

GAN网络混合编码的行人再识别

杨琦, 车进, 张良, 张玉霞

引用本文:

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 Qi, CHE Jin, ZHANG Liang, et al. Person re-identification of GAN network hybrid coding[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(2): 334–342.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2020-0167>

您可能感兴趣的其他文章

Articles you may be interested in

多粒度特征融合的行人再识别研究

Multi-granularity feature fusion for person re-identification

液晶与显示. 2020, 35(6): 555–563 <https://doi.org/10.3788/YJYXS20203506.0555>

基于条件生成对抗网络的手写数字识别

Handwritten digit recognition based on conditional generative adversarial network

液晶与显示. 2020, 35(12): 1284–1290 <https://doi.org/10.37188/YJYXS20203512.1284>

基于深度卷积生成对抗网络的图像识别算法

Image recognition algorithms based on deep convolution generative adversarial network

液晶与显示. 2020, 35(4): 383–388 <https://doi.org/10.3788/YJYXS20203504.0383>

基于改进YOLOv3的棉花异性纤维检测

Detection of foreign fiber in cotton based on improved YOLOv3

液晶与显示. 2020, 35(11): 1195–1203 <https://doi.org/10.37188/YJYXS20203511.1195>

多尺度特征融合与极限学习机结合的交通标志识别

Traffic sign recognition based on multi-scale feature fusion and extreme learning machine

液晶与显示. 2020, 35(6): 572–582 <https://doi.org/10.3788/YJYXS20203506.0572>

文章编号:1007-2780(2021)02-0334-09

GAN 网络混合编码的行人再识别

杨 琦^{1,2}, 车 进^{1,2*}, 张 良^{1,2}, 张玉霞^{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:由于摄像机视角造成的类内差异明显,研究学者开始利用 GAN 扩充数据保持类内不变性。针对现有 GAN 生成图像模糊、背景不真实,本文提出一种利用姿态与外观特征混合编码的行人再识别算法。在训练阶段,将人物图像分解为姿态特征和外观特征,生成网络通过切换外观特征与姿态特征,混合两幅图像中的特征生成高质量图像。判别网络将生成图像的外观特征反馈给生成网络的外观编码器以实现联合优化,采用多损失联合进一步提高生成图片的质量。在测试阶段,使用原数据集对网络模型进行测试,在 Market-1501 和 DukeMTMC-reID 数据集上的 rank-1/mAP 分别达到 93.4%/82.2%、84.3%/70.5%。

关键词:行人再识别;深度学习;GAN 网络;外观特征;姿态特征

中图分类号:TP392 **文献标识码:**A **doi:**10.37188/CJLCD.2020-0167

Person re-identification of GAN network hybrid coding

YANG Qi^{1,2}, CHE Jin^{1,2*}, ZHANG Liang^{1,2}, ZHANG Yu-xia^{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,

Ningxia University, Yinchuan 750021, China)

Abstract: Due to the obvious intra-class differences caused by camera perspective, many researchers begin to use GAN to expand data to maintain intra-class invariance. Nevertheless, the images generated by existing GAN have some defects such as blurred image and unreal background. To solve the above existing problems, a person re-identification algorithm that uses mixed coding of posture and appearance features is proposed in this paper. In the training phase, the image of the person is decomposed into posture features and appearance features, and the generated network can generate high-quality images by switching the appearance feature and posture feature and then mixing the features in the two images. The discriminant network feeds back the appearance characteristics of the generated image to the appearance coder of the generated network to achieve joint optimization, and uses multi-loss joint to further improve the quality of the generated image. In the testing phase, the network

收稿日期:2020-06-23;修订日期:2020-09-04.

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

model is tested using the original data set. The rank-1/mAP on the Market-1501 and DukeMTMC-reID data sets can reach 93.4%/82.2% and 84.3%/70.5% respectively.

