

开放式行人再识别研究进展综述

夏道勋^{1,2}, 郭方¹, 刘浩杰¹, 夏勇³

(1. 贵州师范大学大数据与计算机科学学院, 贵阳 550025; 2. 贵州师范大学贵州省教育大数据应用技术工程实验室, 贵阳 550025; 3. 西北工业大学计算机学院, 西安 710129)

摘要: 开放式行人再识别是在一个未知的空间环境中, 候选行人库中并不一定包含有待检索的行人, 被视为图像检索的子问题, 是比封闭式行人再识别更具挑战和更为实用的应用研究。本文首先总结了开放式行人再识别的发展历程, 与封闭式行人再识别的异同, 开放式行人再识别建模过程和数据集的比较分析; 然后重点总结了开放式行人再识别技术的研究方法, 它们分别是以数据为驱动的研究方法、以效率为驱动的研究方法和以实际应用为驱动的研究方法; 最后对开放式行人再识别技术的研究难点和未来发展进行了较为详细的分析和探讨。

关键词: 开放式行人再识别; 目标检测; 度量学习; 深度学习

中图分类号: TP391.41 **文献标志码:** A

Review on Research Progress of Open-World Person Re-identification

XIA Daoxun^{1,2}, GUO Fang¹, LIU Haojie¹, XIA Yong³

(1. School of Big Data and Computer Science, Guizhou Normal University, Guiyang 550025, China; 2. The Engineering Laboratory for Applied Technology of Big Data in Education in Guizhou Province, Guizhou Normal University, Guiyang 550025, China; 3. School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: Open-world person re-identification (Re-ID) is a task that the query person may not occur in the gallery set in an agnostic space environment. Compared with the closed-world person Re-ID, open-world Re-ID is seen as a subproblem of image retrieval, and is a more challenging and practical application research. We firstly compare and analyze the different pedestrian datasets and describe the development of existing open-set Re-ID works, the discrepancies between closed- and open-world Re-ID, and the modeling process of open-world Re-ID. Then, we focus on summarizing the research methods of open-world Re-ID, which are data-driven research methods, efficiency-driven research methods and application-driven research methods. Finally, the issues of the research and future directions for open-world Re-ID are discussed.

Key words: open-world person re-identification; object detection; metric learning; deep learning

引 言

行人再识别(Re-identification, Re-ID)作为计算机视觉和模式识别的一个热门研究领域,近年来引

起了学术界和工业界广泛的关注^[1-2]。行人再识别^[1,3-5]旨在无重叠的视频监控网络中,对于摄像机视场中给定的某个行人,判断其是否出现在其他摄像机视场中的过程,是比行人检测更具挑战性的计算机视觉前沿研究方向。行人再识别领域知名学者郑良博士将开放式行人再识别系统总结为行人检测和行人识别两个重要任务^[1],如图1所示。行人再识别技术将行人检测结果作为先验知识,直接对行人图片进行跨摄像头检索。

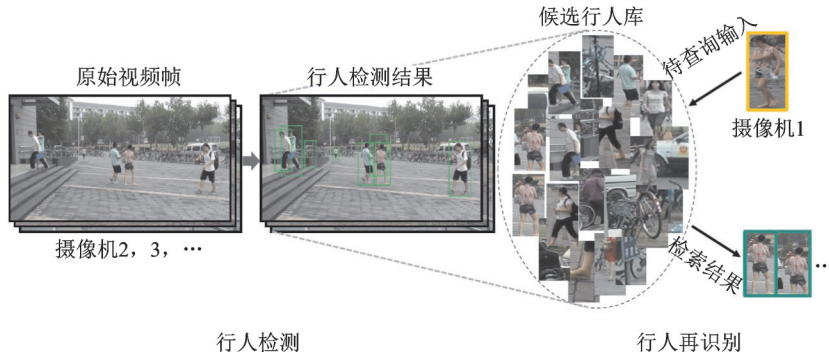


图1 开放式行人再识别系统

Fig.1 System of open-world person Re-ID

经过计算机视觉研究者的不懈努力,行人再识别在很多方面都取得了显著的成功^[1,2,6]。如图2所示,利用一个时间轴来呈现有关行人再识别研究的重要里程碑节点,例如基于度量学习和深度学习的行人再识别已经促进了这一研究领域的发展。然而,现有的行人再识别研究大多是针对封闭场景而不是真实场景下进行,这在很大程度上限制了行人再识别技术的实际应用,通常同一个行人的表现差异可能会大于不同行人之间的表现差异,这种变化差异由视角改变、光照差异、姿势变化、分辨率差异以及遮挡等因素造成。近年来,智慧城市和数字城市推动了智能安防项目的不断发展,大规模视频监控下开放式的行人再识别研究成为主要研究方向,拓宽了计算机视觉的应用领域和研究深度,是一个更具挑战 and 更为实用的开放式应用研究,多镜头下基于深度神经网络的特征学习方法和端到端的学习方法成为行人再识别的主流研究方向^[7-9]。

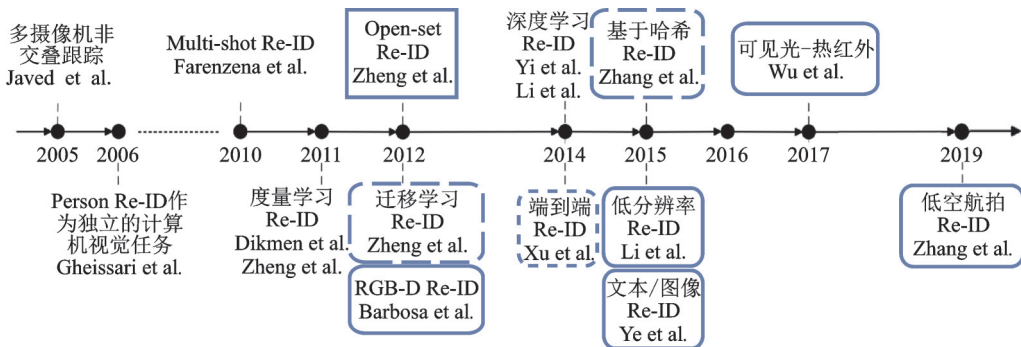


图2 行人再识别的发展历程

Fig.2 Development process of person Re-ID

1 开放式与封闭式的异同

开放式场景行人再识别最早是由 Zheng 等^[1-2,10]提出,是指在一个未知的空间环境中,跨摄像头的

大规模行人 ID 可能存在部分重叠,候选行人库并不一定包含有要检索的行人 ID,若检测到新的行人 ID,算法会提取新的行人特征并加入到候选行人库中。和开放式 Re-ID 相反,封闭式场景行人再识别通常假设需要查询的行人 ID 是一定存在于候选行人库中,研究的重点是确定候选行人库中的图像属于查询行人 ID,并返回查询结果。表 1 分别从识别性质、目标、数据集设置和评价指标细化了它们的差异^[2]。

表 1 开放式与封闭式行人再识别技术的差异
Table 1 Comparisons of closed-set and open-set Re-ID settings

行人再识别	封闭式	开放式
性质	识别	验证
目标	在候选行人库中寻找与查询行人最相似的图像	验证查询行人是否在候选行人库中出现过
数据集设置	摄像机参数基本相同,行人 ID 必须经过指定的摄像机视场	候选行人库中存在大量无关的行人 ID,待查询的行人可能不在候选行人库中
评价指标	CMC、mAP	TTR、FTR

回归本质,行人再识别技术是为了维护公共安全以及安全应用而发展起来的^[2,11],例如在实际监控视频系统中进行视频侦查,寻找特定人员的出现地点以及他的行动轨迹。根据实际应用需求,封闭式行人再识别的这种假设并不一定成立,查询行人可能并没有出现在摄像头的视场中,候选行人库不存在待查询的特定行人 ID。因此,开放式行人再识别比封闭式行人再识别更接近真实场景,挑战性更大^[2,12]。

作为一项识别任务,封闭式和开放式行人再识别息息相关、互为补充^[2,7,12]。封闭式行人再识别的目的是在摄像机精确标注行人 ID 的前提下,确定候选行人库中的图像属于查询行人。开放式行人再识别则侧重于查询行人 ID 在候选行人库中的可能性,评估指标独立于一对一的身份识别,采用高真实目标识别(High-realistic target recognition, TTR)和低假目标识别(Low false target recognition, FTR)。

2 建模过程

开放式行人再识别系统功能模块包含视频数据采集、行人检测和跟踪、行人再识别数据集自动标注、识别模型训练、行人再识别检索和系统性能评估,如图 3 所示。

视频数据采集是从大规模视频监控中众多的摄像机中采集和处理视频流数据,以达到行人样本具有多视角和地理位置多样性等特点;行人检测和跟踪是按照采样规则,将视频帧送入行人检测器,获得

