{fine}} = 0.2\lambda_{\text{prim}}$ 。

2.4 参数设置

本文是基于 PyTorch1.1、Python3.6 深度学习框架下实现的。网络模型中采用 SGD 优化器来优化外观编码器 E_a ,使用 Adam 优化器来优化结构编码器 E_s 、生成器 G 和鉴别器 D 。本文参数

$\beta_1=0, \beta_2=0.999$, 设置初始学习率为 0.001, 共 20 个 epoch, 每个 epoch 迭代训练 10 k 次, 批尺寸 (Batch size) 为 16。训练过程中保持所有图像的长宽比, 将大小统一为 256×128 。

3 实验结果与分析

3.1 数据集和实验环境

本文在行人再识别算法的公开数据集 Market1501^[21]、DukeMTMC-reID^[22] 上进行实验。Market1501 数据集共有 1 501 个行人的 32 668 个检测行人矩形框, 共 6 个摄像头, 包括训练集、测试集、查询集。DukeMTMC-reID 共有 1 812 个行人的 36 411 个检测行人矩形框, 由 8 个摄像机采集。因为有 408 人没有出现在两个摄像头以上, 所以只选择 1 404 个行人作为训练集与测试集。该数据集包括训练集、测试集、查询集。公开数据集的详细信息如表 1 所示。实验室拥有一台具备 4 块 TITAN V GPU 的图形工作站, 具备良好的开发环境。本实验将在 Ubuntu 16.04.5 LTS 操作系统, TITAN V 12G $\times$ 4 显卡, Pytorch1.1 深度学习框架, python3.6 的服务器下实现。

表 1 公开数据集详细信息

Tab. 1 Public datasets details

数据集	ID	Train ID	Train image	Test ID	Test image	Bbox/ID	Cam
Market1501 ^[21]	1 501	751	12 936	750	19 732	21.8	6
DukeMTMC-reID ^[22]	1 404	702	16 522	702	17 661	25.9	8

3.2 评价标准

行人再识别模型用首位命中率 (Rank-1)、平均精度均值 (mean Average Precision, mAP) 和累计匹配曲线 (Cumulative Match Characteristic, CMC) 3 种评价指标衡量模型。这里, Rank-k 表示在查询集中行人特征相似度大小的排序列表中前 k 位存在检索目标的准确率。

3.3 生成图像示例

改进网络通过切换姿态与外观生成大量的高质量图像。给定输入图像 x_i 和 x_j , 网络模型提取输入图像的外观特征以及结构特征, 既可以生成具有 x_i 的外观、 x_j 的结构假图像 x_j^i , 又能生成具有 x_j 的外观、 x_i 的结构假图像 x_i^j 。两个数据

集生成的图像如图 8 所示。

假设每行每列输入图像都一致, 通过切换外观或结构代码在 Market-1501、DukeMTMC-reID 上生成图像的示例如图 9 所示, 可以看出横轴行人生成外观不同、结构相同的假图像; 同样地, 纵轴行人生成外观相同、结构不同的假图像。在 Market1501、DukeMTMC-reID 数据集上, 模型分别生成共计 140 625 张、123 201 张假图像。虽然改进生成对抗网络方法能够生成大量的高质量图像, 但是由于数据集本身存在模糊、分辨率不高的问题, 并且图像的背景影响导致生成的图像会出现少许效果差的结果图像, 如图 10 所示。



