JOURNAL OF IMAGE AND GRAPHICS

主办：中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2020
12
VOL.25

ISSN1006-8961
CN11-3758/TB

数字浮雕建模 *P2647*

中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）

第25卷第12期（总第296期）

2020年12月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主　　编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮　　编 100190
电子信箱 jig@radi.ac.cn
电　　话 010-58887035
网　　址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订　　购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
（邮政信箱：北京399信箱　邮编：100048）
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

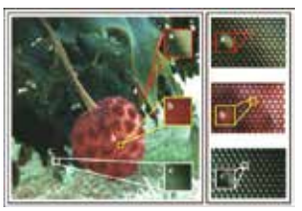
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048,P.R.China))
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

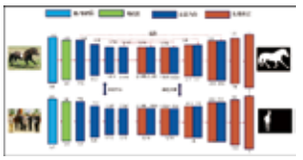
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元

2020年12月 第25卷 第12期（总第296期）
中国图象图形学报
Zhongguo Tuxiang Tuxing Xuebao

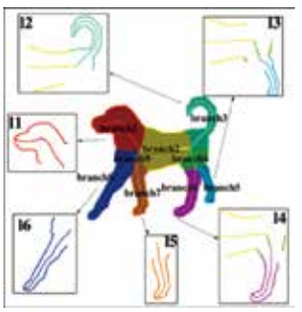
目次



光场显著性检测研究综述(第2465页)



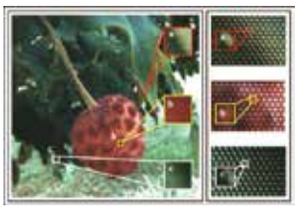
融合注意力机制与知识蒸馏的孪生网络压缩(第2563页)



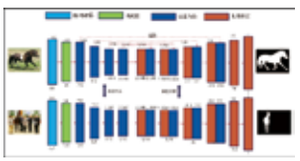
面向可解释性的物体拓扑结构骨架表征方法(第2587页)

Volume 25, Number 12
Published December 16, 2020

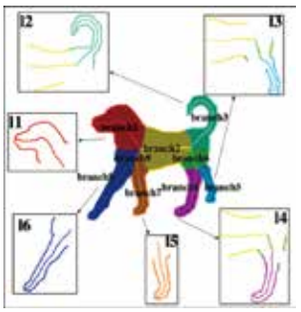
CONTENTS
JOURNAL OF IMAGE AND GRAPHICS



Review of saliency detection on light fields(P2465)



Combining attention mechanism and knowledge distillation for Siamese network compression (P2563)



Skeleton characterization of object topology toward explainability (P2587)

E-mail: jig@radi.ac.cn
Website: www.cjig.cn
Tel: 010-58887035

中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS

中图法分类号:TP301.6　　**文献标识码**:A　　**文章编号**:1006-8961(2020)12-2530-11

论文引用格式:Shi W D, Zhang Y Z, Liu S W, Zhu S D and Bao J N. 2020. Person re-identification based on deformation and occlusion mechanisms. Journal of Image and Graphics,25(12):2530-2540(史维东,张云洲,刘双伟,朱尚栋,暴吉宁. 2020. 针对形变与遮挡问题的行人再识别. 中国图象图形学报,25(12):2530-2540)[DOI:10.11834/jig.200016]

针对形变与遮挡问题的行人再识别

史维东,张云洲,刘双伟,朱尚栋,暴吉宁

东北大学信息科学与工程学院, 沈阳　110819

摘　要:**目的**　姿态变化和遮挡导致行人表现出明显差异,给行人再识别带来了巨大挑战。针对以上问题,本文提出一种融合形变与遮挡机制的行人再识别算法。**方法**　为了模拟行人姿态的变化,在基础网络输出的特征图上采用卷积的形式为特征图的每个位置学习两个偏移量,偏移量包括水平和垂直两个方向,后续的卷积操作通过考虑每个位置的偏移量提取形变的特征,从而提高网络应对行人姿态改变时的能力;为了解决遮挡问题,本文通过擦除空间注意力高响应对应的特征区域而仅保留低响应特征区域,模拟行人遮挡样本,进一步改善网络应对遮挡样本的能力。在测试阶段,将两种方法提取的特征与基础网络特征级联,保证特征描述子的鲁棒性。**结果**　本文方法在行人再识别领域3个公开大尺度数据集Market-1501、DukeMTMC-reID和CUHK03(包括detected和labeled)上进行评估,首位命中率Rank-1分别达到89.52%、81.96%、48.79%和50.29%,平均精度均值(mean average precision,mAP)分别达到73.98%、64.45%、43.77%和45.58%。**结论**　本文提出的融合形变与遮挡机制的行人再识别算法可以学习到鉴别能力更强的行人再识别模型,从而提取更加具有区分性的行人特征,尤其是针对复杂场景,在发生行人姿态改变及遮挡时仍能保持较高的识别准确率。

关键词:行人再识别;形变;遮挡;空间注意力机制;鲁棒性

Person re-identification based on deformation and occlusion mechanisms

Shi Weidong, Zhang Yunzhou, Liu Shuangwei, Zhu Shangdong, Bao Jining

College of Information Science and Engineering, Northeastern University, Shenyang　110819, China

Abstract: **Objective** Person re-identification (re-ID) identifies a target person from a collection of images and shows great value in person retrieval and tracking from a collection of images captured by network cameras. Due to its important applications in public security and surveillance, person re-ID has attracted the attention of academic and industrial practitioner sat home and abroad. Although most existing re-ID methods have achieved significant progress, person re-ID continues to face two challenges resulting from the change of view in different surveillance cameras. First, pedestrians have a wide range of pose variations. Second, some people in public spaces are often occluded by various obstructions, such as bicycles or other people. These problems result in significant appearance changes and may introduce some distracting information. As a result, the same pedestrian captured by different cameras may look drastically different from each other and may prevent re-ID. One simple, effective method for addressing this problem is to obtain additional pedestrian samples. Using abundant practical scene images can help generate more post-variant and occluded samples, thereby helping re-ID systems achieve

收稿日期:2020-01-13;**修回日期**:2020-04-27;**预印本日期**:2020-05-04

基金项目:国家自然科学基金项目(61973066,61471110,61733003);中央高校基本科研业务费专项资金资助(N172608005,N182608004)