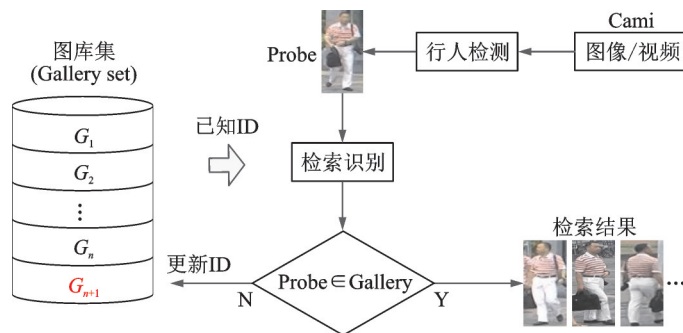


图 3 开放式行人再识别建模过程

Fig.3 Modeling process of open-world person Re-ID

标记有“good”的行人样本再送入行人跟踪器,把同一个行人的检测框 Bounding boxes 关联起来,形成 Tracklet 序列,最后获得该行人一系列不同视角和姿态的样本;行人再识别数据集自动标注是将新获得的 Tracklet 序列与现有 Gallery 数据集进行度量,根据度量结果进行不同的归并操作,归并操作有两种,一种是与现有数据集中某行人 ID 类合并,另一种是在现有数据集中新建行人 ID 类,更新 Gallery 数据集;识别模型训练是按照分配规则,将数据集分为训练集和测试集,再将测试集划分为 Gallery 和 Probe,采用深度学习训练识别模型;行人再识别检索是在识别模型和 Gallery 的支持下,判断 Probe 是否包含在 Gallery 中,如包含则获得检索结果,如果不包含,则增量更新 Gallery 数据集;系统性能评估是采用 TTR 和 FTR 指标来衡量系统的识别性能。

3 数据集比较分析

3.1 数据集

目前行人数据集主要包括封闭式数据集和开放式数据集两大类^[2, 8, 13],未来发展趋势主要表现为:(1)数据集所包含的行人图像在不断增加,但还不能满足实际应用的需求;(2)部分研究工作尝试通过利用一些标注算法来自动标注数据集,但相对于人工标注来讲,自动标注准确率很低;(3)采用了多摄像机组拍摄行人图像,但还没有达到真实场景所需要的规模;(4)构建航拍图像大数据下的行人再识别数据开放集受到重视^[8]。在开放式环境下,行人再识别被看作为验证而不是识别,这意味着会有许多无关的行人 ID 存在,而且在图库中并没有正确的匹配。下面以特定数据为驱动对数据集进行分述。

(1)可见光-热红外基准数据集(Benchmarks for visible-thermal Re-ID)。RegDB^[14]和 SYSU-MM01 是两个特定的可见光-热红外基准数据集。RegDB 数据集包含了 421 个行人 ID,被双摄像机系统采集,对于每个行人 ID,RGB 摄像机采集 10 张可见光 RGB 图像,热红外摄像机获取 10 张热感应图像。SYSU-MM01^[15]数据集由中山大学创建,使用了 4 个 RGB 摄像机和 2 个热红外摄像机进行数据采集,每个行人至少在两个不同的摄像机视场下经过,共采集了 491 个行人 ID 图像,共包含 30 071 张 RGB 图像和 15 792 张热感应图像。在可见光-热红外数据集下进行行人再识别研究非常具有挑战性。

(2)RGB-D 基准数据集(Benchmarks for RGB-depth Re-ID)。现有的 4 个公开发布的 RGB-D 数据集,它们分别是 PAVIS^[16]、BIWI RGBD-ID2、IAS-Lab RGBD-ID3 和 DPI-T^[17]。PAVIS 数据集包含 79 个不同行人 ID,行人具有 4 种运动状态,它们分别是 Collaborative、Walking1、Walking2 和 Backwards,数据集使用 Walking1 和 Walking2 分别作为训练集和测试集。BIWI RGBD-ID 数据集包含 50 个不同行人 ID,含有 50 个训练序列和 56 个测试序列,同步了 RGB 图像和 Depth 图像,分割映射和骨骼数据由 Microsoft Kinect 提供。与 BIWI RGBD-ID 数据集提供的 10 fps 视频帧率相比,IAS-Lab RGBD-ID 数据集提供的是 30 fps 视频帧率,包含 11 个不同行人 ID、11 个训练序列和 22 个测试序列。DPI-T 数据集是一个新的基于深度图像的行人再识别基准数据集,采用自上而下的摄像机历时数天拍摄而成,含有 24 个行人 ID 和 25 段视频。

(3)文本-图像基准数据集(Benchmarks for text-to-image Re-ID)。现阶段,已有的标准数据集均未自带自然语言描述集,用于文本-图像的数据集通常是在现有的数据集基础上重新标注。例如, CUHK-PEDES 数据集^[18]就是将 CUHK03^[19]、Market-1501^[20]、VIPeR4 和 CUHK01^[21] 等数据集聚集起来,进行二次标注,包含 13 003 个行人 ID 和 40 206 张图像,每张图像包含 2 条文本描述信息。

(4)低分辨率基准数据集(Benchmarks for low-resolution Re-ID)。现有几项出色的低分辨率行人再识别工作^[21-23],是将已有的公共数据集(例如 VIPeR、CUHK03、PRID450s 和 3DPES)进行下采样,产生众多低分辨率行人图像。以 SALR-VIPeR 数据集^[24]为例,VIPeR 数据集包含两个摄像机采集的图像,共采集 632 个行人 1 264 张户外图像,将摄像机 A 采集得到的图像设为高分辨率查询图像,再将摄像机 B 采集得到的图像用随机下采样法,得到不同尺度的低分辨率图像。虽然采取随机下采样法这类

模拟方式获得的低分辨率图像,不能很好地适应真实场景应用,但是能提供不同分辨率的行人图像,对低分辨率行人再识别研究起到了积极作用。CAVIAR4ReID 数据集由两个摄像机在购物中心采集而得,包含 72 个行人 ID,摄像机 A 的图像分辨率比摄像机 B 的图像分辨率低很多,非常适合交叉分辨率行人再识别方法研究。PRW^[25] 模拟了视频监控场景,利用 6 台水平摄像机采集行人目标,6 台摄像机中有 5 台是高分辨率摄像机,另一台是低分辨率摄像机。

(5)基于哈希检索基准数据集(Benchmarks for Hash-based Re-ID)。在现已公布的行人再识别数据集中,例如 CUHK03,Market-1501,LSPS^[26],SMT17^[27] 和 Airport^[28]等,行人 ID 有数千个,是大型的行人再识别数据集。然而,相对于实际场景所需的行人 ID 数量,现有的数据集规模实际上很小。因此,现有基于哈希检索的方法必须在相对较大的数据集上进行验证。为了获取有用的数据,可以利用 GANs 算法生成复杂的样本,这些样本带有标记和未带标记,有效地扩充了现有的数据集规模,行人图像的数据量就可以得到快速的发展和自动标注。

(6)低空影像基准数据集(Benchmarks for aerial imagery Re-ID)。无人机低空对地行人再识别技术拓宽了行人再识别的应用领域和研究深度,无人机航拍行人分辨率低,尺度和角度变化剧烈,行人外观特征非常弱,挑战性极大,具有代表性的低空影像基准数据集有 PRAI-1581^[4]、VisDrone2019^[29]、AVI^[30]、MRP(Dataset1)^[12] 和 MRP(Dataset2)^[12]。PRAI-1581 是由空地一体化大数据应用技术国家工程实验室、西北工业大学和西安交通大学联合构建,是专门为空地一体化行人再识别技术研究而构建的开放式数据集,数据集共采集了 39 461 张图像和 1 581 个行人 ID,训练集由 19 523 张图像 782 个行人 ID 组成,测试集由 19 938 张图像 799 个行人 ID 组成,查询图像(Query)和候选行人库(Gallery)的图像数量分别是 4 680 和 15 258 张。VisDrone2019 数据集由天津大学机器学习和数据挖掘实验室的 AISKYEYE 团队收集,采集遍布中国 14 个城市,包含 288 个视频片段、272 117 张视频帧/图像和 260 万个标注框(Bounding boxes)。AVI和MRP数据集的详细内容见图 4 和表 2,由于篇幅所限,其他未介绍的行人再识别数据集见 Person Re-identification Datasets 网址所列举,这里不再赘述。



图 4 Market-1501、MRP、AVI 和 PRAI-1581 数据集部分图片示例

Fig.4 Some illustration images of Market-1501, MRP, AVI and PRAI-1581 datasets

表 2 低空影像相关基准数据集详细参数表

Table 2 Detailed parameter table of benchmarks datasets for aerial imagery

数据集	图像数据(#Images)	行人 ID 数据(#ID)	采集平台	飞行高度/m	摄像机	无人机数量
Market1501 ^[12]	32 668	1 501	固定机座	<10	固定	0
MRP(Dataset1) ^[23]	1 501	23	无人机	<10	移动	3
MRP(Dataset2) ^[23]	4 096	28	无人机	<10	移动	6
AVI ^[22]	10 863	5 124	无人机	2~8	移动	
PRAI-1581 ^[4]	39 461	1 581	无人机	20~60	移动	2
VisDrone2019 ^[21]	39 988	94 396	无人机		移动	14

3.2 软件库

在行人再识别研究方面,本课题申请人已做了大量基础性研究,已完成行人再识别总框架代码的编写,并在上述主要数据集中进行了验证。也一直致力于行人再识别的很多优秀研究者,将自己的研究成果包括代码库进行发布共享,加快了该领域的研究步伐。下面列举部分成果:在数据增强(Data Augmentation)算法方面就吸引了很多学者的注意,比如翻转、裁剪和 Random Erasing, Yi 等^[31]、Zhong 等^[32]、Deng 等^[33] 和 Wei 等^[27] 利用 GAN 技术对行人再识别进行数据增强,分别提出了 DualGAN、Camera Style Adaptation、CycleGAN 和 PTGAN 等模型,并公布了代码库;在特征学习方面,Zhao 等^[34] 提出了基于中层滤波学习的行人再识别研究方法,引入了结构支持向量机(Support vector machine, SVM)以及 DenseColor 和 DenseSift 特征,包含了 Patch Matching 和 Graph Clustering 等工具包;中山大学 iSEE 智能科学与系统实验室研究团队和香港中文大学电子工程系机器学习研究团队发布的行人再识别代码库。