图 8 交换外观与结构生成图像示例

Fig. 8 Examples of images generated by swapping appearance and structure

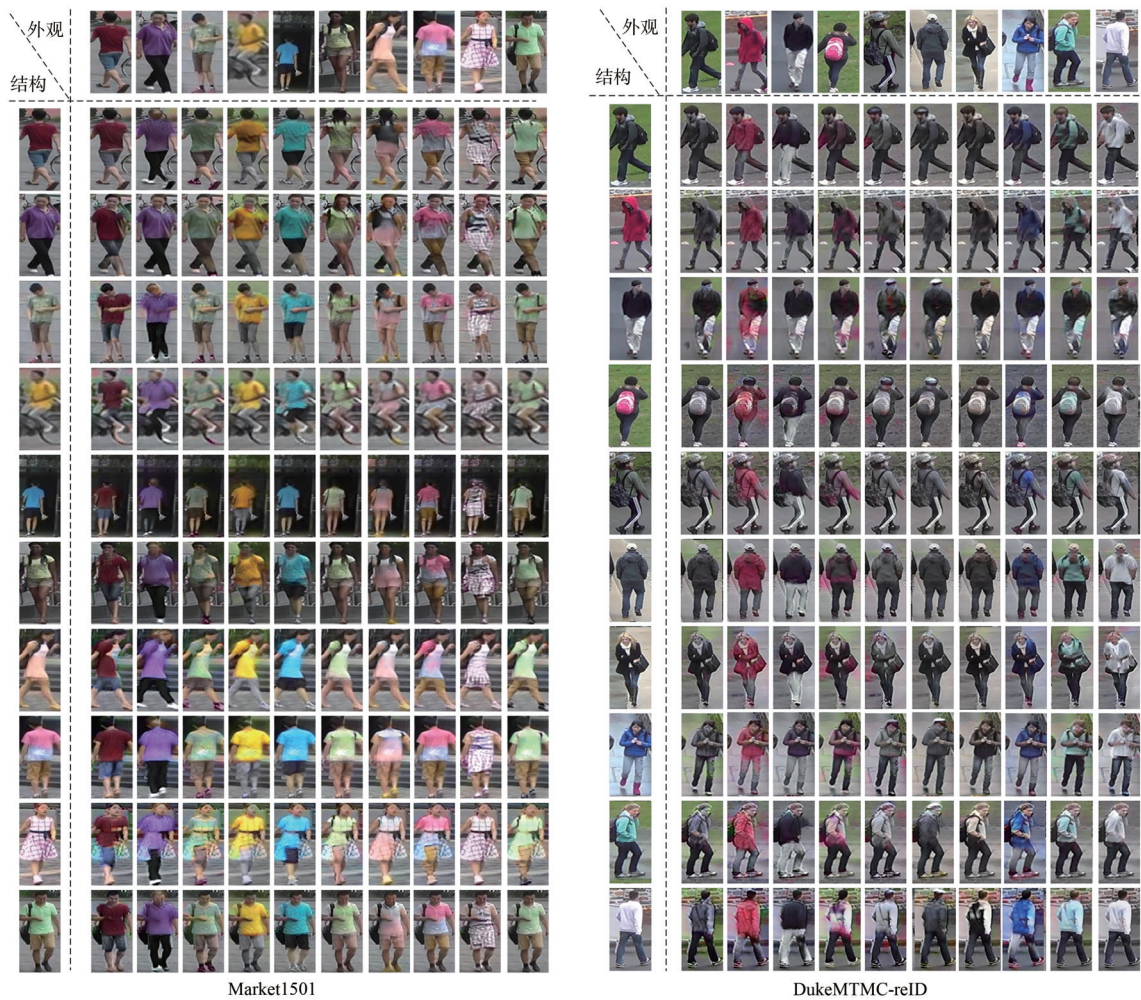


图 9 切换外观或结构代码生成图像的示例

Fig. 9 Examples of images generated by switching the appearance or structure code