Supported by: National Natural Science Foundation of China (61973066, 61471110, 61733003); Fundamental Research Funds for the Central Universities (N172608005, N182608004)

excellent robustness in complex situations. Some researchers have considered the representations of both the image and the key point-based pose as inputs to generate target poses and views via the generative adversarial networks (GAN) approach. However, GAN usually suffers from a convergence problem, and the generated target images usually have poor texture. In random erasing, a rectangle region is randomly selected from an image or feature map, and the original pixel value is discarded afterward to generate occluded examples. However, this approach only creates hard examples by spatially blocking the original image and (similar to the methods mentioned above) is very time consuming. To address these problems, we propose a person re-ID algorithm that generates hard deformation and occlusion samples. **Method** We use a deformable convolution module to simulate variations in pedestrian posture. The 2D offsets of regular grid sampling locations on the last feature map of the ResNet50 network are calculated by other branches that contain a multiple convolutional layer structure. These 2D offsets include the horizontal and vertical values X and Y. Afterward, these offsets are reapplied to the feature maps to produce new feature maps and deformable features via resampling. In this way, the network can change the posture of pedestrians in both horizontal and vertical directions and subsequently generate deformable features, thereby improving the ability of the network in dealing with deformed images. To address the occlusion problem, we generate spatial attention maps by using the spatial attention mechanism. We also apply other convolutional operations on the last feature map of the ResNet50 backbone to produce a spatial attention map that highlights the important spatial locations. Afterward, we mask out the most discriminative regions in the spatial attention map and retain only the low responses by using a fixed threshold value. The processed spatial attention map is then multiplied by the original features to produce the occluded features. In this way, we simulate the occluded pedestrian samples and further improve the ability of the network to adapt to other occluded samples. In the testing, we cascade two features with the original features as our final descriptors. We implement and train our network by using Pytorch and an NVIDIA TITAN GPU device, respectively. We set the batch size to 32 and rescaled all images to a fixed size of 256 × 128 pixels during the training and testing procedures. We also adopt a stochastic gradient descent (SGD) with a momentum of 0.9 and weight decay coefficient of 0.000 5 to update our network parameters. The initial learning rate is set to 0.04, which is further divided by 10 after 40 epochs (the training process has 60 epochs). We fix the reduction ratio and erasing threshold to 16 and 0.7 in all datasets, respectively. We adopt random flip as our data augmentation technique, and we use ResNet50 as our backbone model that contains parameters that are pretrained on the ImageNet dataset. This model is also trained end-to-end. We adopt cumulative match characteristic (CMC) and mean average precision (mAP) to compare the re-ID performance of the proposed method with that of existing methods. **Result** The performance of our proposed method is evaluated on public large-scale datasets Market-1501, DukeMTMC-reID, and CUHK03. We use a uniform random seed to ensure the repeatability of the equity comparison and the results. In the Market-1501, DukeMTMC-reID, and CUHK03 (detected and labeled) datasets, the proposed method has obtained Rank-1 (represents the proportion of the queried people) values of 89.52%, 81.96%, 48.79%, and 50.29%, respectively, while its mAP values in these datasets reach 73.98%, 64.45%, 43.77%, and 45.57%, respectively. In the detected and labeled CUHK03 datasets, the proposed method shows 9.43%/8.74% and 8.72%/8.0% improvements in its Rank-1 and mAP values, respectively. These experimental results validate the competitive performance of this method for small and large datasets. **Conclusion** The proposed person re-ID system based on the deformation and occlusion mechanisms can construct a highly recognizable model for extracting robust pedestrian features. This system maintains high recognition accuracy in complex application scenarios where occlusion and wide variations in pedestrian posture are observed. The proposed method can also effectively mitigate model overfitting in small-scale datasets (e.g., CUHK03 dataset), thereby improving its recognition rate.

Key words: person re-identification; deformation; occlusion; spatial attention mechanism; robustness

0 引 言

随着国家对智慧城市建设的大力投入以及监控设备的迅速发展，大量监控摄像头安装在各大商场、小区以及易发生安全事故的人流密集场所，不仅给人民生活带来了保障，也为警方侦察带来了便捷。公安机关可以根据摄像头拍摄的内容进行嫌疑犯的

锁定和跟踪,极大加快了案件的侦破速度,提高了办案效率。但是,实时监控网络多采用人工观察的方法,海量的监控数据和长时间观察对工作人员来说是一种沉重负担,使其无法长时间集中注意力。因此,人工监控的方法不适用于海量监控数据的管理和分析。行人再识别技术的出现,有效解决了这个问题。行人再识别也称为行人重识别,其任务是给定一个监控设备中的行人图像,检索其他设备中的该行人图像。由于监控设备拍摄角度不同,导致拍摄的行人往往呈现多种外观变化,同一行人在不同摄像头中的外观存在较大差异。其中,最为显著的是行人姿态严重形变且可能伴随部分遮挡,给行人再识别带来了极大挑战。

以卷积神经网络为代表的深度学习方法(Liu 等,2019;朱福庆 等,2018;陈亮雨和李卫疆,2019)在行人再识别任务中取得了良好成绩。相比于传统的手工特征(齐美彬 等,2018),卷积神经网络在复杂场景下可以提取更具有鉴别性的行人特征,从而得到更好的识别性能。但当物体发生形变时,传统的卷积神经网络无法根据物体的形变做出卷积形式的改变。此外,当目标出现遮挡导致关键信息丢失时,模型的性能会急剧下降。针对上述两方面问题,以往的解决方式是通过生成对抗网络(generative adversarial networks,GAN)(Goodfellow 等,2016)生成同一行人的多种姿态(Liu 等,2018)和在原始数据库中对每一幅图像进行不同程度的遮挡(Huang 等,2018)。这类方法都是基于数据扩充或增强的方式来增加训练样本的多样性,从而使得模型可以在更加复杂的场景下进行训练。但该方法得到的困难样本无论是形变还是遮挡程度,都是众多复杂情况下的少数情况。因此,模型适应未知形变或遮挡的能力有限,仍会存在性能下降的问题。此外,采用 GAN 网络进行迁移学习训练难度更高,存在收敛困难问题,且不是端到端的方法。

为此,本文提出一种更加有效的端到端融合形变与遮挡机制的行人再识别算法,可以同时解决形变和遮挡的问题。本文的主要贡献如下:

1)在现有卷积神经网络的行人再识别框架下,引入了可变形卷积层,根据当前识别的内容动态调整卷积核的作用位置来模拟行人的姿态改变,从而增加训练样本的形变多样性,提高模型对行人姿态改变的处理能力。

2)针对行人再识别遮挡问题,提出通过擦除模型学习到的高响应区域,仅保留鉴别性较差的低响应区域来生成遮挡样本,从而增加样本的复杂性,通过强迫模型从低响应区域学习区分性特征来提高模型解决遮挡问题的能力。

3)本文提出的解决形变和遮挡的行人再识别算法在公开数据集 Market-1501、DukeMTMC-reID 和 CUHK03 上进行实验评估,取得了较高的准确率。