在识别模型方面,特别是最近几年深度学习、无监督学习和自监督学习等的兴起,行人再识别研究者们公布了多个优异的算法模型,具有代表性的模型有 Zheng 等^[35] 的 Person Re-ID Baseline Pytorch、Zhou 等^[36] 的 Deep Person Re-ID、Lv 等^[37] 的 Rank Re-ID 和 Yu 等^[38] 的 MAR 模型等。在空地一体化行人再识别技术方面,已构建了大型完善的空地行人再识别 PRAI-1581^[8] 数据集、低空行人目标检测 Vis-Drone2019^[29] 和 AVI^[30] 数据集,Ardeshir 等^[7]、Grigorev 等^[9] 和 Layne 等^[12] 研究者已在空地行人再识别做出了具有原创性和开拓性的研究,为后续研究指明了方向。另外,大数据产业已上升为国家战略,大数据开源软件平台也取得飞速发展,重要的开源平台有 Hadoop 和 Spark 等。HDFS、HBase、MapReduce 和 Spark Streaming 等子框架提供了图像数据的海量存储、高速运算和实时流数据处理,Spark MLlib13 提供了可扩展的机器学习库,包含分类、回归、聚类、协同过滤、降维,以及底层的优化原语等算法和工具,有助于项目的研究。

4 以数据为驱动的开放式研究方法

最近几年,行人再识别研究的关注正在迅速增长,基准数据集的数据量有了很大的提高,早期的 VIPeR 和 GRID 只有数以百计的行人 ID 和数千张图像,然而更新过后的 MARS14 数据集^[39] 已有上千个行人 ID 和上百万张图像,大规模和易用性的数据集将是整个行人再识别领域发展的关键点。现有的许多基准数据集考虑了常见的人物外观变化的影响,例如视角、光照、分辨率和遮挡,以及跨摄像机的视角变化^[1,11]。但是,为了更好地将行人再识别技术应用于开放式的实际场景中,最近的一些研究已经提出了通过挖掘原始多模态行人特征,例如热感应、深度和文本特征,以及除一般可视图像之外的低分辨率行人数据图像研究,下面分类介绍最新的以实际数据为驱动的开放式研究^[2,7]。

4.1 可见光-热红外行人再识别

RGB 摄像头无法在低亮度或者是无照明场景(如夜间)下捕捉到有效的行人目标外观信息,热红外成像监控摄像头在夜间或者黑暗场景中被广泛使用,使用 RGB 摄像机和热红外摄像机可以交替地采集到 24 小时行人图像,这对于视频侦查应用至关重要,并且意义重大^[40-41]。在恶劣光照环境下,将 RGB 行人图像与热红外成像进行跨模态匹配,称为可见光-热红外(Visible-thermal)行人再识别,热红外成像行人图像可以增强行人再识别技术的鲁棒性^[41]。RGB 图像包括 3 个可见光描述通道,热红外图像只包含一个不可见光信息通道,增加了跨模态行人匹配的困难。此外,可见光-热红外行人再识别同样需要处理视角变化、低分辨率和遮挡等问题,所以可见光-热红外行人再识别极具挑战性^[40]。

Kai 等^[42] 是最先进行可见光-热红外行人再识别有关的研究者,于 2010 年提出基于局部特征码本的热红外成像行人再识别,他们仅对单模态热红外成像进行了研究,没有涉及到更重要的交叉模态研

究。Wu等^[15]于2017年提出了可见光-热红外行人再识别的基本问题,他们提供了一个极具意义的跨模态行人再识别基准数据集SYSU-MM01,并提出了深度零度填充(Deep zero-padding)方式进行跨模态匹配,在One Stream框架下,算法将两种Mix(Mix-RGB和Mix-IR)以共享参数的方式结合使用,能够自主地进化RGB与IR图像节点。之后,Ye等^[23,43]提出了两阶段、双路径端到端的可见光-热红外行人再识别学习框架,该框架利用了跨模态共享机制和跨模态与模态内变异机制,为跨模态学习指明了新的方向。现有的研究主要是将不同的模态嵌入到同一个特征空间中学习共同的特征表达,然而只学习共同特征意味着信息损失巨大,降低了特征之间的差异性,Lu等^[44]提出了一种新的跨模态共享特征转移算法(Cross-modality shared-specific feature transfer, cm-SSFT)来解决上述问题,以探索模态共享信息和特定特征在提高再识别能力上的潜力。为了解决红外光-可见光跨模态的差异,Li等^[45]引入中间模态X,用轻量级网络,以自监督的方式学习到可见光-红外图像的知识,将较难的双模态跨模态Re-ID任务转化为较容易的三模态跨模态Re-ID任务。Wang等^[46]提出了生成跨模态配对图像的方法,并执行全局集合级和细粒度实例级对齐,这种方法可以解开特定于模态和模态不变的特征来执行集合级对齐,可以显式删除特定于模态的特征,也可以更好地减少模态变化。虽然热红外成像的重要性引起了很多学者的重视,但是热成像过程中损失了大量图像细节,现有模型的识别性能仍然很低,可见光-热红外行人再识别的诸多问题仍未得到解决^[2]。

4.2 RGB-深度行人再识别

与热红外成像相类似,深度图像在极低的光照和颜色条件下可以保持更多的不变性,穿不同衣服的行人通过某台摄像机视场时,深度图像提供的行人体型和骨骼信息非常重要。Barbosa等^[16]于2012年首次使用RGB深度传感器对行人再识别进行研究,他们利用3D软生物特征识别技术有效地避开行人服装变化问题。Mogelmose等^[47]融合了RGB、深度和热红外成像数据,训练成联合分类器。Munaro等^[48]利用三维重建和行人生物特征技术研究行人再识别。Haque等^[17]提出了一种基于注意力机制的行人再识别模型,该模型联合了卷积神经网络和递归神经网络,能指定行人所在的显著性区域。在缺失RGB信息的情况下,为了提取行人的形态和运动学四维时空特征,Wu等^[49]提出了一种描述人体局部旋转深度形状不变描述子,然后研究了基于核函数的隐式特征转移,将估计的深度特征与RGB外观特征相匹配。在行人再识别跨模态研究中,Hafner等^[50]深入研究了RGB图像和深度图像之间较少被探索的问题,提出了一种新的跨模态提取网络,该网络学习了RGB图像和深度图像中行人表现的共享特征,使得行人再识别模型更具有鲁棒性。姜国权等^[51]提出一种基于RGB信息与灰度信息融合的行人再识别方法,所提方法能够对RGB信息和灰度信息进行有效融合,并且融合信息有助于提高算法对颜色变化的鲁棒性,提升算法性能。

深度信息有利于提取行人特征,能提供行人的体型和骨骼信息,有利于提升行人再识别的准确率^[2,49]。但是,在实际场景中,有两个显而易见的问题需要继续研究和探索:(1)深度摄像机(例如Kinect)常被应用于室内场景,很少在户外场景中使用,因为深度信息会随着行人和摄像机之间的距离增加而迅速减小;(2)当行人图像的行走角度发生变化时,深度图像无法区分人体形状和骨骼信息。

4.3 文本-图像行人再识别

除了由各种摄像机拍摄到的行人图像外,目击者对行人(或者嫌疑人)的自然语言文本描述也是候选行人库的待查询输入(Probe),这是一种重要的侦查方式^[2,52]。如何将自然语言待查询输入与候选行人库进行匹配,这是一项有趣且有意义的工作,这个研究方向被称作文本-图像(Text-to-image)行人再识别^[52]。这项研究近似于基于行人属性的行人再识别研究,主要思想是区分不同行人的中级属性或高级语义,而不是行人的底层特征。与传统的基于行人属性的行人再识别研究相比,基于语义验证的行人再识别具有两个优点:(1)行人属性与行人图像块有关,例如行人上身穿衬衫和下身穿裤子;(2)与分

类任务中的行人属性相比,用于行人属性分类的数量相对有限。因此,行人属性自然语言描述的行人再识别方法研究得到发展^[53],它们可以在其他的行人再识别综述中找到^[1,2,4,11]。

在缺乏行人(或者嫌疑人)图像的情况下,目击者对行人的描述是自然语言语句,跨模态文本-图像行人再识别在开放式真实场景中具有很强的实用性^[54-55]。Ye等^[54]于2015年通过不完整的行人文本描述来解决特定的行人检索问题,是首次针对文本-图像行人再识别得到的研究成果。Li等^[55]提出了孪生网络识别框架解决文本-图像的匹配问题,第一阶段利用跨模态交叉熵损失学习嵌入式跨模态特征CNN-LSTM网络,第二阶段构建了隐式注意机制CNN-LSTM网络,以细化匹配结果。在此基础上,Li等^[18]进一步提出了一种具有门控神经注意机制的递归神经网络,并构建了带有语言注解功能的大规模人物描述数据集。Chen等^[56]使用自然语言描述符辅助训练识别模型,利用行人整体语言描述符监督学习,获得更好的行人全局视觉特征,并构建了全局和局部图像语言关联规则,实现局部视觉特征和语言特征之间的语义一致性。

