

前 言

自深度学习的概念提出以来，人工智能经历了从理论研究到技术落地的多次飞跃。在此过程中，深度学习凭借强大的数据处理能力和特征学习能力，推动了计算机视觉、自然语言处理、语音识别等诸多领域的蓬勃发展。作为计算机视觉领域的重要方向之一，行人重识别技术逐渐受到学术界和工业界的广泛关注。行人重识别技术在智能安防中能够提升公共安全保障能力，在智慧交通中能够优化通行效率，在智慧城市的精细化管理中也发挥着关键作用。这项技术的重要性不仅体现在学术研究中，更在实际应用中展现出巨大的潜力。

近年来，随着深度学习技术的迅猛发展，行人重识别技术在精度和效率方面都取得了显著提高。然而，该领域仍面临诸多挑战，例如，复杂环境下的光照变化、行人姿态差异、数据标签不足等问题严重制约了技术的应用落地。面对这些难点，研究者提出了许多创新性的解决方案，逐步推动行人重识别技术走向实际应用。但这些研究成果多见于论文或技术报告中，缺乏系统的梳理和全面的总结，尤其是在行人重识别核心问题的深入解析和实际应用场景的具体分析方面。为了进一步推动行人重识别技术的研究与应用，帮助更多的研究人员、开发者及高等院校的学生掌握这一领域的核心理论与方法，作者撰写了本书，以为读者提供一条清晰的学习路径，全面解读行人重识别领域的核心挑战、最新进展和未来方向。希望本书不仅是对现有研究现状的总结，更是激发读者深入探索行人重识别技术的起点。

本书系统梳理行人重识别领域的基础理论、关键技术和实际应用，内容涵盖从基础到前沿的多层次主题，包括行人重识别基础理论、行人姿态转换技术、基于图像和视频的行人重识别技术、无监督多源域适应行人重识别技术以及未来的发展趋势。作者在书中不仅分析了领域内的研究现状，还探讨了多种具有代表性的经典方法，展现了行人重识别技术在安防、智能交通等领域中的应用前景。

本书力求兼顾理论深度与实践价值，适合不同背景 and 需求的读者。对于高校人工智能、计算机科学等相关专业的本科生和研究生，本书通过清晰的讲解和丰富的案例，帮助他们快速掌握行人重识别领域的核心知识与方法；对于计算机视觉和行人重识别领域的研究人员和开发者，本书系统地梳理了代表性的理论成果和实践经验，可为其深入研究与技术开发提供参考。希望本书的出版能够启发更多人关注并应用行人重识别技术，为这一领域的持续发展注入新的动力。

本书的完成凝聚了集体的智慧与辛勤付出，是作者研究团队多年研究成果的总

结。其中，王科俊统筹全书，刘美辰为本书的撰写和修改做出主要贡献，张小凤、朱明超参与了部分写作和校稿工作，吕春艳、阎涛和王昊霖等对本书的成果也作出了重要贡献。在此对所有为本书撰写和出版作出贡献的人员致以诚挚的感谢。同时，作者也深知本书的内容离不开前人的研究成果和实践经验，对在行人重识别领域中提出理论方法和应用成果的研究者和实践者表示由衷的感谢。

限于作者水平，本书难免存在不妥之处，敬请广大读者批评指正。

王科俊

2026年2月

目 录

前言

第 1 章 行人重识别基础理论	1
1.1 行人重识别定义	1
1.2 研究价值与应用场景	2
1.2.1 安防和刑侦	2
1.2.2 搜寻救援	3
1.2.3 智能家居养老	3
1.2.4 智能商业	4
1.3 常见数据集和评价指标	5
1.3.1 常见数据集	5
1.3.2 评价指标	10
1.4 本章小结	11
参考文献	12
第 2 章 行人姿态转换技术	13
2.1 行人姿态转换的定义	13
2.2 行人姿态转换与行人重识别的关系	14
2.3 行人姿态转换方法概述	14
2.3.1 评价指标	14
2.3.2 行人姿态转换相关方法	16
2.4 基于分割掩模引导的行人姿态转换方法	21
2.4.1 方法概述	21
2.4.2 网络结构与关键技术	22
2.5 基于语义解析引导的行人姿态转换方法	28
2.5.1 基于语义解析注意力网络的行人姿态转换方法	28
2.5.2 基于语义注射网络的行人姿态转换方法	36
2.6 方法性能分析	45
2.6.1 实验设置	45
2.6.2 性能评估与分析	45
2.7 在行人重识别中的应用	49
2.7.1 数据增强模型	50