4)本文方法扩展性较强,可以扩展到其他一些基于卷积神经网络结构的识别任务中。

1 相关工作

行人再识别工作具有极大的应用价值,受到越来越多的关注,但受限于行人姿态改变以及遮挡的影响,行人再识别工作仍然是一项艰巨任务。现有的大部分工作主要从两个方面进行研究,一是提取鲁棒性特征,二是设计合理的距离度量。传统算法(Liao 等,2015;Matsukawa 等,2016)往往设计一种鲁棒的手工特征来应对行人姿态的改变以及遮挡问题。此外,一些研究工作同样尝试采用鲁棒的距离度量。Köstinger 等人(2012)提出一种简单直接的距离度量学习算法(keep it simple and straight forward metric,KISSME),用于大尺度数据的距离度量学习。Pedagad 等人(2013)提出局部 Fisher 判别分析度量学习算法(local Fisher discriminative analyze,LFDA),不对所有的样本点赋予相同权重,而是考虑局部样本点,应用局部 Fisher 判别分析方法为降维的特征提供有识别能力的空间。虽然传统算法实现了有效的行人特征的表达,但传统算法提取的特征无法针对行人遮挡以及形变提取更加有效的特征,因此传统算法的鲁棒性较差,在复杂环境下的识别性能往往很低。

基于深度学习的方法(Sun 等,2017;Zhao 等,2017;Si 等,2018;Zheng 等,2018)逐渐成为行人再识别的主要方式。相比于传统方法,通过深度学习提取到的特征更加具有判别性,可以得到更好的识别准确率。但是,深度行人再识别模型的性能很大程度上受限于可训练数据的数量,如果数据量较大,那么模型可以在更复杂的场景下学习,更有利于后期模型的识别工作。但大规模的数据集需要人工裁剪和标记,这消耗了更多的时间和经济成本。为此,

Zhong 等人(2017a)提出随机擦除算法,在图像上随机选择一个矩形区域,并将其像素值设置为随机值来产生遮挡式的样本,这在一定程度上提高了模型应对遮挡样本的能力。相比于在图像上擦除样本的方式,Dai 等人(2019)进一步提出了在特征层面上进行擦除的方法,目的是不希望网络过于关注那些显而易见的全局特征。但上述方法都是基于随机遮挡的方式来生成遮挡样本。本文方法是擦除高响应区域,相比之前的方法更具有针对性,可以生成更加困难的行人遮挡样本。目前,注意力机制的方法广泛应用于场景分割(Long 等,2015)、图像分类(Krizhevsky 等,2012)和目标检测(Girshick 等,2014)等计算机视觉领域,基于注意力机制的方法可以让网络更关注区分性特征。受注意力机制的启发,本文通过遮挡空间注意力所关注的显著性特征来生成遮挡困难样本。相比于随机遮挡的形式,本文方法更加具有针对性,遮挡造成的训练难度更大,从而模型可以捕获更加鲁棒性的特征。

此外,对于行人再识别中的形变问题,传统算法提取形状不变特征(scale-invariant feature transform,SIFT)(Lowe 等,1999),但手工设计的特征往往鲁棒性较差,在遇到未知复杂的形变时,模型的性能会急剧下降。近年来,采用姿态迁移的方法生成固定形变样本成为解决行人形变的主要方法。Ma 等人(2017)提出了一种基于姿态指导的行人图像生成方法,采用两阶段的生成策略,通过预估行人大体轮廓后再合成精细的残差图来增强合成图像的细节纹理,虽然完成了姿态的迁移,但生成的图像分辨率较差,不利于后续的模型训练。相比于 Ma 等人(2017)的方法,Qian 等人(2018)提出了一种姿态归一化的生成对抗方法,可以生成较清晰的固定姿态的行人图像,在一定程度上缓解了姿态多样性造成的模型学习视角不敏感带来的特征差的问题,但为每个行人仅生成 8 种简单姿态的图像,仍然无法提高模型应对行人复杂姿态情况的识别效果。最近,Zhu 等人(2019)提出了一种基于注意力的姿态迁移的方法,将原始图像、原始图像姿态关键点以及目标姿态关键点输入至姿态注意力迁移网络,从而将目标姿态更加有效地替换原始图像中的姿态。上述工作取得了一定的效果,但仍然存在 3 个缺点:1)基于姿态迁移的方法需要预先生成目标行人关键点来完成姿态的迁移,因此目标行人姿态的丰富性受限于关键点的丰富性,且较为复杂的关键点很难完成迁移;2)基于姿态迁移的工作(Ma 等,2017;Qian 等,2018;Siarohin 等,2018;Zhu 等,2019)大多采用生成对抗网络实现,图像分辨率较差且存在失真问题,不利于模型的训练;3)采用 GAN 的方法往往存在收敛问题且不是端到端的训练形式,无法直接应用于行人再识别模型。相比于上述方法,本文采用可变形卷积的方式来模拟行人的形变,通过在线学习的方式产生多样且复杂的形变特征。由于本文的形变方法是一种端到端的形式,不需要生成形变的行人图像,因此没有图像失真以及关键点迁移问题。为了进一步提高行人识别的准确率,研究者提出了重排序(re-ranking)(Zhong 等,2017b)和多帧排序(multiple query)的方法,从而进一步提高行人识别的准确性。值得注意的是,上述策略均可以与本文方法联合使用。

2 形变和遮挡生成算法

本文提出一种融合形变与遮挡机制的行人再识别算法。通过可变形卷积生成行人的形变样本,通过擦除空间注意力高响应对应的特征区域生成行人遮挡样本。整体的网络结构如图 1 所示。

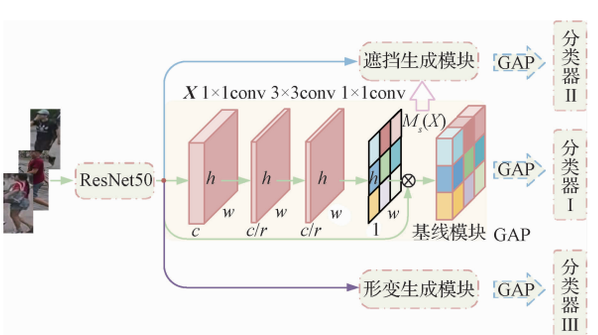


图 1 本文网络结构图

Fig. 1 The proposed network structure

本文的基线模块是在主干网络 ResNet50(He 等,2016)后应用空间注意力来提取行人的全局特征。在此基础上,提出遮挡生成模块与形变生成模块来分别提取行人的遮挡与形变特征。在训练阶段,每种类型的特征经过全局平均池化(global average pooling,GAP)后在 3 个独立的分类器监督下分别进行训练;在推理阶段,3 种类型的特征级联作为行人特征描述子。