近期,文本-图像行人再识别的关注度有所下降,可能有如下原因^[2]:(1)研究文本-图像行人再识别技术极具挑战性,已经公布的模型识别能力非常低下;(2)对行人搜索和目标搜索的通用跨模态文本-视觉匹配框架已经实现,与文本-图像行人再识别之间的界限不够清晰。但是,从视频侦查应用的角度来看,文本-图像行人再识别可以找到两个重要的研究点^[2,18]:(1)目击者对嫌疑人的文本陈述总是不完整,通常是在犯罪发生后才会回忆起嫌疑人;(2)目击者的文本陈述并不一定都正确(例如在光照条件差的环境下,目击者将嫌疑人的衣服颜色识别错误)。因此,文本-图像行人再识别有其自身的独特性,值得研究者进一步探索。

4.4 低分辨率行人再识别

在监控区域部署高分辨率摄像机,代价非常昂贵,通常只能在重点监视区域部署,中低分辨率行人图像远比高分辨率行人图像更为常见,导致行人样本图像的分辨率各不相同^[11,57-59]。但是,现有大多数识别模型都是将行人样本图像归一化到特定的尺度,不考虑图像分辨率差异的影响。如果简单地将高分辨率图像归一化到中低分辨率图像,会导致部分重要视觉信息丢失;反之,图像的缩放会产生大量的噪声^[57]。为了弥补行人再识别中高分辨率和中低分辨率行人图像之间的差距,Jiang等^[58]提出了通用超分辨率(SR)技术对数据集进行预处理,将中低分辨率行人图像映射到高分辨率图像。

Li等^[57]于2015年最先开始低分辨率行人再识别研究,通过最小化异构均值差异优化跨模态尺度图像对齐,提出了一种多尺度距离度量学习鉴别模型。Jing等^[60]提出一种半耦合鉴别词典技术,采用低秩正则化技术学习鲁棒的高分辨率图像和中低分辨率图像特征空间字典,并将高分辨率图库图像特征与中低分辨率查询图像特征进行映射。Wang等^[24,59]实现了一种尺度自适应公式,将高分辨率查询图像与不同尺度的中低分辨率图库图像进行匹配。针对尺度自适应交叉分辨率识别问题,先后有人提出了判别曲面学习模型和级联超分辨率GAN框架。Jiao等^[61]联合超分辨率行人图像和行人ID学习混合深度CNN网络,解决低分辨率行人再识别问题。Guo等^[62]提出了分辨率不变性网络,探索高分辨率和低分辨率特征高度相关的子空间,并从共享空间中提取出具有判别性的特征。沈庆等^[63]提出一种多分辨率特征注意力融合的行人再识别方法,利用主干网络HRNet实现注意力机制,达到高效地表征行人的目的,并利用交错卷积构建4个不同的分支来抽取多分辨率行人图像特征,实现行人不同粒度特征的抽取和不同分支特征的交互。

低分辨率问题不仅存在于中低分辨率摄像机中,也存在于高分辨率摄像机中^[2,59]。例如,如果一个行人出现在距离高分辨率摄像机很远的地方,行人成像规模很小,也就变成了低分辨率图像。还可以对不同尺度的单模态RGB图像进行匹配,从而很容易获得数据。因此,近年来开放式低分辨率行人再识别比其他特定数据驱动的行人再识别研究更受关注^[60]。从现有的低分辨率行人再识别模型的性能

可以看出,经过分辨率增强或多分辨率融合改进后的行人再识别模型性能都超过了传统行人再识别模型的性能,这表明低分辨率问题不仅能有效提升行人再识别模型的精度,而且对于行人再识别技术的实际应用也十分重要。

综上,获取原始数据是实际视频应用研究的首要要求,行人再识别模型的好坏最主要的影响因素是数据集质量,除了确保数据标注精度的质量,还需要适应光照变化、昼夜变化、视觉信息缺失和分辨率差异,以数据为驱动的跨模态研究就显得尤为重要。跨模态需要同时从多个异质信息源提取行人对象的特征,挖掘不同模态之间子部件的相关或对应关系,从而促使学习到的跨模态表示更加精确。

5 效率驱动的开放式研究方法

MARS 数据集^[39]包含了数千个行人ID和大量行人图像,是目前用于行人再识别研究最多的数据集之一,但远远没有达到实际应用的数据规模^[1-3,64]。如果在仅有10个视频监控摄像机的小型城市里,监视时间范围为24小时,每秒25帧,就能获得 $10 \times 24 \times 3\,600 \times 25 = 2.16 \times 10^7$ 张完整的视频帧,利用行人检测器(如DPM、Faster RCNN)检测行人Bounding boxes,其数量将超过108个,是MARS数据集的数百倍。如果在一个中等规模的城市里,视频监控摄像头的数量通常都会超过1万台,行人Bounding boxes的数量则迅猛增长。现阶段,还没有研究者在如此大规模的数据集下进行行人再识别研究工作,Zheng等^[20]却做了一个有意义的实验,在Market-1501数据集上测试了两个数据大小相同的行人再识别基准数据集。实验结果表明,与19KB的数据集相比,500KB的Market-1501识别性能mAP降低了7.0%,换句话说,数据量的增长会使行人再识别的精度降低。与25倍的数据量增长相比,7.0%的精度下降是可以接受的。但是,如果数据规模与模型运行时间之间存在近似的线性关系,其检测和识别效率很难满足实际应用需求。

从方法论角度而言,效率问题主要指行人再识别算法的时间复杂度。近年来,深度学习已经成为主流趋势^[1,3-4],基于深度学习的识别模型几乎达到了行人再识别基准的最佳精度,但这精度的提升是以高昂的计算成本为代价,提取深度特征的算法复杂度与图像像素数量成线性关系,即使是利用简单的CNN行人特征训练行人再识别模型,也不适用于大规模视频监控系统。为了平衡效率和准确性,同时克服大规模行人ID标注效率问题,倒排索引、哈希检索和迁移学习是行人再识别重要的研究方向^[2]。

5.1 基于哈希检索行人再识别

行人再识别的计算时间严重地依赖数据集规模、行人特征和模型复杂度,当数据集大小远远超过大规模时,哈希检索能有效地减少搜索近邻的成本^[65-67]。Zhang等^[65]提出了第一个基于哈希检索的行人再识别框架,该框架从原始图像中学习紧密且位可扩展的哈希编码,该工作构建了一种从特征提取到哈希函数学习的端到端体系结构,其效率明显比一般基准模型的效率高出了几十倍。Zhu等^[68]提出了一种基于局部的深度哈希方法,将深度学习和哈希检索集成到一个框架中,以提升识别模型的精度和效率。深度哈希检索模型的输入为大量的三重态样本,通过行人类内和类间汉明距离约束的三重态损失函数进行学习。实验结果表明,该模型可以平衡大规模行人再识别的精度和效率。Wu等^[69]提出了一种深度哈希检索框架,框架将正样本对(Positive pairs)和硬负样本(Hard negatives)增加到Mini-Batches中,学习到一个结构化损失函数,同时学习到CNNs网络和哈希检索函数,从而获得鲁棒的深度行人特征和能区分行人相似度的哈希编码。

5.2 基于迁移学习行人再识别

受数据集有无标注信息(监督信息或先验信息)的影响,有监督行人再识别因其性能明显优于无监督方法,一度成为行人再识别技术主要的研究方向^[2-3]。近年来,虽然出现了几种具有代表性的无监督学习方法^[70-72],但它们的性能仍然不如最先进的有监督学习方法。与无监督学习方法不同,迁移学习

不仅利用了目标域的未标注数据,还利用了原域中已标注的数据,因此基于迁移学习的行人再识别方法通常优于无监督学习,也接近于有监督学习模型的性能^[73]。

Ma等^[74]提出了一种自适应排序支持向量机模型,用于优化从源域学习到的距离模型,该模型对目标正均值有较高的置信度,对目标负图像对有较低的置信度。在不同行人再识别场景下,Wang等^[75]利用非对称方式联合学习特征相似性度量模型,获得串扰数据的共享组件,并在联合学习过程中增强目标类间的距离。弱监督迁移学习使用了目标数据集的部分标注信息,Peng等^[76]则提出了基于多任务字典学习的跨数据集自适应模型,完全没有使用目标数据集的标注信息。Chen等^[77]构建了深度迁移学习框架,提出了两步微调策略,通过协同训练从辅助数据集中迁移知识。Li等^[78]提出了一种自适应行人再识别网络,网络聚焦在跨数据集信息的利用和域不变特征的推导。Wang等^[79]提出了一种可迁移属性-身份深度学习框架,框架从带有标注的源域到未标注的目标域,共同学习属性语义和具有判别力的身份特征表示空间。Lv等^[38]在目标域中部署了一种基于行人时空模式的增量迁移学习排序方法。Deng等^[33]引入了一种持续生成对抗网络,该网络包含一个孪生神经网络和一个CycleGAN网络,用于域自适应行人再识别,网络学习过程有两个约束条件:(1)图像迁移前后的自相似性;(2)源域图像迁移到目标域图像的域差异性。

