



# 基于计算机视觉的行人检测方法研究进展

姜翔飞<sup>1a</sup>, 吕文涛<sup>1a</sup>, 叶冬<sup>2</sup>, 郭庆<sup>3</sup>, 鲁竞<sup>3</sup>, 陈影柔<sup>1b</sup>

(1. 浙江理工大学, a. 信息科学与工程学院; b. 浙江省智能织物与柔性互联重点实验室, 杭州 310018;

2. 浙江移动信息系统集成有限公司, 杭州 311217; 3. 浙江省技术创新服务中心, 杭州 310007)

**摘要:** 基于计算机视觉的行人检测方法可有效提高行人检测效率, 已广泛应用于智慧城市、辅助驾驶等场景。文章对行人检测涉及的图像分割、特征提取、机器学习和分类与定位等方法进行了归纳, 综述了各种方法的主要思想、适用性和局限性; 同时介绍了行人检测算法的评价指标, 对算法性能进行了分析; 最后总结了行人检测方法的研究进展, 并对未来的发展方向进行了展望。计算机视觉作为目标检测中的一项重要技术, 在行人检测领域仍有待发展, 算法结构改进、分类器优化、复杂场景下的行人检测等是未来的研究重点。

**关键词:** 计算机视觉; 行人检测; 图像分割; 特征提取; 机器学习; 分类与定位

**中图分类号:** TP391.4; TP183

**文献标志码:** A

**文章编号:** 1673-3851(2023)05-0318-13

**引文格式:** 姜翔飞, 吕文涛, 叶冬, 等. 基于计算机视觉的行人检测方法研究进展[J]. 浙江理工大学学报(自然科学), 2023, 49(3): 318-330.

**Reference Format:** LOU Xiangfei, LÜ Wentao, YE Dong, et al. Research progress of pedestrian detection methods based on computer vision[J]. Journal of Zhejiang Sci-Tech University, 2023, 49(3): 318-330.

## Research progress of pedestrian detection methods based on computer vision

LOU Xiangfei<sup>1a</sup>, LÜ Wentao<sup>1a</sup>, YE Dong<sup>2</sup>, GUO Qing<sup>3</sup>, LU Jing<sup>3</sup>, CHEN Yingrou<sup>1b</sup>

(1a. School of Information Science and Engineering; 1b. Key Laboratory of Intelligent Textile and Flexible Interconnection of Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou 310018, China; 2. Zhejiang Mobile Information System Integration Co., Ltd., Hangzhou 311217, China; 3. Zhejiang Technology Innovation Service Center, Hangzhou 310007, China)

**Abstract:** Pedestrian detection methods based on computer vision can effectively improve the efficiency of pedestrian detection, and have been widely used in smart cities, assisted driving and other scenes. In this paper, the methods of image segmentation, feature extraction, machine learning, classification and location involved in pedestrian detection are summarized, and the main ideas, applicability and limitations of each method are summarized. At the same time, the evaluation index of pedestrian detection algorithm is introduced, and the algorithm performance is analyzed. Finally, the research progress of pedestrian detection methods is summarized, and the future development direction is prospected. As an important technology in object detection, computer vision still needs to be developed in the field of pedestrian detection. The improvement of algorithm structure, classifier optimization, and pedestrian detection in

收稿日期: 2022-11-18 网络出版日期: 2023-03-01

基金项目: 国家自然科学基金项目(61601410); 浙江省科技厅重点研发计划项目(2021C01047, 2022C01079); 东北大学流程工业综合自动化国家重点实验室联合基金项目(2021-KF-21-03, 2021-KF-21-06)

作者简介: 姜翔飞(1998—), 男, 浙江金华人, 硕士研究生, 主要从事机器学习、图像处理方面的研究。

通信作者: 吕文涛, E-mail: alvinlwt@zstu.edu.cn

complex scenes are the focus of future research.

**Key words:** computer vision; pedestrian detection; image segmentation; feature extraction; machine learning; classification and position