2.7.2	实验设置	50
2.7.3	性能评估与分析	51
2.8	本章小结	53
	参考文献	53
第3章	基于图像的行人重识别技术	59
3.1	面临的挑战与难点	59
3.1.1	姿态差异	59
3.1.2	空间错位问题	60
3.1.3	遮挡问题	60
3.2	基于图像的行人重识别方法概述	61
3.2.1	基于局部特征提取的行人重识别方法	61
3.2.2	基于注意力网络的行人重识别方法	62
3.2.3	基于姿态估计或语义分割的行人重识别方法	63
3.2.4	基于生成模型的行人重识别方法	64
3.3	基于对抗思想消除姿态差异的行人重识别方法	65
3.3.1	方法概述	66
3.3.2	网络结构与关键性技术	67
3.4	针对空间错位问题的多分支注意力行人重识别方法	72
3.4.1	方法概述	72
3.4.2	网络结构与关键性技术	73
3.5	针对遮挡问题的行人重识别方法	81
3.5.1	基于姿态对齐网络和信息交互的自遮挡行人重识别方法	81
3.5.2	基于语义解析和信息传递网络的遮挡行人重识别方法	90
3.6	多分支姿态估计和语义解析网络的空间错位和遮挡行人重识别方法	95
3.6.1	方法概述	95
3.6.2	网络结构与关键性技术	96
3.7	方法性能分析	102
3.7.1	实验设置	102
3.7.2	性能评估与分析	103
3.8	本章小结	108
	参考文献	108
第4章	基于视频的行人重识别技术	115
4.1	面临的挑战与难点	115
4.2	基于视频的行人重识别方法概述	116
4.2.1	视频行人表征学习研究及发展现状	116

4.2.2 基于异常帧的视频行人重识别方法	118
4.3 基于时空注意力融合的视频行人重识别方法	120
4.3.1 方法概述	121
4.3.2 网络结构与关键技术	121
4.4 基于运动注意力与远程时间聚合的视频行人重识别方法	127
4.4.1 方法概述	127
4.4.2 网络结构与关键技术	127
4.5 基于图卷积与图自适应池化的视频行人重识别方法	132
4.5.1 方法概述	133
4.5.2 网络结构与关键技术	133
4.6 方法性能分析	139
4.6.1 实验设置	139
4.6.2 性能评估与分析	139
4.7 本章小结	141
参考文献	141
第5章 无监督多源域适应行人重识别技术	146
5.1 面临的挑战与难点	146
5.1.1 多源域间的分布差异问题	146
5.1.2 伪标签中的噪声问题	147
5.1.3 计算成本的高昂问题	148
5.2 无监督行人重识别方法概述	149
5.3 基于多源域特征融合的行人重识别方法	152
5.3.1 方法概述	152
5.3.2 网络结构与关键技术	153
5.4 基于耐噪学习的行人重识别方法	157
5.4.1 方法概述	158
5.4.2 网络结构与关键技术	158
5.5 基于混合架构与剪枝的模型轻量化行人重识别方法	163
5.5.1 方法概述	163
5.5.2 网络结构与关键技术	164
5.6 方法性能分析	170
5.6.1 实验设置	170
5.6.2 性能评估与分析	170
5.7 本章小结	172
参考文献	173