Zhang等^[80]针对目前模型很难有效地泛化到无标签的目标域问题,提出了一种逐步增强的自训练方法,他们将模型的训练分为两个阶段,分别提高特征表达能力和匹配源域与目标域的数据分布,并利用了基于排序的三元组损失函数,提高三元组样本的可靠性。Liu等^[81]将复杂的域迁移问题分解成3个不同的子转换问题,每个子转换专注于处理某种图像转移风格,比如光照、分辨率、照相机视角。域不变性映射网络(Domain-invariant mapping network, DIMN)为了保证域的不变性^[82],采用了元学习算法,每次训练都在源域中采样子集训练Task,模型学习一张行人图片与其行人ID分类向量的映射,产生一个单样本分类器。在目标域上,每张Gallery图像都会输入到网络中产生一个ID分类向量,Probe图像与行人ID分类向量做点积运算,完成行人的匹配过程。Yang等^[83]提出了一种新的非对称协作网络,利用两个网络分别接收尽可能纯净的样本和尽可能多样的样本,在类协同框架和样本噪声过滤的同时,将更多低置信度样本纳入到模型的训练中,从而保证训练样本的多样性,提升域适应精度。

Huang等^[84]提出了一种新的域自适应注意力模型,将具有区分性的特征从带有标签的源域转移到没有标注信息的目标域中,该模型包含域共享特征映射和域指定特征映射两部分,设有两个约束条件,它们分别是域间相似度损失和基于聚类过程的加权交叉熵损失。Park等^[85]为行人再识别提出了新的关系网络,引入了人体各个部分与其他部分之间的关系,模型可使一个单独的部件级特征中包含人体其他部分的信息,使得模型更具有识别力。Jin等^[86]提出了基于语义对齐的特征学习网络进行行人重识别,引入人体空间语义对齐的全视图重建任务,实现了单(视角)张图像预测全视角人体外观的能力,解决了行人重识别中图像间空间语义不对齐的难题。Yu等^[87]引入了一种批相关卷积单元(BConv cell),它将数据集的全局信息逐步收集到一个潜在状态,并利用这个潜在状态对提取的特征进行校正,利用BConv细胞单元进一步提出了渐进迁移学习方法,通过联合训练BConv细胞单元和预训练Re-ID模型来完成模型的微调过程。为了解决实际数据的特征样本不平衡问题,林通等^[88]设计了迁移姿态生成对抗网络模型,能很好地解决目标不同姿态的干扰以及数据集中图片数量不足的问题。

迁移学习行人再识别研究已经引起了广泛的关注和研究,从现有研究的实验结果来看,源域的监督知识可以很容易地引入到目标域中^[2,72]。但是,当前的迁移学习机制通常都是假设未标注的目标域是已知或已建立,很少有研究涉及到目标域并不确定的实际情况,从一般行人目标到特定行人目标的无监督迁移学习是一项重要但尚未被探索的研究任务^[2,33]。

综上,深度学习的兴起加快了研究者对行人再识别的研究,并且取得了前所未有的识别效果。但是,随着深度学习特征维度的增加以及图库(Gallery)规模的加大,其算法的处理时间显著性地增加,也

对高能耗的计算设备依赖更为严重,这对开放式行人再识别的实际应用并不可取。因此,以效率为驱动的行人再识别也逐渐受到研究者的重视,也可以从 CVPR2020 和 ECCV2020 等国际会议关于行人再识别技术论文的研究主题中可以得出,已有基于深度哈希算法、基于行人图像图像检索和行人自动标注等算法出现。

6 以实际应用为驱动的开放式研究方法

以实际应用为驱动的开放式行人再识别是在全帧图像图库中寻找一个待查询的行人 ID,这个过程分为两个阶段^[1,25,89]:(1)对图像中的行人进行检测和跟踪预处理;(2)身份匹配再识别。在一般封闭式行人再识别中,仅仅利用行人检测和行人跟踪提供所需的行人数据,类似于手动标注数据集的过程。现有大多数基准数据集的行人 Bounding boxes 清晰且噪声很少,这在实际应用中并不可取;小部分数据集利用行人检测和行人跟踪自动生成数据集的 Bounding boxes,例如 Market-1501 数据集使用 DPM 检测器自动产生不同 Bounding boxes 的宽高比,并标记为“good”“distractor”或“junk”,但是行人检测器无法避免地产生大量带有“distractor”或“junk”标记的 Bounding boxes,从而大大降低了行人再识别系统的性能^[2]。为了缩小行人再识别技术与实际应用需求之间的差距,必须解决行人检测器(包括检测和跟踪)对于行人再识别精度的影响,从检测到识别的端到端框架是一个有效的开放式行人再识别研究方向。

6.1 端到端行人再识别

端到端行人再识别方法有两种描述^[2,90]:(1)从行人检测到行人身份匹配的端到端行人再识别;(2)从特征提取到距离度量的联合学习框架。后者从手动标注 Bounding boxes 开始,侧重于特征和度量联合学习,后面重点阐述前者。Xu 等^[91]认为错误的行人检测 Bounding boxes 被带入行人再识别不可避免,他们提出了一种滑动窗口搜索策略,用于行人检测和身份匹配。实验表明,该联合模型的搜索性能优于 Two-stages 的搜索性能。随后,Xiao 等^[26]基于在线实例匹配损失函数提出了一个端到端的深度学习框架,用于定位和匹配场景图像中出现的查询行人 ID,同时他们还构建了大规模行人搜索数据集 LSP15。Alemu 等^[92]提出了一种端对端约束聚类模式来解决待检索行人图像在相似性扩散的过程中容易造成噪声传播的问题,他们将行人图像检索问题看成是一个约束聚类问题,通过端对端的方式来优化约束聚类,可以更好地利用行人图像的上下文知识。Li 等^[93]提出外观表征与运动增强模型,外观表征增强模块(Appearance enhancement module, AEM)是利用伪标签进行有监督学习,挖掘行人属性特征,运动增强模块(Motion enhancement module, MEM)是通过预测未来帧获取具有身份判别性的运动模式。Wu 等^[94]提出了一种新的轨迹自监督学习方法,利用大量未标记的轨迹数据来优化视频和图像的无监督特征空间,可以在缺少源数据标注、行人 ID 标签和摄像机先验知识的条件下进行端对端的模型训练。Chen 等^[95]提出了一种框架引导区域对齐模型,用在端对端训练中学习具有判别性的特征表达,解决了视频中行人通常处于运动状态从而导致的严重空间错位问题。Li 等^[96]提出了两个关系指导模块来学习增强的行人特征表达,空间注意力模块用来探索全局判别性区域,每个位置的权重由它的特征以及关系特征决定,揭示了人体局部特征与全局特征之间的依赖关系,改进的时序模块通过融合帧与帧之间的关系信息来提取具有鲁棒性的视频级特征。针对目前基于视角的 Re-ID 方法大部分只考虑单一视角下不同行人的分布,而忽略了不同视角之间潜在关系的问题,Zhu 等^[97]提出了基于角度正则化的视角感知损失,将不同视角下的特征全都投影到一个统一的超平面中来获得 Identity-level 和 Viewpoint-level 的分布信息,同时改进了传统视角分类中的硬样本标签,采用软标签解决视角标签聚类中的模糊问题。在一个未知配置的单镜头训练(Single camera training, SCT)环境中,Zhang 等^[98]对 Re-ID 进行了研究,假设每个行人只出现在一台摄像机中,这样有利于低成本的数据收集和注释,从而简化了 Re-ID 在全新的环境中进行训练,并且提出了新的多摄像机负损失函数(Multi-camera negative

loss, MCNL), 该方法显著提高了 SCT 设置环境中 Re-ID 的识别准确率。

文献[13] 重点关注端到端学习的识别效率, 提出了一种级联微调方法, 依次训练检测和分类度量模型, 并设计了复杂的权重调整方案, 将检测置信度合并到相似度得分中。实验结果表明, 同时优化行人检测器和数据集是改进端到端学习框架的必要条件, 否则 Two-stage 框架表现出来的性能可能比端到端框架的性能更好。原因在于行人检测的本质是区分行人和非行人的二值分类器, 而行人再识别重点在行人身份的匹配或者验证。因此, 端到端行人再识别需要统一合适的模型和数据, 才能表现出优异的性能^[2,13]。

6.2 低空航拍行人再识别

作为安防领域空中的一只“眼睛”, 工业无人机正在凭借其巡查范围大、滞空时间长、高效率、低风险、低投入、灵活性高、对复杂环境适应性强等特征, 成为空地一体安防体系建设的关键一环, 并特别适用于执行时敏任务, 为地面指挥提供第一手的准确情报, 成为近几年安防行业又一热门研究领域, 低空航拍图像下的行人再识别逐步成为研究热点。

Zhang 等^[8] 利用卷积特征图的子空间池化奇异值分解表示输入的行人图像, 有效地学习了低空航拍影像中的行人再识别特征描述子, 通过端到端策略进行模型训练。为了方便低空航拍行人再识别的研究, 他们采集了 39 461 幅低空航拍图片, 包含有 1 581 个行人 ID 的大型低空航拍行人再识别数据集, 由两架大疆消费型无人机在距离地面 20~60 m 的低空拍摄, 被命名为 PRAI-1581。VisDrone (Vision meets drone: A challenge) 挑战赛^[29] 的任务是从无人机获取的视觉数据中进行物体检测、跟踪和计数, VisDrone 数据集由天津大学机器学习与数据挖掘实验室的 AISKYEYE 团队收集。