2.1 基线方法

本文采用的基线方法为:在 ResNet50 的 Conv5_3 末端嵌入空间注意力机制,Conv5_3 输出的特征图经过空间注意力修正后作为基线的特征描述子。具体地,空间注意力模块由两个 1×1 卷积和一个 3×3 卷积组成。在公式表达上,定义 $\boldsymbol{X} \in \mathbf{R}^{c\times h\times w}$ 为 Conv5_3 输出特征图,其中,c, h, w 分别代表特征图的通道数量、高度和宽度。首先利用 1×1 卷积降低特征图通道数量为 $\mathbf{R}^{c/r\times h\times w}$,其中 r 为降维比例,然后经过 3×3 卷积扩大感受野,最后再经过 1×1 卷积将维度降为 $\mathbf{R}^{1\times h\times w}$。空间注意力模型计算为

$$M_s(\boldsymbol{X}) = BN(g_2^{1\times1}(g_1^{3\times3}(g_0^{1\times1}(\boldsymbol{X}))))$$
$$\boldsymbol{F}_1 = \boldsymbol{X}\odot M_s(\boldsymbol{X}) \tag{1}$$

式中,g 表示卷积操作,其右上角数字代表卷积核大小,BN(batch normalization)为数据归一化操作,$M_s(\boldsymbol{X}) \in \mathbf{R}^{h\times w}$ 为空间注意力图,$\odot$ 表示哈达玛积(Hadamard product),$\boldsymbol{F}_1$ 表示基线输出特征图。

2.2 遮挡样本生成方法

在行人再识别算法中,遮挡会造成模型识别性能下降,因为可区分性特征一旦遮挡,便丢失了具有判别性的视觉线索。为了提取更加鲁棒的行人特征,受注意力机制模型的启发,本文通过遮挡空间注意力高响应对应的特征图,仅保留空间注意力低响应对应的特征图,在卷积特征空间生成遮挡样本,并输入给分类器学习。如图 2 所示,基线模块中的空间注意力模型在特征图空间方向上学习的位置信息非常关键,更值得模型关注。当得到空间注意力响应图后,将其中高响应区域的响应值置为 0,其他区域数值保持不变,得到掩模 $\overline{M(\boldsymbol{X})}$,再将 $\overline{M(\boldsymbol{X})}$ 与原始特征图作哈达玛乘积,得到本文的遮挡样本。$\overline{M(\boldsymbol{X})}$ 可表示为

$$\overline{M(\boldsymbol{X})} = \begin{cases} 0 & M_s(\boldsymbol{X})_{i,j} > t \\ M_s(\boldsymbol{X})_{i,j} & \text{其他} \end{cases}$$
$$\text{s.t.} \quad i \in w, j \in h \tag{2}$$

式中,$M_s(\boldsymbol{X})_{i,j}$ 表示基线模块学习到的空间注意力图 $M_s(\boldsymbol{X})$ 中第 i,j 位置的数值。当数值大于阈值 t 时,赋值为 0,否则保留原值。最后 $\overline{M(\boldsymbol{X})}$ 与原始特征 $\boldsymbol{X}$ 相乘得到特征图 $\boldsymbol{F}_2$。上述操作将特征图 $\boldsymbol{F}_2$ 中的高响应区域置为 0,因此 $\boldsymbol{F}_2$ 可以理解为行人判别性区域遮挡更具困难性的样本。

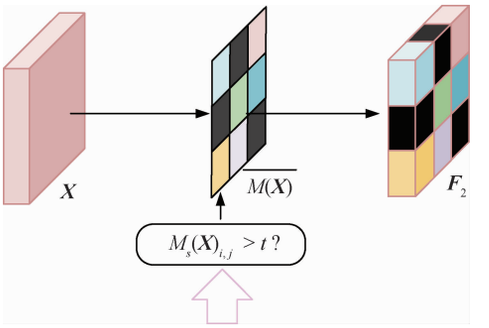


图 2 遮挡生成模块网络结构图

Fig. 2 Network structure of occlusion generation module

2.3 形变样本生成方法

行人姿态变化同样影响着模型的识别性能,为了提高模型对行人姿态变化识别的鲁棒性,本文进一步提出在线生成行人形变特征。具体地,采用一种可变形卷积模块(Dai 等,2017),如图 3 所示。

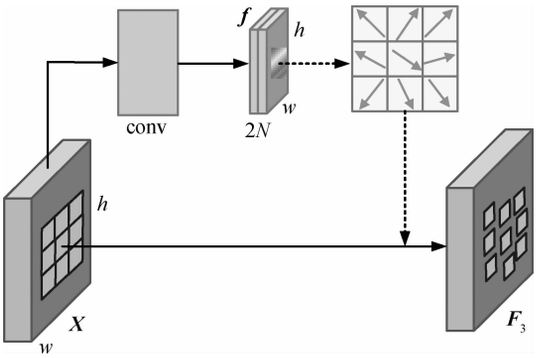


图 3 形变生成模块网络结构图

Fig. 3 Network structure of deformation generation module

该模块由两个分支构成,上面的分支通过两个卷积层生成一个 2D 的采样大小网格,将学习到的偏移量作用于原始的特征图,然后再进行卷积操作。此时卷积操作采样的点并不是标准采样形式,而是需要结合学习到的偏移量进行选择性采样。由于卷积网格是以一种考虑偏移量后的非规则形状,因此卷积后生成的特征图 $\boldsymbol{F}_3$ 称为形变特征。在公式描述上,以 3×3 的卷积核大小为例,卷积核采样位置定义为

$$\boldsymbol{L} = \{(-1,-1),(-1,0),(-1,1),\cdots,$$
$$(1,-1),(1,0),(1,1)\} \tag{3}$$

式中,$\boldsymbol{L}$ 中的数字代表卷积核中 9 个格点在一个卷积核中的位置,$N = |\boldsymbol{L}|$ 为格点数。由于偏移量包括水平和垂直两个方向,为此,通过额外的卷积操作(convolution,conv)生成维度为 $2N$、空间大小不变的

特征偏移量 $\boldsymbol{f}$，此时，特征图 $\boldsymbol{X}$ 中每个特征值的位置为原始坐标值加上偏移值。对于常规卷积特征图 $\boldsymbol{Y}$，其空间位置 (i_o,j_o) 的特征值计算为

$$Y(i_o,j_o) = \sum_{(i,j\in \boldsymbol{L})} \omega(i,j) \otimes X((i_o,j_o) + (i,j)) \quad \text{s.t.} \quad i_o \in w, j_o \in h \tag{4}$$

