

# 行人重识别技术与 YOLOv7 的结合应用研究

姬晨杰<sup>1</sup> 陈祥毅<sup>2</sup> 乔昭鑫<sup>1</sup>

(1 太原科技大学 山西晋中 030024 2 福建省邮电学校 福建福州 350008)

**摘要:** 随着社会经济蓬勃发展,人身安全和财产安全备受关注。安防视频监控系统应运而生,但监控数据激增和目标筛选难题使人工处理不再足够。智能视频监控系统基于计算机视觉技术,采用图像处理和模式识别等技术对视频数据进行筛选、检测和分析。其中,行人重识别是核心功能,旨在准确跨监控镜头检测同一行人。

**关键词:** 行人重识别; 行人检测; 深度学习; YOLOv7 算法

2020 年我国视频监控设备市场规模从 2015 年的 553.5 亿元增长至 982.8 亿元,预计 2023 年我国视频监控设备市场规模有望达到 1294.8 亿元,拥有数百万个监控摄像头,构建了全球最大的监控系统网络和基础平台,为智慧城市建设提供了重要技术支持。然而,随着监控系统的不断发展,其性能和功能也受到了新技术的不断影响。本文的研究工作主要集中在行人重识别领域。随着社会对安全性的不断强调和效率需求的提高,行人重识别技术面临更多挑战。行人重识别技术在不同领域的需求日益多样化,因此对于行人重识别任务的复杂性也不断增加。该技术在自动驾驶、视频监控等领域有着广泛的应用,特别是在多目标检测和跟踪中具有重要作用。因此,本文的研究侧重于行人重识别技术,包括行人目标的识别方法和行人目标之间的关联策略。

## 1. 国内外研究现状

国内外对于目标跟踪技术的研究已经取得了显著进展,这一领域备受专家学者的关注。目标跟踪技术具有广泛的应用前景,尤其在智能监控系统中,已经引起了广泛的研究兴趣。世界各国纷纷推出了智能监控系统,以提高对目标的检测和跟踪能力。美国国防高级研究计划局(DARPA)于 1997 年启动了 VSAM 项目,该项目旨在实现在已知环境下对目标人员或车辆的检测和跟踪。马里兰大学在 2000 年推出了 W4 监控系统,该系统使用投影直方图来区分个体和群体,并能够通过动态外观模型来跟踪人物目标。此外,法国的多家科研机构也开发了智能交通监控系统,用于监测交通状况并通过数据分析来进行交通管理。美国在 2005 年推出了 S3 智能监控系统,该系统通过分析来自分布式摄像头的视频信息,在发现可疑行为时能够自动发出警报。中国也在 2010 年研发了 VS-Star 系统,该系统可以检测和跟踪监控视频中的目标,并通过分析目标行为来实现异常警报。

### 1.1 经典目标跟踪算法

经典目标跟踪算法通常可以分为两大类,即生成式和判别式。生成式方法侧重于利用历史帧的信息来对目

标特征进行统计建模,以减少目标丢失的情况。然而,当面对复杂的背景环境时,生成式方法的效果可能不尽如人意。其中,一种典型的生成式方法是基于概率分布的均值漂移方法。它通过建模目标、初始化跟踪目标的外部矩形框,然后利用梯度上升来搜索目标,计算特征值概率,最终得到均值漂移向量,通过迭代计算概率分布函数的收敛。尽管这种方法在目标与背景的颜色差异较大时具有较快的收敛速度,但在目标与背景颜色差异较小的情况下效果较差。

### 1.2 深度学习行人检测方法

近年来,深度学习方法在行人检测领域取得了显著进展,广泛应用于计算机视觉任务,如物体分类、特征匹配、场景识别和行为分析等。深度学习以其卓越的性能效果成为计算机视觉领域的焦点之一,特别是卷积神经网络(CNN),被认为是深度学习中最经典的模型之一。深度学习在行人检测领域的发展可以追溯到 2012 年,当时 Krizhevsky 等人提出了 Alexnet 卷积神经网络,在 ImageNet 大规模视觉识别挑战中取得了优异的分类成绩。此后,研究者提出了一系列深度学习方法,包括 Conv-Net、多阶段深度学习模型、JointDeep 等,这些方法利用深度神经网络解决了复杂环境中行人遮挡和形变问题<sup>[1]</sup>。

## 2. 基于改进 YOLOv7 的行人检测算法

### 2.1. YOLOv7 的技术理论

YOLOv7 是一个单级实时目标检测器,是基于 YOLOv5 代码进行修改的,是用于计算机视觉任务的最快,最准确的实时对象检测算法模型。YOLO 的基础版本有三个,分别是 YOLOv7, YOLOv7-tiny, YOLOv7-w6。YOLOv7 是针对普通 GPU 计算优化的基础模型。其中 YOLOv7-tiny 是针对边缘 GPU 优化的基础模型, YOLOv7-w6 是针对云 GPU 计算优化的基础模型。现在改进的 YOLOv7 网络使用高效聚合网络 ELAN,考虑内存访问的成本和 GPU 计算效率,提供了更强的网络架构和更准确的目标检测,也提供了更有效的特征集成方法。