比较地面监控体系下的行人再识别, 无人机低空航拍图像中的目标所占像素小, 目标通常稀疏且分布不均匀, 使其很难与周围背景分开, 检测效率很低。受高度、角度和光照等干扰因素影响, 无人机低空航拍图像还存在低分辨率、外观特征弱、相机运动模糊和目标长期遮挡等必须解决的问题。

综上, 计算机视觉的相关研究能为智能安防和智慧城市的兴起构建了一个智能“天眼”系统, 具备重大安防的监测与防御能力、重大安全事故的预警与救援能力, 以及自然灾害的有效监测能力, 为构建立体化、智能化监测预警与综合应对平台提供有力支撑。然而, 开放式行人再识别图像中的目标所占像素小, 目标通常稀疏且分布不均匀, 使其很难与周围背景分开, 检测效率很低。受高度、角度和光照等干扰因素影响, 摄像机视频或者图像还存在低分辨率、外观特征弱、相机运动模糊和目标长期遮挡等必须解决的问题。因此, 构建速度快、精度高、鲁棒性强的开放式行人再识别系统成为近几年安防行业又一热门研究领域。

7 存在的问题与发展方向

7.1 存在的问题

开放式行人再识别主要应用在社会综合治理、新型犯罪侦查、自然灾害救助和预防恐怖袭击等构筑国家立体化、信息化社会治安防控体系上。当前, 已有很多学者参与了开放式行人再识别方法的研究, 本文总结和归纳了 Leng 等^[2] 和 Ye 等^[99] 等的研究工作, 推导出开放式行人再识别存在着以下几个亟待解决的问题。

(1) 训练高泛化能力的模型, 用以适应广义行人再识别问题。在实际应用中, Re-ID 数据可能是从多个不同模态中获得, 查询行人和候选行人库的数据可能处于不同的模态, 不同模态数据包含可见光、热红外、深度图像或文本描述, 行人图像还存在较大的分辨率差异, 现有的跨数据集训练方法大多采用域自适应技术。高泛化能力的 Re-ID 模型应当能自动处理分辨率变化、不同模态、多种环境和多数据域的情况, 这就导致了广义 Re-ID 更具挑战性。未来此方向的研究工作有望促进行人再识别在实际场景

中的应用。

(2)利用最小化行人标注策略构建大规模数据集问题。目前已存在的行人再识别数据集规模小,限制了深度学习模型性能的发挥,但构建大规模数据集需要大量的时间资源和人力资源,如何利用最小化行人标注策略构建大规模数据集,仍然是行人再识别在实际场景中应用的一大挑战。除了无监督学习、弱监督学习、数据生成和小样本学习已经有较多的研究成果以外,利用最小化行人标注策略的还有:一是主动学习或人机交互相结合的方法为最小化行人标注提供了另一种可能的解决方案,以减轻对行人标注的依赖,甚至可以训练一个具有人工回圈监督的 Re-ID 网络;二是虚拟数据学习为最小化人工标注提供了一种替代方法,但是如何弥补合成图像和真实数据集之间的差距仍然是一个问题。

(3)行人再识别系统的可扩展性问题。随着行人再识别领域和真实场景中大规模摄像机网络的数据量规模显著增长,未来应该尝试着解决行人再识别的可扩展性问题。为了快速检索行人,哈希算法被广泛用于提高检索速度,近似于最近邻检索。解决可扩展性问题的另一个方向是设计一个轻量级的行人再识别模型,但是现有开放式行人再识别模型的特征维度高,模型精度的提升是以高昂的计算成本为代价,严重影响了开放式行人再识别技术的实际应用。根据硬件配置自适应地调整行人再识别系统也为解决可扩展性问题提供了解决方案。

(4)自适应动态摄像机网络和环境变化因素问题。在实际应用中,新摄像机可以临时插入现有的监控系统中,这称为动态网络摄像机。开放式行人再识别系统集成成了密集人群和社会关系中的多种环境约束,摄像机模型的自适应和环境因素在实际的动态摄像机网络中至关重要。在一个动态更新的摄像机网络中,很可能包含大量更换服饰的同一行人,未来进一步探讨解决行人服饰变化的问题很有意义。

(5)设计特定域的行人再识别体系结构。现有的深度行人再识别方法通常采用图像分类设计的体系结构作为特征提取的主干网络,随着神经网络结构自动搜索技术的发展,设计特定域的行人再识别体系结构也是亟待解决的问题。譬如基于全尺度特征学习可以在一定范围内挖掘具有判别性的行人特征,以及非常轻量级的行人再识别系统可以获得具有竞争力的识别性能。

7.2 发展方向

在人工智能与大数据技术迅速发展的今天,对适应真实环境的开放式行人再识别技术的需求越来越显著,行人再识别系统获得了更为广阔的研究前景。关于未来的研究方向应在技术创新与新应用方向两方面进行探索,论文建议性地给出了如下几个方面的解决思路。

(1)构建开放式环境下的完备数据库。开放式行人再识别受诸多外界因素的影响,如光照、天气、遮挡、昼夜和摄像机视角等,需要构建真实环境下的完备数据库,承接各种算法在统一的数据库上进行性能测试。可使用生成对抗网络(Generative adversarial network, GAN)等数据增强技术生成更多数据,辅助以数据驱动的深度深度学习。可借助行人检测算法,探索开放式行人再识别数据集的自动标注算法。利用热红外成像监控摄像头,充实夜间或者黑暗场景下的数据集。深度图像在极低的光照和颜色条件下可以保持更多的不变性,能提供行人的体型和骨骼等重要信息。

(2)设计高效简洁的行人再识别特征描述子。近年来,深度学习对行人再识别和图像分类等领域做出了巨大贡献,其超强的特征学习和分类能力引起了众多学者的关注,具有一定的研究价值和应用价值。基于深度学习的识别模型几乎达到了行人再识别基准的最佳精度,但是识别模型设计复杂,特征维度高和计算代价大。为了平衡模型的效率和准确率,设计高效简洁的行人再识别特征描述子就显得尤为重要。

(3)引用先进的机器学习模型。机器学习属于方法论,开放式行人再识别属于应用研究,解决开放式行人再识别的问题需要大量用到机器学习中的很多方法,引用先进的机器学习模型有利于解决开放

式行人再识别面临的问题。半监督、无监督、小样本学习和迁移学习的方法有利于解决现有数据集体量小、大规模数据集标注困难和昂贵等问题。在恶劣光照环境下,跨模态学习有利于增强RGB行人图像与热红外成像之间的跨模态匹配。

(4)搭建适应真实环境的实时应用系统。当前,大部分的行人再识别研究方法是基于已经检测出行人的先验条件下进行的,离真实场景下的实际应用还有一段距离。为开展大都市视频监控网络中图像大数据下行人再识别的实际场景应用研究,借助大数据平台的大容量存储能力和快速并发的计算能力,可搭建起以行人检测、跟踪和检索三位一体开放式的行人再识别应用系统。

(5)探索基于自然语言检索的行人再识别研究。通过对行人的语言描述来检索到指定行人,如寻找白色衬衫的男性,这个研究方向被称作文本-图像行人再识别,或者称为基于自然语言检索的行人再识别研究。在缺乏行人(或者嫌疑人)图像的情况下,目击者对行人的描述是自然语言语句,跨模态文本-图像行人再识别在开放式真实场景中具有很强的实用性,值得研究者进一步探索和研究。

8 结束语

本文重点总结了开放式行人再识别技术的研究方法,介绍了开放式行人再识别的发展历程、与封闭式行人再识别的异同、开放式行人再识别建模过程和数据集的比较分析,最后对开放式行人再识别技术的研究难点和未来发展进行了较为详细的分析和探讨。未来的开放式行人再识别技术应当朝着效率更高、适用性更强的方向进行发展。

参考文献:

- [1] ZHENG Liang, YANG Yi, HAUPTMANN A G. Person Re-identification: Past, present and future[EB/OL]. (2016-10-10). <https://arxiv.org/pdf/1610.02984.pdf>.
- [2] LENG Qingming, YE Mang, TIAN Qi. A survey of open-world person re-identification[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 1(2): 1-18.
- [3] 罗浩,姜伟,范星,等.基于深度学习的行人重识别研究进展[J].*自动化学报*, 2019, 45(11): 2032-2049.
LUO Hao, JIANG Wei, FAN Xing, et al. A survey on deep learning based person re-identification[J]. *Acta Automatica Sinica*, 2019, 45(11): 2032-2049.
- [4] 赵翎娟,郭方,夏道勋.行人再识别技术研究综述[J].*贵州师范大学学报(自然科学版)*, 2019, 37(6): 114-122.
ZHAO Lingjun, GUO Fang, XIA Daoxun. Survey on person re-identification technology[J]. *Journal of Guizhou Normal University (Natural Sciences)*, 2019, 37(6): 114-122.
- [5] 李双群.渐进式行人重识别关键技术研究[D].北京:北京邮电大学,2019.
LI Shuangqun. Research on key technologies of progressive person re-identification[D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [6] 刘晓凯.智能监控系统中行人重识别方法研究[D].大连:大连理工大学,2017.
LIU Xiaokai. Research on pedestrian re-identification in intelligent monitoring system[D]. Dalian: Dalian University of Technology, 2017.
- [7] ARDESHIR S, BORJI A. Egocentric meets top-view[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(6): 1353-1366.
- [8] ZHANG Shizhou, ZHANG Qi, YANG Yifei, et al. Person re-identification in aerial imagery[EB/OL]. (2019-08-09). <https://arxiv.org/pdf/1908.05024.pdf>.
- [9] GRIGOREV A, TIAN Zhihong, RHO S, et al. Deep person re-identification in UAV images[J]. *Journal on Advances in Signal Processing*, 2019, 19(54): 1-10.
- [10] ZHANG Xuan, LUO Hao, FAN Xing, et al. AlignedReID: Surpassing human-level performance in person re-identification [EB/OL]. (2017-11-12). <https://arxiv.org/pdf/1711.08184.pdf>.
- [11] 杨婉香,严严,陈思,等.基于多尺度生成对抗网络的遮挡行人重识别方法[J].*软件学报*, 2020, 31(7): 1-16.