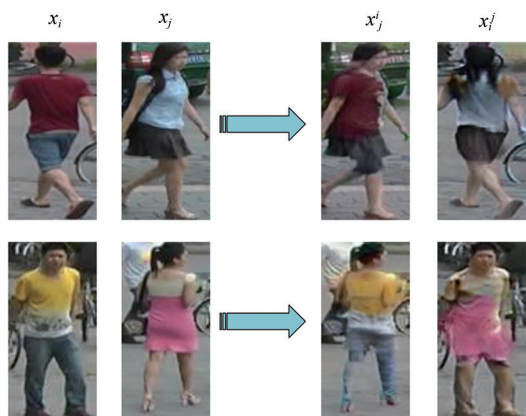


图 10 少许失败案例

Fig. 10 A few failure cases

3.4 实验结果

为了进一步证明该网络的有效性,使用相应的评价标准计算得到了该算法的 Rank-1 和 mAP,并与现有算法在 Market1501、DukeMTMC-reID 数据集上的结果进行对比分析,如表 2 所示。

表 2 本文算法与现有算法的比较

Tab. 2 Comparison of the algorithm in this paper with existing algorithms

Model	Market1501		DukeMTMC-reID	
	Rank-1/%	mAP/%	Rank-1/%	mAP/%
PTGAN ^[12]	87.7	68.9	78.5	56.9
PNGAN ^[14]	89.4	72.6	73.6	53.2
FDGAN ^[15]	90.5	77.4	80.0	64.5
DG-Net ^[17]	94.8	86.0	86.6	74.8
CSL+MGL ^[23]	95.3	86.1	88.1	76.4
PAGAN ^[24]	94.1	85.2	84.7	74.2
AP-GAN ^[13]	93.4	82.2	84.3	70.5
TransReID ^[25]	95.2	89.5	90.7	82.6
Euc ^[26]	93.7	79.7	88.3	76.8
FGSAM ^[27]	91.5	85.4	85.9	74.1
HOReID ^[28]	94.2	85.1	86.9	75.6
MOPN ^[29]	94.6	85.1	88.0	76.2
PAFM ^[30]	95.1	88.2	90.6	78.6
Ours	95.6	88.6	87.1	75.7

注:粗体数字表示本文算法最优值

将本文算法与 8 种基于 GAN 的 ReID 方法和 5 种有监督的 ReID 方法进行性能比较。针对基于 GAN 的 ReID 方法性能比较可知:对于 Mar-

ket1501 数据集,从表 2 的实验结果可以分析到本文算法在 Rank-1 和 mAP 的指标上均优于其他算法(除 TransReID 之外),Rank-1 相比于 DG-Net、CSL+MGL、PAGAN、AP-GAN、TransReID 分别提高了 0.8%、0.3%、1.5%、2.2%、0.4%;mAP 相比于 DG-Net、CSL+MGL、PAGAN、AP-GAN、Trans-ReID 分别提高了 2.6%、2.5%、3.4%、6.4%、-0.9%。对于 DukeMTMC-reID 数据集,本文算法在 Rank-1 和 mAP 的指标上仅次于 CSL+MGL、TransReID 算法。Rank-1 相比于 DG-Net、CSL+MGL、PAGAN、AP-GAN、TransReID 分别提高了 0.9%、-0.7%、1.5%、5.2%、-6.9%。针对 5 种有监督的 ReID 方法性能比较可知:对于 Market1501 数据集,本文算法的 Rank-1 和 mAP 均优于这 5 种算法,Rank-1 和 mAP 分别比 Euc、FGSAM、HOReID、MOPN、PAFM 提高了 1.9%、4.1%、1.4%、1.0%、0.5% 和 8.9%、3.2%、3.5%、3.5%、0.4%;对于 DukeMTMC-reID 数据集,本文算法的 Rank-1 和 mAP 高于 FGSAM、HOReID 算法,Rank-1 和 mAP 分别比 FGSAM、HOReID 高 1.2%、0.2% 和 1.6%、0.1%。说明本文算法在 Market1501 上的效果更显著。同时进一步说明本文算法具有一定的优势。综上所述,所提算法不仅能够较好地提高行人识别的精度,还使模型有较强的泛化能力。

3.5 不同改进点的消融实验

本文主要有 2 点改进:(1)使用新型注意力机制 NAM;(2)将原有的 DGNet 网络中的 Res-Blocks 替换为 DenseBlocks,采用像素损失、对抗损失、身份损失等多种损失联合优化提高生成图的质量。以 Market1501 为例,对不同改进点进行性能评估,如表 3 所示。