Key words: person re-identification; deep learning; GAN network; appearance features; attitude features

1 引言

行人再识别^[1] (Person ReID) 可以看作是一个跨摄像机视角的人物检索问题,旨在建立多个摄像机图像之间的身份对应关系。由于拍摄角度、光照、姿势、视角、图像分辨率、相机设置、遮挡和背景杂波的影响,会导致同一行人的不同图像可能会有显著不同,造成较大的类内差异,这使得行人再识别仍然是一项充满挑战性的任务。

随着深度学习在行人再识别任务中的广泛应用,卷积神经网络由于其强大的特征表达能力以及学习不变的特征嵌入,近年来涌现出各种各样的深度学习算法,尤其在 GAN 网络方面取得不错的进展。生成对抗网络最初是由 Goodfellow 等人^[2]提出,被描述为一个通过对抗训练生成模型的过程。GAN 由生成图像的生成器(G)和鉴别器(D)组成,这两个组件在极小极大值之间进行博弈。文献[3]提出将 GAN 扩展到 CNN 领域,使得利用 GAN 获取的训练样本更加可控,进一步说明了 GAN 网络在计算机视觉任务中的可行性。

众所周知,深度学习的发展得益于大数据的发展,而在如今行人再识别课题中,面临着数据不足与类内差异明显等问题。作为 GAN 网络的先行者,文献[4]提出一种标签平滑的方法,利用生成的数据扩充原始数据集,一定程度上提高了行人再识别的精度。不同于文献[4]采用标签平滑对生成图像的标签采用平均的策略,文献[5]采用伪标签的策略,对生成图像采用最大概率预测为其分配身份,作为具备真实标签的数据使用。文献[6]提出一种识别模型与 GAN 中的判别器共享权重进行联合优化。文献[7]基于不同摄像机类内差异,生成不同相机风格的行人图像。此外,最近的一些研究学者开始将姿态估计应用到 GAN 网络中。文献[8]为减小姿态不同对行人外表的影响,使用 PN-GAN 将数据中的所有行人归一化到 8 个姿态中,将真实数据中提取到的行

人特征和生成数据中提取到的行人特征融合之后做 ReID 匹配。文献[9]提出一种基于姿态迁移的 ReID 框架,通过引入姿态样本库,进而生成多姿态标签样本。文献[10]利用姿态引导的 GAN 网络,学习与身份相关且与姿态无关的特征,使得生成的行人图像与姿态特征无关。不同于上述算法,也有学者将不同特征进行融合得到新的特征表示。文献[11]提出一种多尺度残差网络模型,融合不同的特征得到最终的特征表示。文献[12]提出一种融合了全局特征、局部特征以及人体结构特征的行人再识别算法,该算法无需引入任何人体框架先验知识,并采用多级监督机制优化网络。文献[13]提出一种利用姿态迁移来生成行人图片,并利用两种不同的独立卷积神经网络提取图像特征,融合两种特征得到最后的特征表示。

不同于上述 GAN 网络,本文提出一种基于外观特征和姿态特征混合编码的行人再识别网络,生成模型通过切换外观特征以及姿态特征,结合两幅图像中的特征混合编码生成高质量图像,进一步降低了类内差异造成的影响。网络采用外观损失、姿态损失、对比损失、判别损失等多损失函数对生成的图像进行监督,进一步提高生成图像的质量。利用扩充数据集对网络进行训练,使得网络模型更加健壮。

2 网络架构

网络架构如图 1 所示。将原数据集中的人物图像输入到生成对抗网络,利用输入图像的姿态特征以及外观特征进行自图像以及互图像生成,将生成的人物图像结合原数据集中的人物图像共同输入到卷积神经网络,对网络进行训练,一方面扩充了原数据集中图像不足的问题,另一方面利用这种自图像与互图像生成模式进一步挖掘了图像的细粒度特征,使得训练的模型更加鲁棒。

2.1 生成网络

本文采用 $X = \{x_i\}_{i=1}^N$ 和 $Y = \{y_i\}_{i=1}^N$ 表示数据集中的人物图像以及对应的标签,其中 N 表示

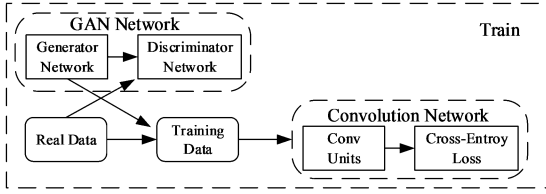


图 1 网络架构

Fig.1 Network architecture

数据集中人物图像的总数量, $y_i \in [1, K]$ 其中 K 表示数据集中类别的数量。在训练阶段, 给定两幅人物图像 x_i 和 x_j , 生成网络分别提取各自对应的外观特征以及姿态特征, 利用两种特征的混合编码生成新图像^[14]。生成网络包括自生成网络和互生成网络两部分, 其中生成网络由外观编码 $I_a: x_i \rightarrow a_i$ 、结构编码 $I_s: x_i \rightarrow s_i$ 、生成解码 $O: (a_i, s_i) \rightarrow x_j^i$, 以及鉴别器 D 来区分生成的图像和原始图像。