- YANG Wanxiang, YAN Yan, CHEN Si, et al. Multi-scale generative adversarial network for person re-identification under occlusion[J]. *Journal of Software*, 2020, 31(7): 1-16.
- [12] LAYNE R, HOSPEDALES T M, GONG Shaogang. Investigating open-world person reidentification using a drone[C]// *Proceedings of the 13th European Conference on Computer Vision*. Zurich, Switzerland: Springer, 2014: 225-240.
- [13] KARANAM S, GOU Mengran, WU Ziyang, et al. A comprehensive evaluation and benchmark for person re-identification: Features, metrics, and datasets[EB/OL]. (2016-02-03). <https://arxiv.org/pdf/1605.09653.pdf>, 2016.
- [14] NGUYEN D T, HONG H G, KIM K W, et al. Person recognition system based on a combination of body images from visible light and thermal cameras[J]. *Sensors*, 2017, 17(3): 605-634.
- [15] WU Ancong, ZHENG Weishi, YU Hongxing, et al. RGB-infrared cross-modality person re-identification[C]// *Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE, 2017: 5390-5399.
- [16] BARBOSA I B, CRISTANI M, BUE A D, et al. Re-identification with RGB-D sensors[C]// *Proceedings of the 12th European Conference on Computer Vision*. Florence, Italy: Springer, 2012: 433-442.
- [17] HAQUE A, ALAHI A, LI Feifei. Recurrent attention models for depth-based person identification[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, Nevada: IEEE, 2016: 1229-1238.
- [18] LI Shuang, XIAO Tong, LI Hongsheng, et al. Identity-aware textual-visual matching with latent co-attention[C]// *Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE, 2017: 1890-1899.
- [19] LI Wei, ZHAO Rui, XIAO Tong, et al. DeepReID: Deep filter pairing neural network for person re-identification[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, Ohio: IEEE, 2014: 152-159.
- [20] ZHENG Liang, SHEN Liyue, TIAN Lu, et al. Scalable person re-identification: a benchmark[C]// *Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015: 1116-1124.
- [21] 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 European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 17-35.
- [22] WANG Taiqing, GONG Shaogang, ZHU Xiayan, et al. Person re-identification by video ranking[C]// *Proceedings of the 13th European Conference on Computer Vision*. Zurich, Switzerland: Springer, 2014: 688-703.
- [23] YE Mang, LAN Xiangyuan, LI Jiawei, et al. Hierarchical discriminative learning for visible thermal person re-identification [C]// *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. New Orleans, USA: AAAI, 2018: 7501-7508.
- [24] WANG Zheng, HU Ruimin, JIANG Junjun, et al. Scale-adaptive low-resolution person re-identification via learning a discriminating surface[C]// *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. New York, USA: Morgan Kaufmann, 2016: 2669-2675.
- [25] ZHENG Liang, ZHANG Hengheng, SUN Shaoyan, et al. Person re-identification in the wild[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii: IEEE, 2017: 3346-3355.
- [26] XIAO Tong, LI Shuang, WANG Bochao, et al. Joint detection and identification feature learning for person search[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii: IEEE, 2017: 3376-3385.
- [27] WEI Longhui, ZHANG Shiliang, GAO Wen, 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, Utah: IEEE, 2018: 79-88.
- [28] ZHENG W S, GONG S, XIANG T. Towards open-world person re-identification by one-shot groupbased verification[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(3): 591-606.
- [29] ZHU Pengfei, WEN Longyin, BIAN Xiao, et al. Vision meets drones: A challenge[EB/OL]. (2018-01-23). <https://arxiv.org/pdf/1804.07437.pdf>.
- [30] SINGH A, PATIL D, OMKAR S N. Eye in the sky: Real-time drone surveillance system (DSS) for violent individuals identification using scatternet hybrid deep learning network[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, Utah: IEEE, 2018: 1710-1718.
- [31] YI Z, ZHANG H, TAN P, et al. DualGAN: Unsupervised dual learning for image-to-image translation[C]// *Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE, 2017: 2868-2876.
- [32] ZHONG Zhun, ZHENG Liang, ZHENG Zhedong, et al. Camera style adaptation for person re-identification[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, Utah: IEEE, 2018: 5157-5166.

- [33] DENG Weijian, ZHENG Liang, YE Qixiang, et al. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person reidentification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah: IEEE, 2018: 994-1003.
- [34] ZHAO R, OUYANG W, WANG X. Learning mid-level filters for person re-identification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Columbus, Ohio: IEEE, 2014: 144-151.
- [35] ZHENG Zhedong, YANG Xiaodong, YU Zhiding, et al. Joint discriminative and generative learning for person re-identification [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, USA: IEEE, 2019: 2133-2142.
- [36] ZHOU Kaiyang, YANG Yongxin, ANDREA C, et al. Omni-scale feature learning for person re-identification[C]// Proceedings of 2019 IEEE International Conference on Computer Vision. Seoul, South Korea: IEEE, 2019: 133-142.
- [37] LV Jianming, CHEN Weihang, LI Qing, et al. Unsupervised cross-dataset person reidentification by transfer learning of spatial-temporal patterns[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah: IEEE, 2018: 7948-7956.
- [38] YU Hongxing, ZHENG Weishi, WU Ancong, et al. Unsupervised person re-identification by soft multilabel learning[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, USA: IEEE, 2019: 2143-2152.
- [39] ZHENG Liang, BIE Zhi, SUN Yifan, et al. MARS: A video benchmark for large-scale person re-identification[C]// Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer, 2016: 868-884.
- [40] WU Ancong, ZHENG Weishi, YU Hongxing, et al. RGB-infrared cross-modality person re-identification[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 5390-5399.
- [41] DAI Pingyang, JI Rongrong, WANG Haibin, et al. Cross-modality person re-identification with generative adversarial training [C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. Stockholm, Sweden: Morgan Kaufmann, 2018: 677-683.
- [42] KAI J, ARENS M. Local feature based person reidentification in infrared image sequences[C]//Proceedings of the Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance. Boston, USA: IEEE, 2010: 448-455.
- [43] YE Mang, WANG Zheng, LAN Xiangyuan, et al. Visible thermal person re-identification via dual-constrained top-ranking [C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. Stockholm, Sweden: Morgan Kaufmann, 2018: 1092-1099.
- [44] LU Yan, WU Yue, LIU Bin, et al. Cross-modality person re-identification with shared-specific feature transfer[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2020: 13379-13389.
- [45] LI D, WEI X, HONG X, et al. Infrared-visible cross-modal person re-identification with an x modality[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 4610-4617.
- [46] WANG G, ZHANG T, YANG Y, et al. Cross-modality paired-images generation for rgb-infrared person re-identification[C]// Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 12144-12151.
- [47] MOGELMOSE A, BAHNSEN C, MOESLUND T B, et al. Person re-identification with RGB, depth and thermal features [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops. Portland, USA: IEEE, 2013: 301-307.
- [48] MUNARO M, BASSO A, FOSSATI A, et al. 3D reconstruction of freely moving persons for re-identification with a depth sensor[C]//Proceedings of 2014 IEEE International Conference on Robotics and Automation. Hong Kong, China: IEEE, 2014: 4512-4519.
- [49] WU Ancong, ZHENG Weishi, LAI Jianhuang. Robust depth-based person re-identification[J]. IEEE Trans Image Process, 2017, 26(6): 2588-2603.
- [50] HAFNER F M, BHUIYAN A, KOUIJ J F P, et al. RGB-depth cross-modal person re-identification[C]//Proceedings of the 16th IEEE International Conference on Advanced Video and Signal Based Surveillance. Taipei, China: IEEE, 2019: 1-8.
- [51] 姜国权,肖祺祺,霍占强. 基于RGB与灰度信息融合的行人再识别方法[J]. 计算机工程, 2020. DOI: 10.19678/j.issn.1000-3428.00.57115.
- JIANG Guoquan, XIAO Zhenzhen, HUO Zhanqiang. Person re-identification based on the fusion of RGB and grayscale information[J]. Computer Engineering, 2020. DOI: 10.19678/j.issn.1000-3428.00.57115.