第 6 章 未来发展与探索	176
6.1 未来发展趋势	176
6.1.1 方法优化与性能提升	176
6.1.2 多模态融合	177
6.1.3 边缘计算与云计算	178
6.1.4 隐私保护与安全性	178
6.2 未来研究方向	179
6.2.1 空间错位与遮挡	179
6.2.2 跨视角与跨域识别	180
6.2.3 隐私保护与数据安全	181
6.2.4 轻量化与嵌入式应用	182
6.2.5 多任务学习与场景理解	183

第1章 行人重识别基础理论

随着智能监控系统的不断发展，如何在复杂的场景下高效、精准地识别并追踪特定目标，成为公共安全领域中的一个重要课题。为了应对这一挑战，行人重识别 (person re-identification, person Re-ID) 技术逐渐兴起。行人重识别^[1]是一种跨摄像头环境下对特定行人身份进行匹配和追踪的技术，能够在海量监控数据中快速定位目标，为智能安防和公共管理提供关键支持。

那么，什么是行人重识别？行人重识别技术是如何发展起来的？它又能在哪些场景中发挥作用呢？本章主要介绍行人重识别的基本概念、研究价值，以及实际应用场景，为深入了解行人重识别技术奠定基础。

1.1 行人重识别定义

随着现代科技的飞速发展，计算机视觉技术^[2]应运而生。该技术通过计算机来模拟人类的视觉过程，具备人类感知能力及视觉功能，进而对目标进行检测、分类、分割和跟踪等。当今社会，监控系统对于摄像头数量和质量的需求不断提高，计算机视觉技术不仅在人们的日常生活中得到了广泛的应用，在公共安全领域也发挥着重要的作用。在大多数的计算机视觉任务中，检测或跟踪的目标对象是人，例如，搜寻失踪人口或者追踪可疑人物，需要对人员进行有效识别。由于摄像头像素以及拍摄视角的局限性，无法从监控视频中获得高清晰度、高像素的人脸图像，特别是在交通运输和城市规划等实际场景下，行人图像可能不含有脸特征，此时人脸识别技术较为受限，难以应用于复杂条件下的安防、刑侦及搜救等领域。行人重识别^[1]作为一种计算机视觉技术，是智能视频分析领域中继人脸识别技术后兴起的一种新方法，属于复杂场景下跨摄像头跟踪^[3]的范畴，能够很好地适用于各类监控场景，有效识别并准确定位出监控视频中的特定行人。

作为跨摄像头的行人检索技术，行人重识别旨在检索非重叠视角摄像头下特定身份的行人，研究对象是行人的各类特征，包括服装、发型、体型、姿态等，可作为人脸识别的补充技术。具体来讲，在视频监控系统中首先指定一张待查询的行人图像，利用行人重识别技术跨设备搜索该行人，判断此人是否出现在不同时刻无重叠区域中的不同摄像头画面中^[4]，即查询图像在行人候选集中进行匹配，获得该行人在不同摄像头下出现的时间和位置。因此，该技术可以跨越时间和空间对目标行人进行跟踪、匹配与身份鉴定，在人的背面、侧面或人脸不清晰的情

况下,实现目标行人的识别和检索,弥补人脸识别技术的不足。行人重识别任务具体示意图见图 1.1。摄像头 A、B、C、D 为不同视角下的摄像头,对于任何一个摄像头拍摄的行人图像,需要在其他摄像头中匹配到同一身份标识(ID)的行人,即为行人重识别。