2.1.1 自生成网络

给定一对行人图像 $x_i, x_j \in X$ 且 $y_i = y_j$, 对于输入图像 x_i , 自生成网络首先提取其外观特征 a_i 以及姿态特征 s_i , 利用提取的外观特征以及姿态特征进行图像生成。将原图分解为两个特征分支, 然后利用这两个特征分支对其进行重新生成, 相当于对原图像的重构, 如图 2 的上层分支所示。考虑到生成图像 x_i^i 与原图像 x_i 中的像素对应关系, 因此采用像素级的 L_1 损失对其进行训练, 损失函数为:

$$L_1^i = E [\|x_i - O(a_i, s_i)\|_1] \quad (1)$$

考虑到 $y_i = y_j$, 即为同一行人的不同图像, 行人图像的外观特征是相近的, 因此提出一种利用同一行人的不同图像来生成图像的方法。即采用图像 x_i 的姿态特征, 仅采用图像 x_j 的外观特征。由于外观特征是相似的, 所以基于同一行人的图像生成应该无限接近于原输入图像 x_i , 因此仍然采用像素级的 L_1 损失对其进行训练, 损失函数为:

$$L_1^j = E [\|x_i - O(a_j, s_i)\|_1] \quad (2)$$

输入图像是来自同一行人的不同图像, 外观特征是相近的, 因此使用同一个姿态特征利用相似的外观特征生成的图像也是相似的。所以生成图像 x_i^i, x_j^j 的外观特征也应该是相似的。因此, 针对这种假设, 提出一种针对生成图像的判别方法, 如图 3 所示。

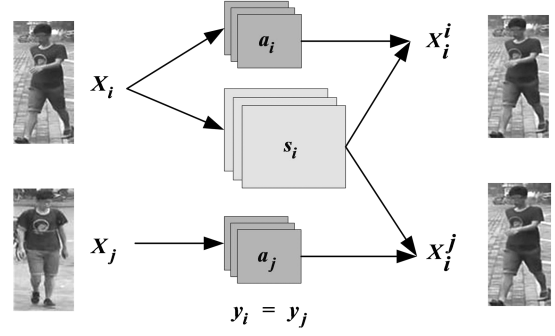


图 2 自生成网络

Fig.2 Self generated network

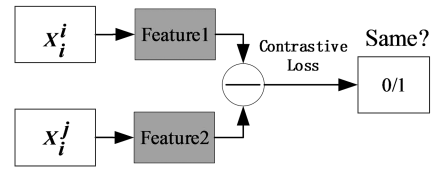


图 3 验证网络

Fig.3 Verified network

生成图像 x_i^i, x_j^j 都是基于特定姿态的生成图像, 基于上述假设分别提取生成图像的外观特征 f_1, f_2 , 采用欧式距离来度量特征相似度:

$$d = \|f_1 - f_2\|_2 \quad (3)$$

采用对比损失^[15]优化网络具体公式如下:

$$L_1^c = \frac{1}{2N} \sum_{n=1}^N y d^2 + (1-y) \max(m-d, 0)^2 \quad (4)$$

式中, d 表示两个样本特征的二范数, y 为两个样本是否匹配的标签, $y=1$ 表示匹配, m 为设定的阈值, N 为样本的个数。

2.1.2 互生成网络

与自生成网络相比, 互生成网络关注的是两幅不相关的图像经过混合编码后生成新图像的过程, 如图 4 所示。给定两幅输入图像 $x_i, x_j \in X$, 且 $y_i \neq y_j$, 首先提取图像 x_i, x_j 外观特征 a_i, a_j 以及姿态特征 s_i, s_j , 其次, 利用外观特征与姿态特征交叉组合生成新图像。对于生成图像 x_i^j , 是由图像 x_i 的姿态特征 s_i 以及图像 x_j 的外观特征 a_j 经过混合编码生成的。所以生成图像的姿态编码应该与图像 x_i 的姿态特征 s_i 相似, 外观编码应该与图像 x_j 的外观特征相似, 采用 L_1 损失对其训练, 损失函数如下:

$$L_1^a = E [\|a_j - I_a(O(a_j, s_i))\|_1] \quad (5)$$

$$L_1^s = E [\|s_i - I_s(O(a_j, s_i))\|_1] \quad (6)$$

利用混合编码对原始数据进行图像生成,使得生成的图像更加逼真,一方面,有效扩充了数据集。另一方面,有效减缓了类内差异的影响。采用多损失优化网络进一步提高了图像的真实性,有效解决了行人不够真实、图像模糊、背景不真实等问题。