表 3 不同改进点的性能评估

Tab. 3 Performance evaluation of different improvement points

不同改进点		Market1501		DukeMTMC-reID	
(1)	(2)	Rank-1/%	mAP/%	Rank-1/%	mAP/%
×	×	94.8	86.0	86.6	74.8
×	✓	95.0	87.6	86.9	75.3
✓	×	95.2	87.9	86.8	75.0
✓	✓	95.6	88.6	87.1	75.7

4 结 论

本文提出改进生成对抗网络的换衣行人再识别算法,是一种将生成对抗网络(GAN)数据生成和行人再识别学习端对端结合起来的联合框架,通过切换外观与结构编码,生成模块中的自生成网络与互生成网络生成大量图像,鉴别模块与生成模块之间在线循环,将生成的图像在线反馈给鉴别器,鉴别器通过主要特征学习、细粒度特征挖掘与注意力机制 NAM 三者

联合学习,进一步提高鉴别器鉴别能力与行人再识别分类识别的准确率。生成器与鉴别器中的残差块 ResBlock 替换为 DenseBlock,加强特征传播并且大幅度减少参数数量。本文算法不仅解决了行人再识别数据集不足的问题,而且提高了跨摄像头下由于行人外观(如衣服)改变的识别率,在图像生成质量和行人再识别分类识别精度方面都有显著的改进,在 Market1501 和 DukeMTMC-reID 数据集上 Rank-1 和 mAP 分别达到了 95.7%/88.6% 和 87.1%/75.7%。

参 考 文 献:

- [1] LIAO S C, HU Y, ZHU X Y, *et al.* Person re-identification by local maximal occurrence representation and metric learning [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston: IEEE, 2015: 2197-2206.
- [2] ZHENG L, YANG Y, HAUPTMANN A G. Person re-identification: past, present and future [J]. *arXiv*, 2016: 1610.02984.
- [3] CHEN B H, DENG W H, HU J N. Mixed high-order attention network for person re-identification [C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul: IEEE, 2019: 371-381.
- [4] CHEN D P, XU D, LI H S, *et al.* Group consistent similarity learning via deep CRF for person re-identification [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 8649-8658.
- [5] GE Y X, CHEN D P, LI H S. Mutual mean-teaching: pseudo label refinery for unsupervised domain adaptation on person re-identification [C]//*Proceedings of the 8th International Conference on Learning Representations*. Addis Ababa: ICLR, 2020.
- [6] SI J L, ZHANG H G, LI C G, *et al.* Dual attention matching network for context-aware feature sequence based person re-identification [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City: IEEE, 2018: 5363-5372.
- [7] YU S J, LI S H, CHEN D P, *et al.* COCAS: a large-scale clothes changing person dataset for re-identification [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020: 3397-3406.
- [8] ZHENG Z D, ZHENG L, YANG Y. Pedestrian alignment network for large-scale person re-identification [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 29(10): 3037-3045.
- [9] SUH Y, WANG J D, TANG S Y, *et al.* Part-aligned bilinear representations for person re-identification [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 418-437.
- [10] 张良,车进,杨琦.多粒度特征融合的行人再识别研究[J].液晶与显示,2020,35(6):555-563.
ZHANG L, CHE J, YANG Q. Multi-granularity feature fusion for person re-identification [J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(6): 555-563. (in Chinese)
- [11] ZHENG Z D, ZHENG L, YANG Y. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro [C]//*Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, 2017: 3374-3782.
- [12] WEI L H, ZHANG S L, GAO W, *et al.* Person transfer GAN to bridge domain gap for person re-identification [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake