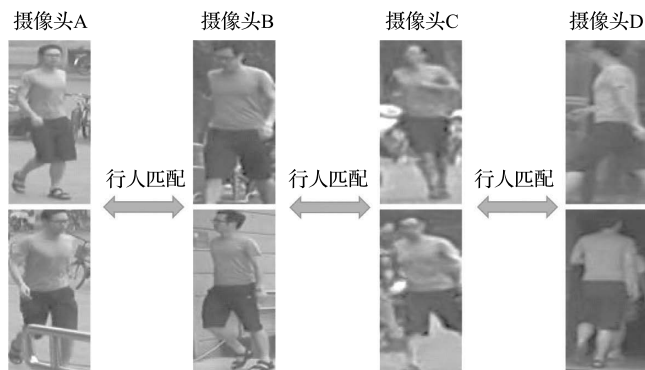


图 1.1 行人重识别任务示意图

1.2 研究价值与应用场景

随着城市化进程的加速和社会治安问题的日益复杂,如何高效地识别和追踪行人已成为公共安全和社会管理中的重要课题。行人重识别技术应运而生,它利用计算机视觉和深度学习方法,通过分析和匹配图像中行人的特征,能够在复杂的环境中实现对目标行人的准确识别和追踪。近年来,行人重识别技术迅速发展,并在多个领域展现了广阔的应用前景和重要的实际意义。行人重识别不仅可以提升公共安全和刑侦工作的效率,还在安防和刑侦、搜寻救援、智能家居养老、智能商业等方面发挥着关键作用。

1.2.1 安防和刑侦

在安防和刑侦领域,行人重识别技术正逐步成为一种高效的辅助工具。警方在追踪和抓捕犯罪嫌疑人过程中,可以通过行人重识别技术快速分析海量监控视频数据,精准锁定嫌疑人。这一技术能够利用目标嫌疑人的图像信息,在不同摄像头的拍摄范围内实现目标的连续跟踪,从而帮助执法人员快速定位目标的行踪轨迹。此外,行人重识别技术能够从历史监控录像中筛选出与嫌疑人特征相符的人员信息,有效缩小排查范围。这不仅显著节省了人力物力,还极大地提高了案件侦破的效率和准确性。

行人重识别技术在复杂的公共场所,如火车站、机场、商场等高密度人群聚

集区域,也具有重要的应用价值。通过实时监控分析系统,该技术能够迅速发现潜在的嫌疑目标,甚至可以通过关联分析技术预警可能的犯罪行为。与传统的人工比对方法相比,行人重识别方法不仅减小了误差,还能在短时间内完成大规模数据的处理和分析,有助于应对复杂案件和紧急突发事件。

随着行人重识别技术的不断发展,其在安防和刑侦领域的应用场景将更加广泛。例如,结合多模态数据(如红外图像、深度图像)和多摄像头数据融合技术,系统可以在更恶劣的光照条件下,如夜间或低能见度环境中,仍然实现高精度的识别。这些进步将为现代化智能安防系统提供更强有力的技术支持,助力社会治安和刑侦工作迈向更高的智能化和精确化水平。

1.2.2 搜寻救援

在失踪或遇险人员的搜寻与救援工作中,行人重识别技术展现出了极大的潜力和价值。通过利用目标人员的影像信息,搜救人员可以快速从大规模的监控视频或现场捕获的影像中锁定目标,显著提高搜救效率,缩短救援所需时间。在突发事件或复杂环境中,如地震、山体滑坡、火灾等灾难现场,这种技术能够帮助救援团队在极短时间内定位到失踪或被困人员的位置,极大地提高救援成功率。

行人重识别技术在复杂场景下的应用尤为重要。例如,在城市环境中,利用分布广泛的摄像头网络,可以实时跟踪失踪人员的动态行踪,绘制人员移动轨迹,从而为搜救工作提供关键线索。而在偏远地区或灾难现场,结合无人机或机器人设备的行人重识别技术,可以快速扫描大面积区域,从中筛选出目标人员,克服传统人工搜寻速度慢、覆盖范围有限的问题。

此外,行人重识别技术还可以与其他智能系统相结合,如生物特征识别(如步态识别、人脸识别)和多模态数据融合(如红外成像、深度图像),进一步提高在恶劣环境下的识别能力。例如,在烟雾弥漫或光线不足的灾难现场,通过红外图像与行人重识别技术的结合,救援人员可以快速找到隐藏在障碍物后或被烟雾遮挡的人员,从而提高搜救效率,挽救更多生命。