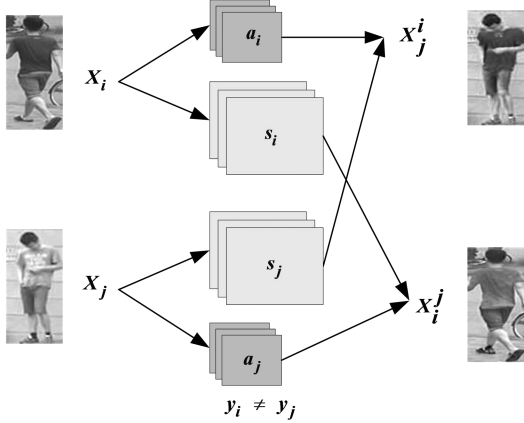


图4 互生成网络框架

Fig.4 Mutual generation network framework

2.1.3 基于外观特征的 ID 分配

网络提取了人物图像的姿态特征以及外观特征,由于行人图像在不同摄像机的视角下姿态是各异的,所以姿态特征并不具备区分不同行人的特性。在跨摄像机视角中,外观特征的不变性可以作为区分不同属性的行人。考虑到这个问题,首先训练一个基于外观特征对行人图像进行身份鉴别的网络模型,提取原始数据集中所有图像的外观特征,保留其标签属性,采用交叉熵损失对网络进行训练,损失函数如下:

$$\hat{y}_i = P(y_i = 1 | a_i), \quad (7)$$

式中, a_i 表示从图像 x_i 中提取的外观特征, $\hat{y}_i$ 是基于外观代码预测 x_i 属于类 y_i 的概率。

式(7)为单个样本的损失,总样本的损失可以表示为:

$$L_2 = -\sum_{i=1}^N y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i), \quad (8)$$

式中, $\hat{y}_i$ 是基于外观代码预测 x_i 属于类 y_i 的概率, N 表示原始数据集中图像总数量。

对于生成图像 x_j^i , 由于姿态特征并不具备身份识别的属性,所以利用生成图像的外观特征作

为身份属性的识别特征。采用交叉熵损失对网络进行训练,为了保持身份属性的不变性以及连续性,采用以下公式对网络进行训练:

$$L_2^s = E [-\log(p(y_i | x_j^i))], \quad (9)$$

式中, $p(y_i | x_j^i)$ 是基于外观代码预测 x_j^i 属于 y_i 的概率。

2.2 判别网络

生成器 G 和判别器 D 在极小极大博弈中扮演了两个竞争对手的角色, D 作为一个判别网络(如图5所示)将原数据集图像与生成图像共同输入到判别网络,提取特征利用交叉熵损失优化判别网络。

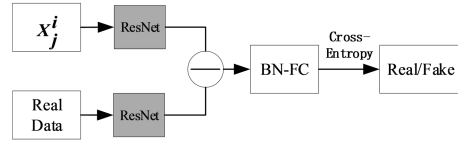


图5 判别网络模型

Fig.5 Discriminant network model

网络优化的目的是让 $D(x_i)$ 无限接近于 1, $D(O(a_i, s_j))$ 尽可能大,使用对抗性损失^[16]来匹配生成图像的分布与真实数据的分布如下:

$$L_3 = E [\log D(x_i) + \log (1 - D(O(a_i, s_j)))], \quad (10)$$

式中, $D(x_i)$ 表示判断真实图片是否真实的概率。

对于相同的框架特征,将使用不同的外观特征进行图像合成的图像属性视为与提供框架特征的行人具有相同的身份属性。也就是说,可以看到同一位行人穿着不同的衣服,这迫使网络模型学习与衣服等特征无关的特征表示,从而迫使网络模型挖掘出更多的判别特征(图,背包等),进一步挖掘图像中的细粒度信息,并增强网络模型的鲁棒性。损失函数可表示为:

$$L_4^j = E [-\log(p(y_j | x_j^i))], \quad (11)$$

其中, $p(y_j | x_j^i)$ 是基于框架特征 x_j^i 属于 y_j 的概率。

2.3 目标函数

在训练阶段,网络优化了外观损失姿态损失、验证损失、以及判别损失作为优化的总目标,如式(12):

$$L = L_1^i + L_1^j + L_1^c + L_1^a + L_1^s + L_2 + L_2^s + L_3 + L_4^j. \quad (12)$$