- [52] LI Shuang, XIAO Tong, LI Hongsheng, et al. Person search with natural language description[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Hawaii: IEEE, 2017: 5187-5196.
- [53] YIN Zhou, ZHENG Weishi, WU Ancong, et al. Adversarial attribute-image person re-identification[C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. Stockholm, Sweden: Morgan Kaufmann, 2018: 1100-1106.
- [54] YE Mang, CHAO Liang, ZHENG Wang, et al. Specific person retrieval via incomplete text description[C]//Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. Shanghai, China: ACM, 2015: 547-550.
- [55] LI Shuang, XIAO Tong, LI Hongsheng, et al. Identity-aware textual-visual matching with latent co-attention[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 1908-1917.
- [56] CHEN D, LI H, LIU X. Improving deep visual representation for person re-identification by global and local image-language association[C]//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer, 2018: 56-73.
- [57] LI Xiang, ZHENG Weishi, WANG Xiaojuan, et al. Multi-scale learning for low-resolution person re-identification[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 3765-3773.
- [58] JIANG Junjun, XIANG Ma, CHEN Chen, et al. Single image super-resolution via locally regularized anchored neighborhood regression and nonlocal means[J]. IEEE Transactions on Multimedia, 2016, 19(1): 15-26.
- [59] WANG Zheng, YE Mang, YANG Fan, et al. Cascaded SR-gan for scale-adaptive low-resolution person re-identification[C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. Stockholm, Sweden: Morgan Kaufmann, 2018: 3891-3897.
- [60] JING X Y, ZHU X, WU F, et al. Super-resolution person re-identification with semi-coupled low-rank discriminant dictionary learning[J]. IEEE Transactions on Image Processing, 2017, 26(3): 1363-1378.
- [61] JIAO Jiening, ZHENG Weishi, WU Ancong, et al. Deep low-resolution person reidentification[C]//Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans, USA: AAAI, 2018: 6967-6974.
- [62] GUO Tongtong, FENG Zhanxiang, CHEN Zeyu, et al. Biometric recognition[M]. New York: Springer, 2019: 447-454.
- [63] 沈庆, 田畅, 王家宝, 等. 多分辨率特征注意力融合行人再识别[J]. 中国图象图形学报, 2020, 25(5): 946-955.
- SHEN Qing, TIAN Chang, WANG Jiabao, et al. Multi-resolution feature attention fusion method for person re-identification [J]. Journal of Image and Graphics, 2020, 25(5): 946-955.
- [64] ZHENG Feng, JI Rongrong, DENG Cheng, et al. Pyramidal person RE-identification via multi-loss dynamic training[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, USA: IEEE, 2019: 8506-8514.
- [65] ZHANG R, LIN L, ZHANG R, et al. Bit-scalable deep hashing with regularized similarity learning for image retrieval and person re-identification[J]. IEEE Transactions on Image Processing, 2015, 24(12): 4766-4779.
- [66] 戴巨超, 王洪元, 倪彤光, 等. 基于深度卷积生成对抗网络和拓展近邻重排序的行人重识别[J]. 计算机研究与发展, 2019, 56(8): 1632-1641.
- DAI Chenchao, WANG Hongyuan, NI Tongguang, et al. Person re-identification based on deep convolutional generative adversarial network and expanded neighbor reranking[J]. Journal of Computer Research and Development, 2019, 56(8): 1632-1641.
- [67] 余晨阳, 温林凤, 杨钢, 等. 结合 BiLSTM 和注意力机制的视频行人再识别[J]. 中国图象图形学报, 2019, 24(10): 1703-1710.
- YU Chenyang, WEN Linfeng, YANG Gang, et al. Video person reidentification based on BiLSTM and attention mechanism [J]. Journal of Image and Graphics, 2019, 24(10): 1703-1710.
- [68] ZHU F, KONG X, ZHENG L, et al. Part-based deep hashing for large-scale person re-identification[J]. IEEE Transactions on Image Processing, 2017, 26(10): 4806-4817.
- [69] WU Lin, WANG Yang, GE Zongyuan, et al. Structured deep hashing with convolutional neural networks for fast person re-identification[J]. Computer Vision and Image Understanding, 2018, 167(1): 63-73.
- [70] YE M, MA A J, ZHENG L, et al. Dynamic label graph matching for unsupervised video re-identification[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 5152-5160.
- [71] YE M, LAN X Y, YUEN P C. Robust anchor embedding for unsupervised video person re-identification in the wild[C]//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer, 2018: 170-186.
- [72] YE M, LI J W, MA A J, et al. Dynamic graph co-matching for unsupervised video-based person re-identification[J]. IEEE

- Transactions on Image Processing, 2019, 28(6): 2976-2990.
- [73] YU Hongxing, WU Ancong, ZHENG Weishi. Cross-view asymmetric metric learning for unsupervised person re-identification [C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 994-1002.
- [74] MA A J, LI J W, YUEN P C, et al. Cross-domain person reidentification using domain adaptation ranking SVMs[J]. IEEE Transactions on Image Processing, 2015, 24(5): 1599-1613.
- [75] WANG Xiaojuan, ZHENG Weishi, LI Xiang, et al. Cross-scenario transfer person reidentification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(8): 1447-1460.
- [76] PENG Peixi, XIANG Tao, WANG Yaowei, et al. Unsupervised cross-dataset transfer learning for person re-identification [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, Nevada: IEEE, 2016: 1306-1315.
- [77] CHEN H, WANG Y, SHI Y, et al. Deep transfer learning for person re-identification[C]//Proceedings of the Fourth IEEE International Conference on Multimedia Big Data. Xi'an, China: IEEE, 2018: 1-5.
- [78] LI Yujhe, YANG Fuen, LIU Yencheng, et al. Adaptation and re-identification network: An unsupervised deep transfer learning approach to person re-identification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah: IEEE, 2018: 172-178.
- [79] WANG Jingya, ZHU Xiatian, GONG Shaogang, et al. Transferable joint attribute-identity deep learning for unsupervised person re-identification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah: IEEE, 2018: 2275-2284.
- [80] ZHANG X, CAO J, SHEN C, et al. Self-training with progressive augmentation for unsupervised cross-domain person re-identification[C]//Proceedings of 2019 IEEE International Conference on Computer Vision. Seoul, South Korea: IEEE, 2019: 8221-8230.
- [81] LIU J, ZHA Z, CHEN D, et al. Adaptive transfer network for cross-domain person re-identification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, USA: IEEE, 2019: 7195-7204.
- [82] SONG J, YANG Y, SONG Y, et al. Generalizable person re-identification by domain-invariant mapping network[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, USA: IEEE, 2019: 719-728.
- [83] YANG F, LI K, ZHONG Z, et al. Asymmetric co-teaching for unsupervised cross domain person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 12597-12604.
- [84] HUANG Y, PENG P, JIN Y, et al. Domain adaptive attention learning for unsupervised person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 11069-11076.
- [85] PARK H, HAM B. Relation network for person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 11839-11847.
- [86] JIN X, LAN C, ZENG W, et al. Semantics-aligned representation learning for person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 11173-11180.
- [87] YU Z, JIN Z, WEI L, et al. Progressive transfer learning for person re-identification[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. Macao, China: Morgan Kaufmann, 2019: 4220-4226.
- [88] 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 Tong, CHEN Xin, TANG Xiao, et al. Dual stream convnet-gan for person re-identification[J]. Information Technology and Network Security, 2020, 39(6): 7-12.
- [89] 种衍杰, 方琰, 沙涛. 基于特征融合网络的行人重识别[J]. 计算机系统应用, 2019, 28(1): 127-133.
- CHONG Yanjie, FANG Yan, SHA Tao. Person re-identification by feature fusion network[J]. Computer Systems & Applications, 2019, 28(1): 127-133.
- [90] ZHEN Zhou, YAN Huang, WEI Wang, et al. See the forest for the trees: Joint spatial and temporal recurrent neural networks for video-based person re-identification[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Hawaii: IEEE, 2017: 6776-6785.
- [91] XU Yuanlu, MA Bingpeng, RUI Huang, et al. Person search in a scene by jointly modeling people commonness and person uniqueness[C]//Proceedings of the 22nd ACM International Conference on Multimedia. Orlando, USA: ACM, 2014: 937-940.

- [92] ALEMU L T, SHAH M, PELILLO M. Deep constrained dominant sets for person re-identification[C]//Proceedings of 2019 IEEE International Conference on Computer Vision. Seoul, South Korea: IEEE, 2019: 9854-9863.
- [93] LI Shuzhao, YU Huimin, HU Haoji. Appearance and motion enhancement for video-based person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 1134-1140.
- [94] WU Guile, ZHU Xiatian, GONG Shaogang. Tracklet self-supervised learning for unsupervised person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 12362-12369.
- [95] CHEN Zengqun, ZHOU Zhiheng, HUANG Junchu, et al. Frame-guided region-aligned representation for video person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 10591-10598.
- [96] LI X, ZHOU W, ZHOU Y, et al. Relation-guided spatial attention and temporal refinement for video-based person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 11434-11441.
- [97] ZHU Z, JIANG X, ZHENG F, et al. Viewpoint-aware loss with angular regularization for person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 13114-13121.
- [98] ZHANG T, XIE L, WEI L, et al. Single camera training for person re-identification[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 12878-12885.
- [99] YE Mang, SHEN Jianbing, LIN Gaojie, et al. Deep learning for person re-identification: A survey and outlook [EB/OL]. (2020-01-04). <https://arxiv.org/pdf/2001.04193.pdf>.

作者简介:



夏道勋(1980-),男,博士,副教授,硕士生导师,研究方向:计算机视觉、大数据技术,E-mail: dxxia@gznu.edu.cn。



郭方(1995-),女,硕士研究生,研究方向:计算机视觉,E-mail: gchunji@163.com。



刘浩杰(1997-),男,硕士研究生,研究方向:计算机视觉,E-mail: 1327094020@qq.com。



夏勇(1978-),通信作者,男,博士,教授,博士生导师,研究方向:数字图像处理,E-mail: yxia@nwpu.edu.cn。

(编辑:夏道家)