1.2.3 智能居家养老

行人重识别技术在智能居家养老领域具有重要的应用价值,能够为独自在家的老年人提供全方位的安全保障和健康监测。通过部署行人重识别系统,可以实时跟踪老年人的活动状态,当检测到异常行为或紧急状况时,系统能够在第一时间发出警报并通知家属或医护人员,显著提升老年人的生活安全性和健康保障。例如,在日常生活中,老年人可能会因健康原因或意外摔倒而陷入危险。行人重识别系统可以通过摄像头实时监测老年人的行为,当检测到摔倒动作或异常的长时间静止状态时,系统能够自动识别并触发预警机制,迅速向家庭成员、社区医

护人员或紧急服务机构发送报警信息。这种实时响应功能，不仅为老年人提供更加安全的居家环境，还为其家人提供更多的安全感。

此外，行人重识别技术还可以结合其他智能家居设备进行综合应用。例如，通过与健康监测设备(如心率监测器、步态分析仪)相结合，系统可以提供更加全面的健康数据监测和分析。当老年人出现步态异常、行动迟缓或心率异常等健康信号时，系统能够主动提醒家属安排健康检查，降低可能出现的健康风险。行人重识别技术还可以实现生活习惯的监测与优化，例如，记录老年人的日常活动轨迹、活动频率和时长，提供数据支持以帮助其调整作息和生活习惯，从而提升生活质量。在夜间，系统可以检测老年人是否安全入睡或离开床位，降低夜间发生意外的风险。

随着技术的发展，行人重识别系统还将与语音交互、人工智能决策系统相结合，为老年人提供更贴心、更智能的服务，从日常陪伴到健康管理，全面提升居家养老的智能化水平，打造更加安全、舒适的老年生活环境。

1.2.4 智能商业

行人重识别技术在智能商业中的应用主要体现在跨摄像头的顾客身份匹配和行为分析上，帮助商家实现更高效的管理和服务。通过部署多个摄像头网络，行人重识别系统可以在大规模商场中跨区域跟踪顾客的活动，为商家提供跨区域行为数据支持。例如，行人重识别技术可以解决顾客跨区域行为的连续性问题。在商场不同楼层或不同分区的摄像头中，重识别系统能够通过匹配顾客的身份信息，将其在不同摄像头捕获到的行为数据关联起来。这一功能对分析顾客的完整购物路径和行为模式具有重要意义，可以帮助商家识别顾客对哪些区域或商品更感兴趣。

在大规模商场的跨区域防损管理中，行人重识别系统可以快速锁定跨区域的潜在目标。例如，当某区域内出现可疑行为时，系统能够自动将其轨迹延展到其他区域，从而提升整个商场的安全监控能力。这种跨摄像头的身份匹配能力使得防损管理更加高效，减少漏检和误判的可能性。行人重识别技术还能为会员系统的精准服务提供支持。通过在不同区域的摄像头中识别会员身份，商家可以为顾客提供定制化的跨区域服务，例如，在顾客进入特定区域时推送个性化推荐信息，或根据其购物路径安排导购人员提供针对性服务。通过分析不同顾客在跨区域购物中的行为模式，商家可以优化商场布局和资源分配，提升购物体验 and 运营效率。同时，基于行人重识别的多区域身份匹配功能，还可以实现多商场之间的顾客行为数据联动，为跨商场的营销策略提供支持。

此外，行人重识别技术在智慧城市和智能监控中也有着广阔的应用前景。该技术能够在人口密集地区实现行人的数字化管理，助力智慧城市建设，提升城市

管理效率和安全水平。

1.3 常见数据集和评价指标

行人重识别技术的研究离不开高质量的数据集和科学合理的评价指标。数据集的多样性和规模直接影响模型的泛化能力，评价指标则是衡量方法性能的重要依据。本节将介绍行人重识别领域中广泛使用的常见数据集和评价指标，为后续读者对方法的研究和性能评估提供支持。