## 0 引言

行人检测技术通过计算机视觉技术来检测图像序列中是否有行人目标,同时对行人目标进行定位。目前,行人检测技术不仅在智慧城市、辅助驾驶等场景中得到了广泛应用,而且为行人跌倒检测、人体行为预测等研究提供理论基础和技术支撑<sup>[1]</sup>。基于计算机视觉的行人检测技术在无人驾驶、辅助驾驶中起到了决定性的作用,该技术可在行车过程中通过成像设备获取车辆周围的图像数据,给出图像序列中行人目标的准确定位,提高了辅助驾驶的安全性,保障了驾驶员及行人的生命安全。在实际应用中,各类行人检测系统需要高实时性、强鲁棒性<sup>[2-4]</sup>,这是行人检测技术的难点,也使其成为目标检测领域的研究热点。行人检测与一般的目标检测存在着较大差异,一般的目标检测方法并不完全适用于行人目标<sup>[5]</sup>,主要原因有以下几点:

a)检测图像中存在目标重叠、相互遮挡等情况。行人检测中通常会出现人与人的遮挡、人与物的遮挡。目前的行人检测方法已能处理局部遮挡问题,但对于大面积、较严重的遮挡问题仍待进一步研究<sup>[6]</sup>。

b)行人检测受背景影响较大。现实背景下光照变化、类似行人轮廓物体的干扰等都会造成行人目标与背景的混淆、难以区分,从而出现漏检、误检,使得准确识别、精确定位行人目标变得非常困难<sup>[7]</sup>。

c)小目标行人定位准确度较低。远距离成像的行人通常目标较小、分辨率较低,缺乏充分的特征信

息,易受噪声干扰,所以检测算法难以精确定位小目标行人。

d)行人多姿态、非刚性导致定位准确度低。行人检测区别于一般的目标检测,不同行人具有不同的姿态和特征,由此带来的不确定性会影响行人目标定位的准确度。

随着图像处理技术的发展,近年来越来越多的研究人员提出了基于图像分割、特征提取、机器学习等行人检测方法来解决以上问题,这也让行人检测技术得到了长足的进步<sup>[8]</sup>。本文综述了近年来典型的行人检测方法,对其主要思想、适用性和局限性进行了总结。首先,阐述了基于图像分割、特征提取、机器学习、分类与定位的行人检测方法;其次,介绍了行人检测算法的评价指标,同时对算法性能进行了分析;最后,总结了行人检测方法的研究进展,并对未来的发展方向进行了展望。

## 1 基于图像分割的行人检测方法

图像分割是将图像划分为若干个特定的区域,再将这些区域按照其特征区分为不同的类别,便于提取感兴趣目标区域(Region of interest, ROI)<sup>[9]</sup>。基于图像分割的行人检测方法计算速度快、节约硬件资源,但该方法受图像背景的影响较大,对于行人目标的检测准确率不高。本文总结了以下3种基于图像分割的行人检测方法,分别为阈值分割检测方法<sup>[10]</sup>、边缘分割检测方法<sup>[11]</sup>和语义分割检测方法<sup>[12]</sup>,该3种方法的原理和优缺点见表1。

表1 基于图像分割的行人检测方法的原理和优缺点

方法类型	原理	优点	缺点	相关文献
阈值分割检测方法	通过设定不同的特征阈值,将图像的像素点划分为若干类	运算速度快	受图像背景影响较大	文献[10,13-17]
边缘分割检测方法	搜索不同区域的边界,根据边界进行图像分割	计算量少,边缘定位准确	容易造成信息丢失,易受噪声干扰	文献[11,18-23]
语义分割检测方法	对图像每一像素点打上语义标签	稳定性高,检测精度高,鲁棒性好	计算量大,无法区分开互相遮挡的行人目标	文献[12,24-29]

### 1.1 阈值分割检测方法

阈值分割法是目前使用最广泛、最常见的图像分割方法。该方法通过设定不同的特征阈值,将ROI和背景区域进行分割<sup>[13]</sup>。唐清<sup>[14]</sup>提出了一种改进的单高斯模型阈值分割方法,该方法利用了红

外热成像技术以获得高亮像素的检测分割结果。在正常的背景环境中,该方法可以对行人目标与背景区域进行有效分割<sup>[15]</sup>。单高斯模型阈值法较为简洁实用,但有一定的局限性,易受复杂背景的影响。如采用红外图像的阈值分割法,当背景中诸如电器、

电线、汽车等接近人体温度的热源较多时,易将温度较高的背景与行人目标混淆。为了改善复杂环境下图像分割准确率较低的问题,Su等<sup>[16]</sup>提出了基于最大类间方差的红外图像阈值分割法,可将图像中的待测目标和背景进行准确区分,但该方法稳定性不佳,且计算量较大,无法兼顾检测的实时性和准确率。

Trivedi等<sup>[17]</sup>提出了一种自适应阈值分割算法,通过动态决策准则提高行人检测模型的性能。该算法对不同背景环境变化和待测目标灰度变化有较好的鲁棒性,同时兼顾了检测精度和实时性,在复杂背景环境、存在噪声干扰的行人检测中,相较其他阈值分割算法有一定的优势。