- City: IEEE, 2018: 79-88.
- [13] 杨琦,车进,张良,等. GAN网络混合编码的行人再识别[J]. 液晶与显示, 2021, 36(2): 334-342.
YANG Q, CHE J, ZHANG L, *et al.* Person re-identification of GAN network hybrid coding [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(2): 334-342. (in Chinese)
- [14] QIAN X L, FU Y W, XIANG T, *et al.* Pose-normalized image generation for person re-identification [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 661-678.
- [15] GE Y X, LI Z W, ZHAO H Y, *et al.* FD-GAN: pose-guided feature distilling GAN for robust person re-identification [C]//*Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montréal: Curran Associates Inc., 2018: 1230-1241.
- [16] 杨海伦,王金聪,任洪娥,等. 基于DCN的无监督域自适应行人重识别方法[J]. 液晶与显示, 2021, 36(11): 1573-1582.
YANG H L, WANG J C, REN H E, *et al.* DCN-based unsupervised domain adaptive person re-identification method [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(11): 1573-1582. (in Chinese)
- [17] ZHENG Z D, YANG X D, YU Z D, *et al.* Joint discriminative and generative learning for person re-identification [C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2019: 2133-2142.
- [18] LIU Y C, SHAO Z R, TENG Y Y, *et al.* NAM: normalization-based attention module [J]. *arXiv*: 2111.12419, 2021.
- [19] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks [C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 2261-2269.
- [20] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 3-19.
- [21] ZHENG L, SHEN L Y, TIAN L, *et al.* Scalable person re-identification: a benchmark [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago: IEEE, 2015: 1116-1124.
- [22] RISTANI E, SOLERA F, ZOU R, *et al.* Performance measures and a data set for multi-target, multi-camera tracking [C]//*Proceedings of European Conference on Computer Vision*. Amsterdam: Springer, 2016: 17-35.
- [23] 张师林,曹旭. 基于Camstyle改进的行人重识别算法[J]. 计算机工程与应用, 2020, 56(15): 124-131.
ZHANG S L, CAO X. Improved person re-identification algorithm on Camstyle [J]. *Computer Engineering and Applications*, 2020, 56(15): 124-131. (in Chinese)
- [24] ZHANG Y X, CHE J, HE Y T. Research on person re-identification based on pose-attention-GAN [C]//*Proceedings 2021 International Conference on Computer Information Science and Artificial Intelligence*, 2021: 261-2.
- [25] HE S T, LUO H, WANG P C, *et al.* TransReID: transformer-based object re-identification [C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021: 14993-15002.
- [26] LI Z Y, JIN Y, LI Y D, *et al.* Learning part-alignment feature for person re-identification with spatial-temporal-based re-ranking method [J]. *World Wide Web*, 2020, 23(3): 1907-1923.
- [27] ZHOU Q Q, ZHONG B N, LAN X Y, *et al.* Fine-grained spatial alignment model for person re-identification with focal triplet loss [J]. *IEEE Transactions on Image Processing*, 2020, 29: 7578-7589.
- [28] WANG P Y, ZHAO Z C, SU F, *et al.* HOReID: deep high-order mapping enhances pose alignment for person re-identification [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2908-2922.
- [29] 唐一星. 融合全局特征与局部特征的行人再识别研究[D]. 合肥: 合肥工业大学, 2021.
TANG Y X. Research on person re-identification based on fusing global features and local features [D]. Hefei: Hefei University of Technology, 2021. (in Chinese)
- [30] YANG J, ZHANG C L, TANG Y P, *et al.* PAFM: pose-drive attention fusion mechanism for occluded person re-identification [J]. *Neural Computing and Applications*, 2022, 34(10): 8241-8252.

作者简介:



张玉霞(1997—),女,宁夏中卫人,硕士研究生,2019年于宁夏大学获得学士学位,主要从事图像处理、行人再识别、智能视频等方面的研究。E-mail: 1064316204@qq.com



车进(1973—),男,宁夏银川人,博士,教授,2014年于天津大学获得博士学位,主要从事图像处理、行人再识别、智能视频等方面的研究。E-mail: koalache@126.com