1.3.1 常见数据集

目前的行人重识别数据集通常由视角非重叠的多个摄像头采集而成，每一张采集到的行人图像均对应一个行人 ID。早期的行人重识别数据集的标注方式主要为手工裁剪，近期的行人重识别数据集是通过行人检测或者行人追踪来自动获取的。相比手工裁剪的方式，自动检测能够更好地还原检测错误的实际场景。在行人重识别的数据集中，对于指定的从摄像头中获取的输入图像，称之为查询图像(query image)；对于从不同的摄像头中获取的检索图像，称之为候选图像(gallery image)。数据集中所有查询图像组成查询集(query set)，所有候选图像构成候选集(gallery set)，两者共同组成测试集(test set)。训练集(train set)是指用于训练的图像集合，主要用于训练神经网络参数，测试集用于评价训练完成的神经网络的性能。行人重识别数据集分为两类，分别是基于图像的行人重识别数据集和基于视频的行人重识别数据集。本节将详细介绍两类数据集的具体细节。

1. 基于图像的行人重识别数据集

常用的基于图像的大规模行人重识别数据集包括 Market-1501^[5]、DukeMTMC-reID^[6]、CUHK03^[7]、MSMT17^[8] 和 Occluded-Duke^[9]等。其中，Market-1501、DukeMTMC-reID、CUHK03 属于通用数据集，Occluded-Duke 是专门针对遮挡场景设计的数据集。表 1.1 汇总了这些数据集的具体信息。

Market-1501 数据集是一个大型的公共行人重识别数据集，在清华大学校园夏天拍摄采集，于 2015 年公开发布，涵盖行人姿态变化、光照变化和遮挡等现象的图像，与实际场景相似。图像来自 6 个非重叠视角的摄像头，其中包括 5 个高分辨率摄像头和 1 个低分辨率摄像头，行人矩形框由检测器自动检测和剪裁。数据集共包含 1501 个行人身份的 32668 张图像，训练集由来自 751 个不同身份的 12936 张图像组成，平均每个人有 17.2 张训练图像。测试集由候选集和查询集组成。候选集由来自 750 个身份的 19732 张图像组成，平均每个人有 26.3 张测试图像。

查询集的行人图像均由手动标注，从 6 个摄像头中为测试集中每个身份的行人选取一张图像组成查询集，共包含 3368 张图像，所以每个身份的行人在查询集中至多有 6 张图像。训练集和测试集的行人身份不重复。图 1.2 展示了 Market-1501 数据集中视角变化、姿势变化和遮挡等因素的图像。

表 1.1 基于图像的行人重识别数据集

特征	Market-1501	DukeMTMC-reID	CUHK03 (detected)	MSMT17	Occluded-Duke
拍摄地点	清华大学	杜克大学	香港中文大学	北京大学	杜克大学
发布年份	2015	2017	2014	2018	2019
行人身份数量	1501	1404	1467	4101	1221
摄像头数量	6	8	10	15	2
行人图像总数	32668	36411	14097	126441	35489
训练集行人身份数量	751	702	767	1041	702
查询集行人身份数量	750	702	700	3060	519
候选集行人身份数量	750	1110	700	3060	1110
训练集图像数量	12936	16522	7365	32621	15618
查询集图像数量	3368	2228	1400	11659	2210
候选集图像数量	19732	17661	5332	82161	17661



图 1.2 Market-1501 数据集示例

DukeMTMC-reID 数据集是 DukeMTMC 数据集的子集, 采集于杜克大学, 由 8 个非重叠视角的高分辨率摄像头拍摄, 于 2017 年公开发布。DukeMTMC-reID 数据集是由 DukeMTMC 数据集的视频中, 以每秒 120 帧帧率采集的图像构成, 提供人工标注的行人边界框。DukeMTMC-reID 数据集共有来自 1404 个行人身份的 36411 张图像, 从 702 个行人身份中选取 16522 张图像作为训练集, 剩余 702 个行人身份的 2228 个查询图像作为查询集。候选集共有 17661 张图像。此外, DukeMTMC-reID 数据集包含很多被各种物体遮挡的图像。图 1.3 展示了 DukeMTMC-reID 数据集中视角变化、姿势变化和遮挡等因素的图像。