2.4 网络设置

基于 PyTorch 深度框架搭建网络模型,在训练阶段,采用 ResNet50 作为基准网络提取外貌特征,训练的基线网络仅仅依据外貌特征对图像进行分类。采用残差块与卷积层组合的轻量级网络^[17]提取姿态特征。验证网络采用了 DenseNet121^[18]提取生成图像的外貌特征。生成网络^[19]是由残差块经过下采样输入到卷积单元组成的,判别网络^[20]是由 6 个卷积层和一个残差块组成。所有图像的宽高比为 128×384 ,参数 m 设置为 1,并且通过 SGD 方法优化和迭代网络。初始学习率设置为 0.001。

3 实验数据库

3.1 数据集

本文提出的行人再识别算法在公开的数据

集 Market1501^[21]、DukeMTMC-reID^[22] 上进行实验并取得不错的效果。本文使用累积匹配特征曲线(CMC)和平均精度均值(mAP)两个指标来衡量模型的性能。表 1 列出了数据集的详细信息。

Market1501 是一个大型的行人数据集,采集了 6 个摄像机的数据,包含 751 个行人的 12 936 张训练图像,750 个行人的 19 732 张测试图像,边界框直接由可变形零件模型(DPM)^[23]检测,这更接近于真实的场景,采用训练集中的 12 936 张图像训练网络,在 single-shot 模式下进行。

DukeMTMC-reID 是由 8 个摄像机采集的 1 812 个行人图像,在数据集中有 1 404 个行人出现在两个摄像机以上的视角中,随机选择 702 个行人的图像作为训练集,剩余 702 个行人图像作为测试集。

表 1 数据集详细信息

Tab.1 Details of the dataset

Dataset	ID	Train ID	Test ID	Bbox	Bbox/ID	Cam
Market1501	1 501	751	750	32 668	21.8	6
DukeMTMC-reID	1 404	702	702	36 411	25.9	8

3.2 生成图像示例

对于输入图像 x_i, x_j , 网络模型分别提取输入图像的外观特征以及姿态特征,利用不同图像的交叉特征生成图像。生成图像 x_j^i 既保存

了图像 x_i 的外观特征,又保存了图像 x_j 的姿态特征。不同数据集的生成图像示例如图 6 所示。

3.3 实验结果

本文采用多损失函数联合监督机制优化网络迭代。网络的总损失包括像素级的损失 $L^1 = L_1^i + L_1^j + L_1^c + L_1^a$;根据外观分配的 ID 损失 $L^2 = L_2 + L_2^g$;区分网络中的生成损失 $L^3 = L_3$;网络框架特征的损失 $L^4 = L_3^j$ 。表 2 列出了不同损失函数对网络模型的影响。从表 2 中数据可以看出,多损失联合监督机制可以有效地挖掘图像中的细粒度特征,增强网络模型的鲁棒性,并进一步提高行人识别的准确性。

本文对提出的利用姿态特征以及外貌特征混合编码的行人再识别算法与现有的行人再识别算法进行了比较,如表 3 所示。表 3 中分割线以上为未采用生成图像扩充数据的算法,分割线

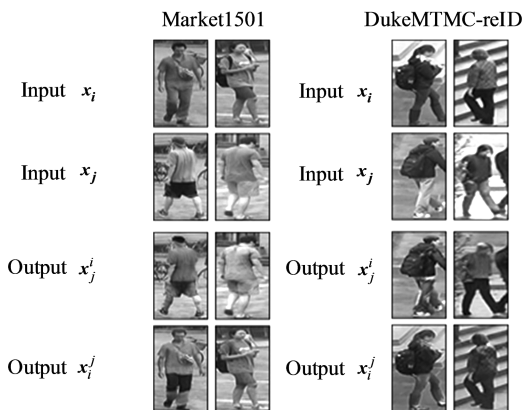


图 6 生成图像示例

Fig.6 Generated image example

表2 不同损失函数对模型的影响

Tab.2 Influence of different loss functions on the model

Model	Market1501		DukeMTMC-ReID	
	Rank-1/%	mAP/%	Rank-1/%	mAP/%
L^1+L^3	93.0	79.9	83.9	69.7
$L^1+L^3+L^2$	93.8	83.1	84.5	71.3
$L^1+L^3+L^4$	93.4	82.2	84.2	70.5
$L^1+L^3+L^2+L^4$	94.1	84.6	84.9	72.1