式中，$\omega(i,j)$ 为卷积核 (i,j) 位置的权重值，$\otimes$ 表示卷积操作。然而，在可变形卷积中，卷积操作需要考虑原始卷积核中 (i,j) 位置的偏移量 $(\Delta i,\Delta j)$，因此将式(4)进一步改写为

$$\boldsymbol{F}_3(i_o,j_o) = \sum_{(i,j\in \boldsymbol{L})} \omega(i,j) \otimes \boldsymbol{X}((i_o,j_o) + (i,j) + (\Delta i,\Delta j)) \tag{5}$$

式中，$\boldsymbol{F}_3$ 为形变特征，$(i,j) + (\Delta i,\Delta j)$ 表示此时的卷积操作是在带有偏移量的特征图上进行的。需要注意的是，$(\Delta i,\Delta j)$ 通常是小数，因此需要采用线性插值方式确定非整数位置的采样，具体为

$$X(I,J) = \sum_{q}^{w} \sum_{p}^{h} G((q,p),(I,J)) X(q,p) \tag{6}$$

式中，非整数坐标 $(I,J) = (i_o,j_o) + (i,j) + (\Delta i,\Delta j)$，$(q,p)$ 为特征图 $\boldsymbol{X}$ 的整数位置坐标，$G(\cdot,\cdot)$ 是 2 维线性插值核，式(6)可以进一步分离为

$$G((q,p),(I,J)) = \max(0,1-|q-I|) \cdot \max(0,1-|p-J|) \tag{7}$$

通过上面的方法，可变形卷积模块可以灵活地改变卷积核的采样位置，实现从行人的不同区域提取特征并融合，从而模拟行人的姿态改变，生成形变特征，进一步提高模型应对行人姿态改变的能力。

3 实 验

3.1 实验数据集

在 Market-1501、DukeMTMC-reID 和 CUHK03 数据集上进行实验。通过与基线方法以及现有的行人再识别方法对比，验证本文方法的有效性。

Market-1501(Zheng 等，2015)数据集由 6 个摄像头拍摄的 32 668 幅图像组成，包含 1 501 个行人。实验将 751 个行人共 12 936 幅图像作为训练集，750 个行人共 19 732 幅图像作为测试集。

DukeMTMC-reID(Ristani 等，2016)数据集由 8 个摄像头拍摄的 36 411 幅图像组成，包含 1 404 个行人。实验将 702 个行人共 16 522 幅图像作为训练集，702 个行人共 17 661 幅图像作为测试集。

CUHK03(Li 等，2014)数据集由两个摄像头拍摄的 14 097 幅图像组成，包含 1 467 个行人。实验将 767 个行人样本作为训练集，700 个行人样本作为测试集。

3.2 实验设置及评价标准

实验基于深度学习框架 Pytorch 完成，采用 GPU 为 TITAN XP 的深度学习服务器进行训练。在训练过程中，采用随机梯度下降算法优化模型，初始学习率、权重衰减系数以及优化器动量数值分别设置为 0.04，0.000 5 和 0.9，batch size 设置为 32，整个网络的训练迭代(epoch)次数设置为 60 次，学习率在 40 次迭代后衰减为 0.004。此外，所有输入网络中的图像尺寸统一调整为 256 × 128 像素，降维比例 r 设置为 16，遮挡样本阈值 t 设置为 0.7。数据增强方式仅采用随机翻转，基础模型(ResNet50)初始化参数采用 ImageNet 上的预训练参数。

本文采用累计匹配特性曲线(cumulative match characteristic，CMC)和平均精度均值(mean average precision，mAP)来评估算法的性能。CMC 曲线中 Rank-k 表示前 k 个搜索结果中找到待查询行人的比率。mAP 指所有查询样本正确率—召回率曲线下面积的平均值，反映了行人再识别方法总体性能。

3.3 实验结果

为了说明本文方法的有效性，与现有的行人再识别算法进行对比分析，包括基于遮挡样本生成的方法 IDE + RE(ID discriminative embedding + random erasing)(Zhong 等，2017a)，AOSReID(adversarially occluded samples for person re-identification)(Huang 等，2018)，基于奇异值分解的显著性特征学习方法 SVDNet(singular vector decomposition network)(Sun 等，2017)、基于困难样本挖掘的方法 Triplet Loss(Hermans 等，2017)、基于注意力机制的方法 DLPAR(deeply learned part aligned represent)(Zhao 等，2017)以及基于姿态迁移的方法 Pose-Transfer(Liu 等，2018)。表 1—表 3 分别给出在 3 个数据集上的实验对比结果。

与同样采用遮挡样本生成方法的 AOSReID 相比，本文方法在 Market-1501、DukeMTMC-reID 和 CUHK03(detected)数据集上的性能仍有较大提升，如表 1—表 3 所示，Rank-1 指标分别提升了 3.03%、

表 1　不同方法在 Market-1501 数据集上的评估结果

Table 1　Evaluation results of different methods on Market-1501 dataset

/%

方法	Rank-1	Rank-5	Rank-10	mAP
DLPAR	81.00	92.00	94.70	63.40
SVDNet	82.30	92.30	95.20	62.10
IDE + RE	85.20	–	–	68.30
Triplet Loss	84.90	94.20	–	69.10
AOSReID	86.49	–	–	70.43
Pose-Transfer	87.65	–	–	68.92
本文	**89.52**	**96.11**	**97.65**	**73.98**

注:加粗字体为每列最优结果,“–”表示未提供数据。

表 2　不同方法在 DukeMTMC-reID 数据集上的评估结果

Table 2　Evaluation results of different methods on DukeMTMC-reID dataset

/%

方法	Rank-1	Rank-5	Rank-10	mAP
Pose-Transfer	68.64	–	–	48.06
SVDNet	76.70	86.40	89.90	56.80
IDE + RE	74.20	–	–	56.20
AOSReID	79.17	–	–	62.10
本文	**81.96**	**90.17**	**93.09**	**64.45**

注:加粗字体为每列最优结果,“–”表示未提供数据。

表 3　不同方法在 CUHK03 数据集上的评估结果

Table 3　Evaluation results of different methods on CUHK03 dataset

/%

方法	detected		labeled	
	Rank-1	mAP	Rank-1	mAP
Pose-Transfer	41.60	38.70	45.10	42.00
SVDNet	41.50	37.30	40.90	37.80
IDE + RE	38.50	34.80	41.50	36.80
AOSReID	47.14	43.33	–	–
本文	**48.79**	**43.77**	**50.29**	**45.58**