图 1.3 DukeMTMC-reID 数据集示例

CUHK03 数据集是一个广泛使用的大型行人重识别数据集, 由香港中文大学校园的两组五对不同视角的摄像头拍摄采集得到, 于 2014 年公开发布。该数据集提供了人工标注和检测器自动获取的行人边界框两种形式, 并且含有新旧两种测试协议。在旧协议下, 数据集包含 1360 个行人身份, 其中训练集 1160 个行人身份, 验证集 100 个行人身份, 测试集 100 个行人身份, 为 single-shot 设置, 每个行人在测试集里只有一张图像, 称作训练/测试协议。在新协议下, 数据集共包含 1467 个行人身份的 14097 张图像。其中训练集有 767 个行人身份的 7365 张图像, 测试集包含 700 个行人身份的 6732 张图像。在测试阶段, 对每一个测试身份, 分别在两组摄像头视角中各选取一张作为查询图像, 共 1400 张图像。其余图像构成候选集, 共 5332 张图像。该方式为 multi-shot 设置, 每个行人在候选集中有多张

图像。本书所有实验均采用自动检测的行人边界框,即 CUHK03 (detected) 数据集,其更接近实际场景。图 1.4 展示了 CUHK03 (detected) 数据集中视角变化、姿势变化和遮挡等因素的图像。



图 1.4 CUHK03 (detected) 数据集示例

MSMT17 数据集是目前行人重识别中最大的基于图像的数据集之一,共包含 126441 张图像和 4101 个行人身份。图像分别在北京大学校园的早上、中午和下午拍摄,并且在四种不同的天气条件背景下拍摄,除此之外,为了保护被采集者的隐私信息,MSMT17 数据集对行人的面部进行了遮挡操作,这也使在该数据集上训练得到的方法能够更加关注于人脸之外的信息。该数据集通常被认为是捕捉了最复杂的多个不同场景,因此相当具备挑战性。

Occluded-Duke 数据集全称是 Occluded-DukeMTMC 数据集,是从 DukeMTMC-reID 数据集内筛选而来,专门针对遮挡问题,是目前最大并且唯一含有训练集的遮挡行人重识别数据集。训练集有 702 个行人身份的 15618 张图像。查询集包含 519 个行人身份的 2210 张遮挡图像,相比 DukeMTMC-reID 数据集减少 183 个行人身份。候选集包含 1110 个行人身份的 17661 张图像,与 DukeMTMC-reID 数据集一致。Occluded-Duke 数据集中遮挡图像示例如图 1.5 所示,该数据集中存在检测错误、光照变化、视角变化和遮挡等情况。

2. 基于视频的行人重识别数据集

基于视频的数据集包含多个连续的帧,这些帧构成了一个视频序列。本书研



图 1.5 Occluded-Duke 数据集示例

究的是基于视频的行人重识别方法，常用的数据集有 MARS^[10]和 DukeMTMC-VideoReID^[11]。这些数据集包含了大量的视频序列，每个序列包含了多个摄像头视角下的行人图像。两个数据集的具体信息如表 1.2 所示。

表 1.2 视频行人重识别主流数据集

数据集	发布年份	ID 数	序列数	标签方式	评价指标
MARS	2016	1261	20478	DPM+GMMCP	CMC+mAP
DukeMTMC-VideoReID	2018	1812	4832	手工	CMC+mAP

MARS 数据集是 Market-1501 的视频扩展数据集。该数据集共包含 1261 个行人身份、20478 个有效行人视频序列，按照标准划分，训练集包含 625 个行人身份的 8298 个视频序列，测试集包含 636 个行人身份的 12180 个视频序列。因为 MARS 数据集的行人轨迹序列由可变形部件模型(deformable part model, DPM)检测器和基于高斯混合模型的变化点(Gaussian mixture model based change point, GMMCP)检测跟踪器生成，所以行人视频序列中还包含 3248 个由检测跟踪错误而产生的干扰视频序列，进一步增加了方法识别的难度。此外，MARS 数据集中包含 6 个不同摄像头视角拍摄的视频，其中 4 个摄像头拍摄场景为室外街景，2 个摄像头拍摄场景为室内商场，涵盖了多种复杂环境，如低分辨率、光照变化、模糊和部分遮挡等。