YOLOv7 网络采用了模块级重参数化和采用辅助头检测以及动态标签分配策略,从而保证精确度更高,可以解决模型随深度降低带来的性能退化,正样本差等问题。

## 2.2. YOLOv7 技术理论在行人检测上的应用

YOLO 算法是进行行人检测的一个非常重要的算法,它可以精确通过图片或者视频中的人物进行检测,确定出目标人物,从而得到目标人物的活动轨迹,但是仍然有一定的限制<sup>[2]</sup>。现在改进 YOLOv7 的技术的应用则更加的精确的检测人物,它能够在保证高识别准确率的同时有着非常快的检测速度,从而大大的缩短检测时间提高效率<sup>[3]</sup>。

## 3. 重识别算法

### 3.1. 重识别的解释

重识别是重点利用现在的计算机视觉技术对图像或者视频序列中判断是否存在特定的目标行人的技术,通常是指通过目标人物的穿着,体态,发型,行为等在不同的场景中能够再次识别出目标人物,并以此在图像中描绘出目标人物的行进轨迹的 AI 视觉技术<sup>[4]</sup>,总的来说,就是利用计算机视觉技术检测出不同摄像头,图片或者视频下的同一个目标人物,它基于图像“行人重识别”本质是图像的描述与匹配,特征提取和度量学习,提取目标人物的特征,这些特征“包括衣服颜色,纹理,边缘,形状或者是鞋子的款式等表现特征”。在行人检测的基础上进一步的识别出目标人物的特征细节,并得到详细的运动轨迹。

### 3.2. 特征提取

在图像表示中保持几何不变性对于图像检索非常重要,因为从监控视频中经过行人检测出的行人图像有些可能被缩放、旋转以及行人自身也有大量姿态变化。为获取更鲁棒的行人特征,我们提出了 S-CNN 模型,融合了 SIFT 和 CNN 特征。考虑到行人重识别任务的特殊性,我们提出了多任务学习策略,包括分类任务和检索任务。我们还对现有的分类和检索损失函数进行深入研究,并验证了损失函数组合的性能,以提高行人重识别的准确性。

### 3.3. S-CNN 模型

在这一节中,我们详细介绍了我们提出的 S-CNN 模型,该模型结合了 SIFT 和 CNN 特征,旨在提取更鲁棒的行人特征以应对不同场景和变换。SIFT 特征在处理图像尺度、旋转等变换时表现出色,而 CNN 在学习更高级别的抽象特征方面具有优势。我们将这两种特征相互融合,旨在构建一个更具鲁棒性的模型,能够适应行人在

不同场景中的多样性。融合方法:我们通过将 SIFT 和 CNN 特征连接或融合到网络的不同层中,使得模型能够同时学习到局部和全局的图像特征。这种融合方法有助于提高模型的表征能力,使其更好地捕捉到行人的身份特征。

## 3.4. 实验结果与分析

在这一节中,我们将详细讨论对我们提出的算法进行的广泛实验,以及实验结果的分析。这些实验旨在验证算法的有效性、准确性以及在不同数据集上的鲁棒性。

泛化能力:我们分析了算法的泛化能力,即在未见过的场景或数据集上的性能表现。结果表明,我们的算法在泛化性能方面表现良好,具备适应不同环境的能力。对关键因素的鲁棒性:我们分析了算法对关键因素(如遮挡、光照变化等)的鲁棒性。通过这些分析,我们可以识别出算法在特定场景下的优势和改进的空间,为进一步优化提供指导。通过对实验结果的全面分析,我们验证了算法在不同条件下的鲁棒性和性能优越性。这为我们的算法在实际行人重识别应用中的推广奠定了坚实的基础。

## 4. 结论

社会安全日益重要,智能视频监控系统应运而生。为解决监控数据庞大和目标筛选难题,采用深度学习方法,融合行人检测和重识别技术。行人检测方面,选择 YOLOv7 模型,并提出多尺度扩展预测的改进算法,增强对小尺度行人目标的敏感性。行人重识别方面,引入 S-CNN 模型,结合 SIFT 和 CNN 特征,采用多任务学习策略,并深入研究损失函数组合以提高准确性。最终,将行人检测和重识别融为一体,实现端到端的行人重识别,适应实际监控场景,提高社会安全管理水平。

## 参考文献:

- [1]Ouyang W , Wang X. Joint Deep Learning for Pedestrian Detection[C].Sydney , Australia: IEEE International Conference on Computer Vision, 2013
- [2]曹伊宁,李超,彭雅坤.基于改进 YOLOv7 的夜间行人检测算法[J].长江信息通信,2022,35(10):57-60.
- [3]郑佳卉.基于 YOLOv3 的行人视频目标检测方法[D].西安电子科技大学,2019.
- [4]邱青.基于FPGA的行人检测YOLO算法轻量化研究与实现[D].华南理工大学,2021.

作者简介:姬晨杰(2003.03-),男,汉族,山西晋中人,本科,研究方向:软件工程专业。