注:加粗字体为每列最优结果,“–”表示未提供数据。

2.79% 和 1.65%,mAP 指标分别提升了 3.55%、2.35% 和 0.44%。原因在于,虽然 AOSReID 方法通过滑窗遮挡的方式生成了遮挡性的样本,但尺寸固定的滑窗方法往往只能遮挡住一部分显著性区域,并不适合环境复杂的情况。而本文考虑到显著性区域往往是不规则区域,因此通过设置阈值的方式获取更具遮挡性的样本,从而实现了更好的实验结果。

与 Market-1501 数据集相比,DukeMTMC-reID 数据集的复杂性更高。在该数据集上,与其他方法相比,本文方法仍能取得性能上的提升,如表 2 所示。与 Pose-Transfer 方法相比,本文方法具有较大优势,原因在于 Pose-Transfer 方法采用 GAN 网络生成多种姿态的行人图像,然后与原始图像同时送进网络训练。虽然在一定程度上解决了行人姿态变化对模型性能的影响,但其姿态多样性与实际场景下的行人姿态仍具有较大差别,且 GAN 网络训练难度大、收敛慢,生成的多种行人姿态图像需重新送入网络中训练行人识别模型,不是端到端的学习过程。而本文采用端到端的在线学习的方式,生成更为多样且复杂的行人形变特征,实现了更加优异的结果。

受限于 CUHK03 数据量不足,模型在该数据集上更容易出现过拟合问题,性能表现往往很差。本文方法在卷积特征空间在线生成困难特征图,丰富了样本的多样性,提高了模型的泛化能力。如表 3 所示,与 Pose-Transfer 和 AOSReID 单独使用姿态迁移或遮挡样本不同,本文方法将形变和遮挡结合使用,模型的性能实现了进一步提升。

3.4　消融实验

为了进一步验证本文方法的有效性,本文从定量和定性分析两方面进行验证。首先,测试基线方法在 3 种数据集上的实验结果。在相同的参数设置下,本文测试分别加入形变、遮挡以及二者结合方法的实验,具体对比结果如表4—表6 所示。此外,通过可视化的形式直观展示了遮挡方法和形变方法的有效性。

表 4　在 Market-1501 数据集上的消融实验

Table 4　Ablation study on Market-1501 dataset

/%

方法	Rank-1	Rank-5	Rank-10	mAP
基线	88.60	95.84	97.51	70.54
形变	89.99	96.19	97.65	73.96
遮挡	**90.38**	**96.62**	**98.10**	73.74
形变 + 遮挡	89.52	96.11	97.65	**73.98**

注:加粗字体为每列最优结果。

3.4.1　定量分析

遮挡和形变样本生成是为了增强模型对视觉线

表 5　在 DukeMTMC-reID 数据集上的消融实验

Table 5　Ablation study on DukeMTMC-reID dataset

/%

方法	Rank-1	Rank-5	Rank-10	mAP
基线	79.08	87.79	91.29	60.53
形变	80.92	89.95	93.09	63.61
遮挡	80.39	89.72	92.77	63.41
形变 + 遮挡	**81.96**	**90.17**	**93.18**	**64.45**

注：加粗字体为每列最优结果。

表 6　在 CUHK03 数据集上的消融实验

Table 6　Ablation study on CUHK03 dataset

/%

方法	detected		labeled	
	Rank-1	mAP	Rank-1	mAP
基线	39.36	35.03	41.57	37.58
形变	46.43	41.96	46.07	43.33
遮挡	47.00	41.65	45.43	42.20
形变 + 遮挡	**48.79**	**43.77**	**50.29**	**45.58**

注：加粗字体为每列最优结果。

索具有不变性的学习能力，从而在复杂环境和行人表观变化时仍能提取有效的行人特征。图 4 是基线方法与本文方法的 CMC 对比。从表 4—表 6 和图 4 可以看出，无论单独或结合使用，本文方法的识别效果都优于基线方法。

在小规模数据集 CUHK03 上，由于该数据集样本数较少、多样性差，因此模型训练时容易出现过拟合问题。然而，本文通过形变与遮挡样本生成的方式增加样本的多样性、复杂性以及样本的学习难度，从而使得模型更加关注鲁棒性的特征，因此可以在 CUHK03 数据集上实现较大的性能提高。具体地，在 CUHK03-labeled 数据集上，本文方法相比于基线方法，Rank-1 提高 8.72%，mAP 提高 8.0%；在更加困难的检测器裁剪下的 CUHK03-detected 数据集上，本文方法相比于基线方法，Rank-1 提高 9.43%，mAP 提高 8.74%。在 DukeMTMC-reID 数据集上，本文方法仍然有较大的提升，说明即便在复杂的数据集下，该方法仍是一种有效的手段。但在 Market-1501 数据集上，两种方法单独使用时的模型性能略高于结合的方式。因此推测，增加形变与遮挡困难样本的方法在 Market-1501 数据集上引入了过多的干扰，导致模型学习了不相干的信息，使识别准确性有所下降，但整体的识别结果仍然高于基