### 1.2 边缘分割检测方法

边缘分割法通过搜索不同区域的边界实现ROI与背景区域的分割。该方法受复杂背景的影响较小,对于边界特征较明显的行人检测,精确度较高。Sobel分割是一种较常见的边缘检测方法,该方法根据图像的每一像素点,计算其与四周邻点灰度的加权差<sup>[18-19]</sup>,该加权差通常在边缘处达到极值,从而实现边缘的检测。

Sobel分割仅对于每个像素的周围像素点进行计算,计算量相对较小,它对灰度渐变的图像往往检测效果较好。但是一般的Sobel分割法不能区分水平边缘和垂直边缘,无法精确检测边缘方向信息,边缘定位精度较低,易受噪声的干扰,无法达到行人检测的应用需求。因此,Ma等<sup>[20]</sup>提出了一种边缘检测算法,将图像边缘特征与经过二值化后分割的边缘特征融合,提高了边缘检测的方向敏感度,解决了易将路灯、车灯等亮度较高的背景与行人目标混淆的问题;然而该方法抗噪声能力较弱,对于检测边缘的精度稍显不足。Wu等<sup>[21]</sup>针对该问题,提出了一种基于改进Sobel的边缘检测算法。该算法先对待测图像中的每一种边缘情况进行数学建模,然后仅对连续的边缘模型进行分析,建立用于边缘方向检测的分类器模型;通过细化处理待测图像的梯度图,增强图像边缘特征,解决了Sobel分割法在边缘定位精度低且易受噪声干扰的问题。这类改进的Sobel边缘检测算法能有效改善传统Sobel分割法的缺陷,提高对于行人目标边缘检测时的抗噪声能力,使边缘定位更加精确,是一种实用性较高的行人检测算法<sup>[22]</sup>。

Shi等<sup>[23]</sup>提出了一种基于改进Prewitt的边缘检测算法,该算法利用图像中各像素周围邻域的灰度加权相关度,结合待测行人目标的特征,生成一个

自适应的动态阈值。该算法是对传统Prewitt算法的改进,具有传统Prewitt算法抗噪能力强的优点,通过自适应生成的动态阈值解决了原先边缘定位精度低的问题,过滤了伪边缘,同时提高了计算速度,可满足行人检测应用中实时、精确的需求。

### 1.3 语义分割检测方法

语义分割在复杂背景环境中的检测效果较好,鲁棒性较优。Hu等<sup>[24]</sup>提出了一种语义分割树算法,该算法可生成更多的语义区域,形成语义分割树的中间节点。语义分割树有助于识别检测ROI,可以较好地过滤掉冗余目标。该算法对所有图像类别的检测效果都较好,即使是分辨率较低的图像检测效果也比其他算法有较大提升。但该算法仍存在语义分割时相邻像素间易混淆、标签不一致的问题。

Gao等<sup>[25]</sup>针对语义分割时相邻像素间标签一致性较差、易混淆的缺点,提出了一种组合多类图像语义分割的方法,该方法将自上而下和自下而上的分割结果合并在一起,增大了图像感受野。Jiang等<sup>[26]</sup>也根据图像的上下文信息,设计了一种基于多尺度卷积神经网络特征融合的图像语义分割方法。

上述两种语义分割方法在行人检测中具有更高的检测准确率,但没有明显延长计算时间。在行人检测过程中融合不同尺度的网络模型,能够获得更多的局部特征,分割结果更加准确,模型泛化能力强。因此,该方法对于小目标行人检测效果较为突出。

CNN、R-CNN、Fast-RCNN等算法能达到不错的语义分割效果,且检测精确度高,模型鲁棒性好;但是这些算法没有充分考虑待测图像的全局语义信息,对于小目标的行人检测误差较大,而且神经网络结构往往导致模型计算量较大,在实时性上达不到应用要求<sup>[27]</sup>。Li等<sup>[28]</sup>提出了基于改进U-Net神经网络的语义分割算法,利用深度卷积、残差网络和树状网络来改进网络模型,在减少参数量的同时提取更丰富的图像信息。该算法在卷积层后加入最大池化,并利用反卷积还原图像特征,进行大量特征融合,从而获得更精确的语义信息;但该算法本质上还是基于卷积神经网络,仍然存在模型参数量大、计算速度慢等不足,在实际应用中会造成检测速度过慢且资源消耗大等问题。