DukeMTMC-VideoReID 数据集是 DukeMTMC-reID 数据集的衍生版本，主要用于评估视频行人重识别方法的性能。该数据集是在架设 8 个摄像头有噪声背景的室外场景中捕获的，并且受到光照、姿势和视点变化和遮挡的影响。由于此数据集是手工标注的，每个身份在摄像机下只有一个视频序列。研究人员以每秒 12 帧的帧率从视频中裁剪行人图像来生成一个轨迹。数据集按照协议进行分割，即 702 个行人身份用于训练，702 个行人身份用于测试，408 个行人身份作为干扰。该数据集总共生成了 369656 帧的 2196 个视频序列用于训练，445764 帧的 2636 个视频序列用于测试和干扰。

1.3.2 评价指标

评价指标用于衡量不同方法在视频行人重识别任务上的性能表现,帮助研究者评估方法的优劣并指导方法的改进。常见的视频行人重识别评价指标主要有累计匹配特征(cumulative matching characteristics, CMC)曲线和平均精度均值(mean average precision, mAP)。

CMC 曲线是在计算机视觉和模式识别领域中广泛使用的一种性能评估曲线,即在给定的数据集上,根据不同的匹配阈值,以及每个匹配阈值下的匹配概率,计算出模型在不同检索列表长度下的累计匹配率。具体对于视频行人重识别而言,针对检索库(query)中的某个待检索的视频(probe),返回搜索库(gallery)的一系列按相似度由高到低排序的结果,若某一查询对应的真实匹配目标出现在前 k 个检索结果中,则记为一次命中。所有查询中命中的数目占总查询数的比例称为 Rank- k 准确率。以 k 为横坐标、Rank- k 准确率为纵坐标,即可绘制出 CMC 曲线,具体的数学表达式为

$$\text{CMC}_k = \frac{1}{n} \sum_{i=1}^n r(i, k) \quad (1.1)$$

式中, n 为测试集中查询样本的数量; $r(i, k)$ 为第 i 个查询样本的真实匹配目标是否出现在前 k 个检索结果中,若出现,则 $r(i, k)$ 等于 1, 否则等于 0。

CMC 指标计算的具体示例见图 1.6。左侧的蓝色方块表示 query 中 5 个不同身份的行人,右侧的红色和绿色方块则表示 gallery 返回的结果,这些结果按照相似度从高到低排序。红色方块表示匹配错误,绿色方块表示匹配正确。在第一列中,只有第四个查询目标匹配正确,因此 Rank-1 准确率为正确匹配数除以检索样本总数,即 $1/5=0.2$ 。在排名前五的检索样本中,第一个、第二个和第四个查询目标均正确匹配,因此 Rank-5 准确率为 $3/5=0.6$ 。同样,在排名前十的检索样本中,所有查询目标均包含正确匹配样本,因此 Rank-10 准确率为 $5/5=1$ 。

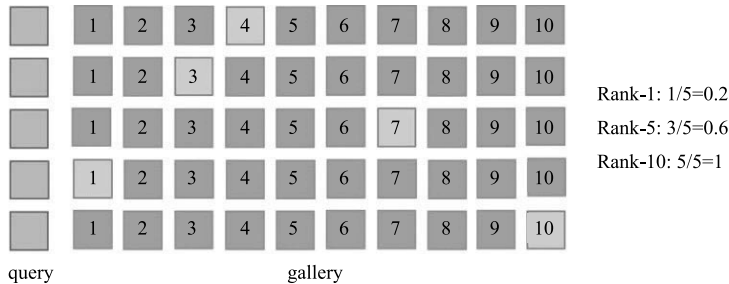


图 1.6 CMC 指标计算示例