以下为采用生成图像辅助训练的算法。由表中的实验数据可以看出,采用姿态特征和外貌特征混合编码的行人再识别算法后,在Market1501数据集上的表现效果较好,Rank-1仅比PCB算法稍低0.4%,但mAP的性能却高于PCB算法0.6%;在DukeMTMC-ReID数据

集上的Rank-1仅仅低于Part-aligned、Manes算法不到一个百分点,而mAP的性能仅低于Manes算法。综上所述,本文提出的算法在两大公开的数据集上表现效果较好,Rank-1、mAP评估指标能优于现有的大部分主流算法,可以看出所提算法的优越性。

表3 本文算法与现有算法进行比较

Tab.3 Comparison of the algorithm in this paper with existing algorithms

Model	Market1501		DukeMTMC-ReID	
	Rank-1/%	mAP/%	Rank-1/%	mAP/%
IDE[24]	72.52	46.0	67.7	47.1
LOMO+XQDA[1]	72.58	51.96	30.75	17.04
DCF[25]	80.3	57.5	—	—
SSM[26]	82.2	68.8	—	—
SVDNet[27]	82.3	62.1	76.7	56.8
PAN[28]	82.8	63.4	71.59	51.51
DPFL[29]	88.9	73.1	79.2	60.0
MLFN[30]	90.0	74.3	81.0	62.8
HA-CNN[31]	91.2	75.7	80.5	63.8
Part-aligned[32]	91.7	79.6	84.4	69.3
Manes[33]	93.1	82.3	84.9	71.8
PCB[34]	93.8	81.6	83.3	69.2
DeformGAN[35]	80.6	61.3	—	—
LSRO[4]	84.0	66.1	67.7	47.1
Multi-pseudo[6]	85.8	67.5	76.8	58.6
PT[8]	87.7	68.9	78.5	56.9
PN-GAN[9]	89.4	72.6	73.6	53.2
FD-GAN[10]	90.5	77.7	80.0	64.5
Ours	93.4	82.2	84.3	70.5

4 结 论

本文提出一种利用姿态以及外观特征混合编码生成图像的行人再识别算法。采用多损失监督的方式修正生成图像,使得生成模块与判别模块是一个在线的交互循环,使得两者相互受益。生成模型通过切换外观特征以及结构特征,结合两幅图像中的特征混合编码生成高质量图像,判别

模型将生成图像的外观特征反馈给生成模型的外观编码器,通过联合优化,进一步提高生成图片的质量。一方面,解决了数据集不足的问题,另一方面,进一步解决了行人图像不真实、模糊、背景不真实等问题。这种利用扩充数据集训练网络的方式,使得网络模型更加鲁棒。两个数据集的实验结果显示,算法的 Rank-1 指标相比于 FD-GAN 方法提升了 2.9%、4.3%,相比于 mAP 提升了 4.5%、6%。

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, MA: IEEE, 2015: 2197-2206.
- [2] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial networks [C]//*Proceedings of the Advances in Neural Information Processing Systems (NIPS)*. New York: NIPS, 2014: 2672-2680.
- [3] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [C]//*Proceedings of the 4th International Conference on Learning Representations (ICLR)*. San Juan, Puerto Rico, 2016.
- [4] 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.
- [5] HUANG Y, XU J S, WU Q, *et al.* Multi-pseudo regularized label for generated data in person re-identification [J]. *IEEE Transactions on Image Processing*, 2019, 28(3): 1391-1403.
- [6] LI X, WU A C, ZHENG W S. Adversarial open-world person re-identification [C]//*Proceedings of the 15th European Conference on Computer Vision-ECCV 2018*. Munich, Germany: Springer, 2018.
- [7] CHOI Y, CHOI M, KIM M, *et al.* StarGAN: unified generative adversarial networks for multi-domain image-to-image translation [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 8789-8787.
- [8] LIU J X, NI B B, YAN Y C, *et al.* Pose transferrable person Re-identification [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 4099-4108.
- [9] QIAN X L, FU Y W, XIANG T, *et al.* Pose-normalized image generation for person re-identification [C]//*Proceedings of 15th European Conference on Computer Vision-ECCV 2018*. Munich, Germany: Springer, 2018.
- [10] 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, Canada: ACM, 2018.
- [11] 贡卫国,史其琦,王民.基于深度卷积神经网络的多特征融合的手势识别[J].液晶与显示,2019,34(4):417-422.
YUN W G, SHI Q Q, WANG M. Multi-feature fusion gesture recognition based on deep convolutional neural network [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(4): 417-422. (in Chinese)
- [12] 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)
- [13] 林通,陈新,唐晓,等.基于双流卷积神经网络和生成式对抗网络的行人重识别算法[J].信息技术与网络安全,2020,39(6):7-12.
LIN T, CHEN X, TANG X, *et al.* Dual stream ConvNet-Gan for person re-identification [J]. *Information Tech-*