姜宏宇<sup>[29]</sup>利用卷积神经网络轻量化技术,提出了一种适用于实时行人检测的轻量级网络MS-PSPNet。该网络利用金字塔池化模块提取多尺度语义信息,利用逐点分组卷积降低计算参数量,提升运算速度;针对参数量减少会导致模型精度损失的问题



题,在模型中加入注意力机制,在最后两层使用空洞卷积来增大感受野,便于获取行人图像的全局特征。该网络适用于复杂背景下的行人检测,解决了语义分割存在的相邻像素间标签一致性差、易混淆的问题。

语义分割检测方法通常具有较好的检测精度,多适用于复杂背景下的行人检测,稳定性较高。使用语义分割方法可以使提取到的行人特征更具典型,增强模型的泛化能力。

表2 基于特征提取的行人检测方法优缺点

方法类型	优点	缺点	相关文献
底层特征检测方法	计算速度快	特征单一,判别能力差,泛化能力弱	文献[30-35]
混合特征检测方法	特征描述较全面,检测准确率高	计算量增加,检测速度慢,影响实时性	文献[36-41]

## 2.1 底层特征检测方法

基于图像颜色、纹理等底层特征的行人检测方法是普遍、最常见的一种。Ma等<sup>[30]</sup>提出了基于Histogram of oriented gradients(HOG)特征的行人目标检测方法,该方法通过计算图像局部的梯度幅值、方向构成特征,并利用支持向量机(Support vector machine,SVM)进行分类预测;其在MIT行人数据库的检测准确率几乎达到100%,鲁棒性也相对较好。这类特征是目前使用最广泛的描述行人的特征之一,大多数行人检测方法都是基于HOG特征和SVM分类器的思路改进的。

然而,基于HOG特征的行人检测方法也有较多缺陷,如特征维度高导致计算速度变慢。Liu等<sup>[31]</sup>针对该问题,提出了基于HOG特征块可变性的特征,设计的算法能够自适应提取行人目标特征;同时使用积分图加快运算速度,用Boosting算法筛选出最具典型性的特征块,该算法对于计算速度的提升较为有效。Bilal等<sup>[32]</sup>提出了类似的行人检测算法,在GPU上计算HOG,可极大提升运算速度;通过结合多尺度的特征提高检测性能。上述两种方法虽然可满足检测的实时性要求,但是忽视了细节特征,不适用于小目标行人的检测。Fan等<sup>[33]</sup>利用人体的多尺度特征,提出一种基于Gabor特征和HOG特征的行人检测算法。该算法对待测图像进行小波变换以提取行人多方位多尺度的振幅谱,使用主成分分析降维,然后提取HOG特征,减少了参数量,最后融合小波变换后得到的Gabor特征与HOG特征,能够有效加快运算速度,同时提高检测精度。

局部二值模式(Local binary pattern,LBP)是一种用于图像纹理分析的算子,多应用于人脸识别、图像纹理特征分类。LBP于2008年提出,其作用是记录待测图像的每一像素点与周围像素点的对比

## 2 基于特征提取的行人检测方法

行人特征描述子的选择对模型的泛化能力有较大的影响,若提取的特征判别力较差则会导致误检率较高,无法达到辅助驾驶、智慧交通等场景的应用需求。基于特征提取的行人检测方法一般包括基于颜色、梯度或纹理等信息的底层特征检测方法以及混合特征检测方法,这两类行人检测方法优缺点见表2。

信息<sup>[34]</sup>。Peng等<sup>[34]</sup>提出了基于HOG和LBP的行人检测算法,该算法能有效减少待测图像中遮挡带来的干扰,提升检测性能。然而,LBP在低分辨率、噪声干扰较多的图像中不能有效描述局部细节特征,判别能力较弱。因此,该算法的鲁棒性较弱,难以应用于实际的无人驾驶、智慧交通等场景中。

Census transform histogram(CENTRIST)与LBP类似,通过计算每一像素点对应的Census transform(CT)值组成的特征直方图,描述图像的全局信息。对比基于HOG特征的行人检测方法,其呈现出的图像特征更为全面。林才纺<sup>[35]</sup>提出了基于CENTRIST特征的行人检测算法,将RGB图像转化为灰度图像,运用GPU并行处理技术实现缩放图像、图像滤波、CT图像构建等步骤,大大提升了检测效率。基于CENTRIST特征的行人检测算法拥有更高的检测效率和更好的抗噪性能,同时能有效处理细节特征,满足行人检测的应用要求。