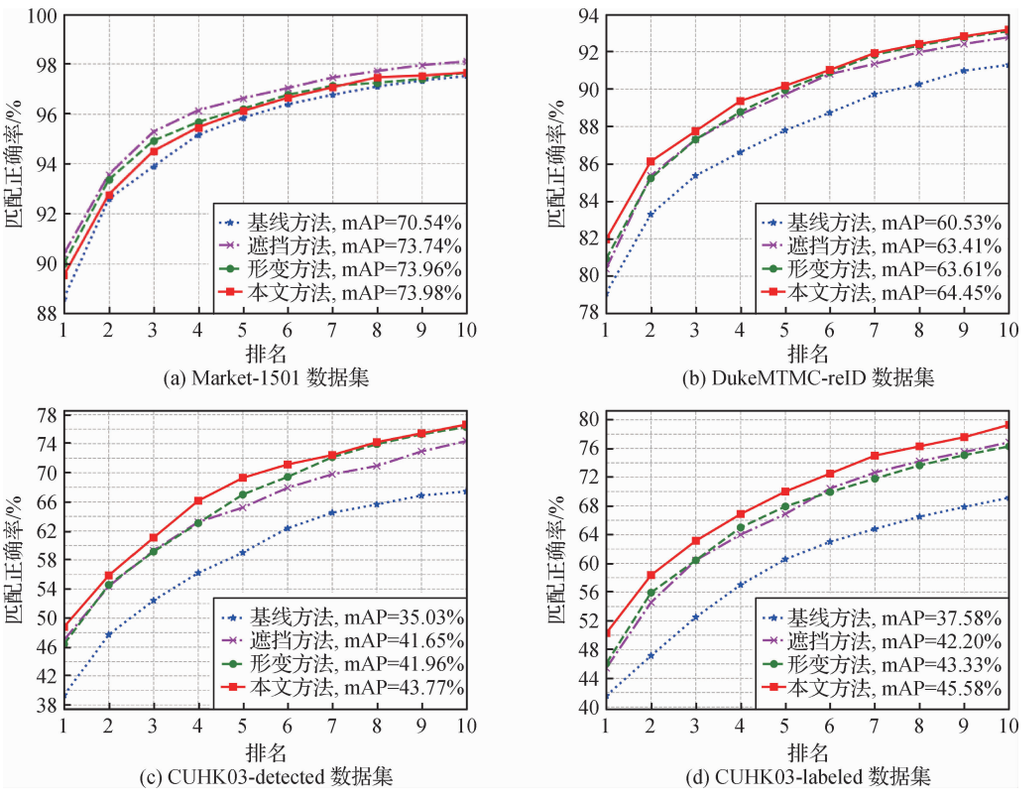


图 4　基线方法与本文方法的 CMC 对比

Fig. 4　CMC comparison of the baseline and the proposed method ((a) Market-1501 dataset; (b) DukeMTMC-reID dataset; (c) CUHK03-detected dataset; (d) CUHK03-labeled dataset)

线的方法。上述消融实验证明了本文每一个设计的有效性。

3.4.2　定性分析

为了从视觉上直观地展现本文方法的有效性，图5给出了基线方法和本文方法的行人查询排序结果。图5为模型检索的前10名排序结果。其中第1列是待查询图像，其他是候选图像。绿色框代表待查询图像与候选图像为同一行人，红色框代表为不同行人。

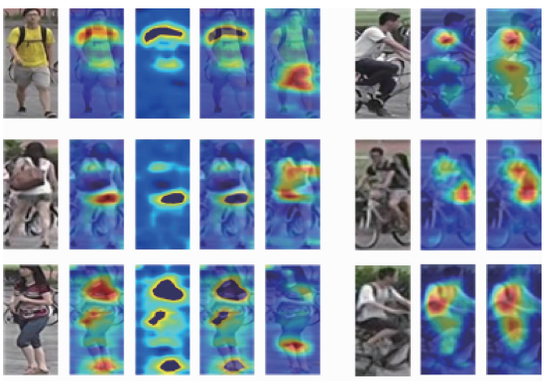
(a) 基线方法　(b) 本文方法

图5　本文方法和基线方法的排序图

Fig. 5　Ranking list of our methods and baseline ((a) baseline; (b) ours)

从图5可以看出，对于完整行人，基线方法和本文方法都可以准确识别。当行人出现遮挡(第2行)时，基线方法前10名检索结果中出现5个错误识别，本文方法出现1个。当行人姿态改变(第3行)时，基线方法出现6个错误识别，本文方法出现2个。为了更加突出本文方法的优越性，将同时发生遮挡及形变的行人(第4行)作为待查询对象，基线方法识别性能进一步下降，排序图中出现7个错误，而本文方法实现了全部正确的查询结果。这从定性的角度证明了本文方法的有效性。此外，对于错误识别出现的次序，相比于基线方法，本文的错误查询结果在排序图中更靠后，说明本文的总体性能更加优越。

同时，对遮挡和形变两种模块分别进行了可视化分析，如图6所示。图6(a)第2列为基线方法对遮挡行人的可视化，通过擦除空间注意力高响应区域(图6(a)第3列)获得低响应区域特征(图6(a)第4列)，送至分类器Ⅱ中学习，从而提高模型对低响应行人区域(图6(a)第5列)的关注能力。实验表明，本文采用遮挡模块可以学习与基线方法互补的特征。尤其图6(a)第3行的可视化结果可以证明：虽然网络关注到较多的行人区域，但经过遮挡模块处理后，仍然可以学习到位于行人身上的互补区域。图6(b)第2列是基线方法对形变行人的可视化，与此相比，采用形变模块处理后模型可以关注到更多行人区域(图6(b)第3列)。具体地，对于行人上半身及腿部的形变，基线方法更容易关注形变较小的上半身区域，而对腿部形变较大的部位关注能力较差；即使同样可以关注腿部区域，但关注区域较小(图6(b)第2行)。对两种模块的可视化分析，再一次证明了本文方法的有效性。



(a) 遮挡　(b) 形变

图6　遮挡及形变模块热度图

Fig. 6　The heatmaps of occluded model and deformable model ((a) occluded model; (b) deformable model)

4　结　论

遮挡和行人姿态变化极大地影响着行人再识别模型的识别准确性。为此，本文提出了一种融合形变与遮挡机制的行人再识别算法，通过可变形卷积生成形变的样本，通过遮挡空间高响应区域生成遮挡样本，从而提高模型在行人姿态改变以及遮挡场

景下的识别能力。在 3 个数据集上的实验结果表明,本文提出的结合形变与遮挡机制的行人再识别算法增加了样本多样性,使模型可以在更加复杂的情况下训练,从而提高了模型应对行人形变与遮挡的鲁棒性。特别地,在规模较小的 CUHK03 数据集上,证明了本文提出的形变与遮挡机制生成的样本可以在一定程度上缓解数据集因数据量少产生的过拟合问题。此外,由于本文采用端到端的形式产生形变与遮挡的困难样本训练网络,与以往工作将离线生成的形变或遮挡图像重新送进网络中相比,本文方法更加简洁有效。目前,遮挡仍然是行人再识别研究的关键问题,本文设计的遮挡模块仍然需要针对特定数据集进行遮挡程度的参数设定,而参数的选择对模型训练稳定性影响较大,因此未来将在学习遮挡方向上进一步开展研究。

参考文献(References)

Chen L Y and Li W J. 2019. Multishape part network architecture for person re-identification. Journal of Image and Graphics, 24(11): 1932-1941 (陈亮雨, 李卫疆. 2019. 多形状局部区域神经网络结构的行人再识别. 中国图象图形学报, 24(11): 1932-1941) [DOI: 10.11834/jig.190042]

Dai J F, Qi H Z, Xiong Y W, Li Y, Zhang G D, Hu H and Wei Y C. 2017. Deformable convolutional networks//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 764-773 [DOI: 10.1109/ICCV.2017.89]

Dai Z Z, Chen M Q, Gu X D, Zhu S Y and Tan P. 2019. Batch dropblock network for person re-identification and beyond [EB/OL]. [2019-09-03]. https://arxiv.org/pdf/1811.07130v2.pdf

Girshick R, Donahue J, Darrell T and Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 580-587 [DOI: 10.1109/cvpr.2014.81]

Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2016. Generative adversarial nets [EB/OL]. [2019-12-13]. https://arxiv.org/pdf/1406.2661.pdf

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/cvpr.2016.90]