- nology and Network Security, 2020, 39(6): 7-12. (in Chinese)
- [14] 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*. Long Beach, CA, USA: IEEE, 2019.
- [15] VARIOR R R, HALOI M, WANG G. Gated Siamese convolutional neural network architecture for human re-identification [C]// *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 791-808.
- [16] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets [C]// *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal, Quebec, Canada: ACM, 2014.
- [17] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]// *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV: IEEE, 2016: 770-778.
- [18] ARAL R A, KESKIN Ş R, KAYA M, *et al.* Classification of TrashNet dataset based on deep learning models [C]// *Proceedings of 2018 IEEE International Conference on Big Data (Big Data)*. Seattle, WA, USA: IEEE, 2018: 2058-2062.
- [19] PERERA P, ABAVISANI M, PATEL V M. Unsupervised multi-image-to-image translation using generative adversarial networks [C]// *Proceedings of the 24th International Conference on Pattern Recognition (ICPR)*. Beijing: IEEE, 2018: 140-146.
- [20] ISOLA P, ZHU J Y, ZHOU T H, *et al.* Image-to-image translation with conditional adversarial networks [C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, 2017: 5967-5976.
- [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 the 14th ECCV Workshops 2016*. Amsterdam, The Netherlands: Springer, 2016.
- [23] LIN L, WANG X L, YANG W, *et al.* Discriminatively trained and-or graph models for object shape detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(5): 959-972.
- [24] ZHENG L, YANG Y, HAUPTMANN A G. Person Re-identification: past, present and future [J]. *arXiv*:1610.02984, 2016.
- [25] LI D W, CHEN X T, ZHANG Z, *et al.* Learning deep context-aware features over body and latent parts for person re-identification [C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, 2017: 7398-7407.
- [26] BAI S, BAI X, TIAN Q. Scalable person re-identification on supervised smoothed manifold [C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, 2017: 3356-3365.
- [27] SUN Y F, ZHENG L, DENG W J, *et al.* SVDNet for pedestrian retrieval [C]// *Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE 2017: 3820-3828.
- [28] 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.
- [29] CHEN Y B, ZHU X T, GONG S G, *et al.* Person re-identification by deep learning multi-scale representations [C]// *Proceedings of 2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*. Venice: IEEE, 2017: 2590-2600.
- [30] CHANG X B, HOSPEDALES T M, XIANG T. Multi-level factorisation net for person re-identification [C]// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 2109-2118.
- [31] LI W, ZHU X T, GONG S G. Harmonious attention network for person re-identification [C]// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018:

- 2285-2294.
- [32] 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-ECCV 2018*. Munich, Germany: Springer, 2018.
- [33] WANG C, ZHANG Q, HUANG C, *et al.* Mancs: a multi-task attentional network with curriculum sampling for person re-identification [C]// *Proceedings of the 15th European Conference on Computer Vision-ECCV 2018*. Munich, Germany: Springer, 2018: 384-400.
- [34] SUN Y F, ZHENG L, YANG Y, *et al.* Beyond part models: person retrieval with refined part pooling [J]. *arXiv*:1711.09349, 2017.
- [35] SIAROHIN A, SANGINETO E, LATHUILIÈRE S, *et al.* Deformable GANs for pose-based human image generation [C]// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 3408-3416.

作者简介:



杨 琦(1994—),男,陕西榆林人,硕士研究生,2017 年于天津城建大学获得学士学位,主要从事图像处理、智能视频方面的研究。E-mail: qyang1028@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、智能视频方面的研究。E-mail:koalache@126.com