Hermans A, Beyer L and Leibe B. 2017. In defense of the triplet loss for person re-identification [EB/OL]. [2019-12-13]. https://arxiv.org/pdf/1703.07737.pdf

Huang H J, Li D W, Zhang Z, Chen X T and Huang K Q. 2018. Adversarially occluded samples for person re-identification//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 5098-5107 [DOI: 10.1109/cvpr.2018.00535]

Köstinger M, Hirzer M, Wohlhart P, Roth P M and Bischof H. 2012. Large scale metric learning from equivalence constraints//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE: 2288-2295 [DOI: 10.1109/cvpr.2012.6247939]

Krizhevsky A, Sutskever I and Hinton G E. 2012. Imagenet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: ACM: 1097-1105

Li W, Zhao R, Xiao T and Wang X G. 2014. DeepReID: deep filter pairing neural network for person re-identification//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 152-159 [DOI: 10.1109/cvpr.2014.27]

Liao S C, Hu Y, Zhu X Y and Li S Z. 2015. Person re-identification by local maximal occurrence representation and metric learning//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 2197-2206 [DOI: 10.1109/cvpr.2015.7298832]

Liu J X, Ni B B, Yan Y C, Zhou P, Cheng S and Hu J G. 2018. Pose transferrable person re-identification//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 4099-4108 [DOI: 10.1109/CVPR.2018.00431]

Liu S W, Zhang Y Z, Qi L, Coleman S, Kerr D and Zhu S D. 2019. Adversarially erased learning for person re-identification by fully convolutional networks//Proceedings of 2019 International Joint Conference on Neural Networks. Budapest, Hungary: IEEE: 1-8 [DOI: 10.1109/IJCNN.2019.8852283]

Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 3431-3440 [DOI: 10.1109/cvpr.2015.7298965]

Lowe D G. 1999. Object recognition from local scale-invariant features//Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE: 1150-1157 [DOI: 10.1109/iccv.1999.790410]

Ma L Q, Jia X, Sun Q R, Schiele B, Tuytelaars T and Van Gool L. 2017. Pose guided person image generation [EB/OL]. [2019-12-13]. https://arxiv.org/pdf/1705.09368.pdf

Matsukawa T, Okabe T, Suzuki E and Sato Y. 2016. Hierarchical Gaussian descriptor for person re-identification//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1363-1372 [DOI: 10.1109/CVPR.2016.152]

Pedagad S, Orwell J, Velastin S and Boghossian B. 2013. Local fisher discriminant analysis for pedestrian re-identification//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE: 3318-3325 [DOI: 10.1109/cvpr.2013.426]

Qi M B, Wang C C, Jiang J G and Li J. 2018. Person re-identification based on multi-feature fusion and alternating direction method of multipliers. Journal of Image and Graphics, 23(6): 827-836 (齐美彬, 王慈淳, 蒋建国, 李佶. 2018. 多特征融合与交替方向乘子法的行人再识别. 中国图象图形学报, 23(6): 827-836) [DOI: 10.11834/jig.170507]

Qian X L, Fu Y W, Xiang T, Wang W X, Qiu J, Wu Y, Jiang Y G and Xue X Y. 2018. Pose-normalized image generation for person re-identification//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 661-678 [DOI: 10.1007/978-3-030-01240-3_40]

Ristani E, Solera F, Zou R, Cucchiara R and Tomasi C. 2016. Performance measures and a data set for multi-target, multi-camera tracking//Proceedings of European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 17-35 [DOI: 10.1007/978-3-319-48881-3_2]

Si J L, Zhang H G, Li C G, Kuen J, Kong X F, Kot A C and Wang G. 2018. Dual attention matching network for context-aware feature sequence based person re-identification//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 5363-5372 [DOI: 10.1109/cvpr.2018.00562]

Siarohin A, Sangineto E, Lathuilière S and Sebe N. 2018. Deformable GANs for pose-based human image generation//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 3408-3416 [DOI: 10.1109/cvpr.2018.00359]

Sun Y F, Zheng L, Deng W J and Wang S J. 2017. SVDNet for pedestrian retrieval//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 3820-3828 [DOI: 10.1109/ICCV.2017.410]

Zhao L M, Li X, Zhuang Y T and Wang J D. 2017. Deeply-learned part-aligned representations for person re-identification//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 3239-3248 [DOI: 10.1109/iccv.2017.349]

Zheng L, Shen L Y, Tian L, Wang S J, Wang J D and Tian Q. 2015. Scalable person re-identification: a benchmark//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 1116-1124 [DOI: 10.1109/iccv.2015.133]

Zheng Z D, Zheng L and Yang Y. 2018. A discriminatively learned CNN embedding for person reidentification. ACM Transactions on Multimedia Computing, Communications, and Applications, 14(1): 1-20 [DOI: 10.1145/3159171]

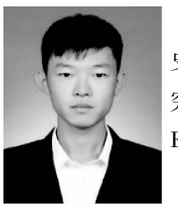
Zhong Z, Zheng L, Cao D L and Li S Z. 2017b. Re-ranking person re-identification with k-reciprocal encoding//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 3652-3661 [DOI: 10.1109/cvpr.2017.389]

Zhong Z, Zheng L, Kang G L, Li S Z and Yang Y. 2017a. Random erasing data augmentation[EB/OL]. [2019-12-13]. https://arxiv.org/pdf/1708.04896.pdf

Zhu F Q, Kong X W, Fu H Y and Tian Q. 2018. Two-stream complementary symmetrical CNN architecture for person re-identification. Journal of Image and Graphics, 23(7): 1052-1060 (朱福庆, 孔祥维, 付海燕, 田奇. 2018. 两路互补对称 CNN 结构的行人再识别. 中国图象图形学报, 23(7): 1052-1060) [DOI: 10.11834/jig.170557]

Zhu Z, Huang T T, Shi B G, Yu M, Wang B F and Bai X. 2019. Progressive pose attention transfer for person image generation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 2342-2351 [DOI: 10.1109/cvpr.2019.00245]

作者简介



史维东,1994 年生,男,硕士研究生,主要研究方向为图像检索、行人再识别。
E-mail: swd1003@sina.com

张云洲,通信作者,男,教授,博士生导师,主要研究方向为智能机器人、计算机视觉。
E-mail: zhangyunzhou@ise.neu.edu.cn

刘双伟,男,硕士研究生,主要研究方向为行人再识别、深度学习。E-mail: 1700915@stu.neu.edu.cn
朱尚栋,男,博士研究生,主要研究方向为计算机视觉和行人再识别、机器学习。E-mail: zhushangdong@gmail.com
暴吉宁,女,博士研究生,主要研究方向为计算机视觉和动态目标跟踪。E-mail: yiyinbaobao@126.com