## 2.2 混合特征检测方法

基于单一特征的行人检测方法一般检测精确度较低,泛化能力较弱。所以,研究人员通常进行多特征融合以提高特征判别力。Park等<sup>[36]</sup>提出的行人目标检测算法将HOG特征和描述纹理的LBP特征进行融合,通过主成分分析减少计算参数量;该算法还加入了加权部件,对待测图像中的每一目标附上权值,可有效降低遮挡或行人目标部分重合造成的影响。Watanabe等<sup>[37]</sup>提出了基于共生梯度方向直方图(Co-occurrence histograms of oriented gradients,CoHOG)的行人检测方法,该方法通过多个梯度方向的特征描述行人目标,相较于基于HOG特征的行人检测方法具有更高的准确率及召回率。

种衍文等<sup>[38]</sup>针对行人检测精度不足的问题,利

用四方向特征(Four direction features, FDF)设计了检测模型,并采用 Entropy-histograms of oriented gradients(EHOG)算子来增强对于小目标行人的敏感度,该算法在行人检测中有较优的检测准确率和鲁棒性,但其实时性较差。针对该情况, Schwartz 等<sup>[39]</sup>提出了基于边缘特征的方法,结合行人颜色、纹理信息,采用偏最小二乘(Partial least squares, PLS)降低维度,提升计算效率。但是上述算法受遮挡及背景干扰影响较大。针对该情况,储珺等<sup>[40]</sup>提出了融合语义和多层特征的行人检测算法,增强行人和背景的分度;该算法在 Caltech 和 CityPersons 数据集上漏检率较低,具有强鲁棒性,可便捷应用于其他的行人检测框架中。Dollár 等<sup>[41]</sup>考虑图像上下文信息,设计了积分通道特征,在此基础上提出了

行人检测方法。该方法的基本原理是利用积分图技术把图像的梯度直方图、积分直方图、Haar 特征等结合起来进行特征计算;该方法集成了多特征信息,同时兼顾了检测速度和准确率。

综上所述,基于不同特征的行人检测方法,其检测速度、准确率、鲁棒性各有不同,而且在不同应用场景下所选择的行人特征描述子也各有差异。

### 3 基于机器学习的行人检测方法

机器学习广泛应用于行人检测领域,基于机器学习的方法能够自动提取判别力强的行人特征,且检测准确率高,鲁棒性好。本文总结并分析了传统机器学习检测方法和深度学习检测方法,这两种方法的优缺点见表 3。

表 3 基于机器学习的行人检测方法优缺点

方法类型	优点	缺点	相关文献
传统机器学习检测方法	泛化能力强,参数量少	受噪声干扰影响大,抗干扰能力弱	文献[42-45]
深度学习检测方法	检测准确率高,稳定性好	参数量大,计算速度慢	文献[46-56]

#### 3.1 传统机器学习检测方法

传统机器学习检测方法通常都采用 Boosting 算法,主要包括自适应增强(Adaptive boosting, AdaBoost)和梯度提升决策树(Gradient boosting decision tree, GBDT)算法。如 Ma 等<sup>[42]</sup>提出了基于 Haar 特征和 AdaBoost 的行人检测方法,采用改进 AdaBoost 来选择判别力强的 Haar 特征,通过积分图像表示多尺度特征,在计算速度上得到明显提升,实时性更好;然而使用积分图像的特征表示方法会受复杂背景的干扰,鲁棒性较差。Yang 等<sup>[43]</sup>针对复杂背景下的行人检测,提出了基于 HOG 特征和 SVM 的方法,利用 GrabCut 和 Mask R-CNN 进行行人图像分割;该方法具有更高的检测精度和稳定性,但是存在过拟合的问题。针对该问题,王焱<sup>[44]</sup>运用 RPN 进行特征提取,再使用 GBDT 对模型进行训练;为了防止过拟合,在训练过程中加入了随机策略。该算法在 Caltech 数据集上的检测性能较优,能较好地检测存在部分遮挡的行人目标。

Edgelet 以局部特征的检测为基础,在分类预测时,结合各部分的局部特征得到最终的分类结果。Shridhar 等<sup>[45]</sup>用 Edgelet 来描述行人的局部轮廓特征,在此基础上设计了行人检测方法。该方法首先将人体划分为若干个区域,用 AdaBoost 算法对每一部分区域进行训练,得到一个强分类器模型;在预测时,结合各个分类器的置信度来进行目标判别。上述方法通常有较好的分类预测效果,对于遮挡较多、

行人重叠等情况下的检测鲁棒性较强。然而该方法在分类预测时计算量较大,需要耗费大量的时间进行匹配计算,实时性较差,无法满足实际的应用需求。

#### 3.2 深度学习检测方法

深度学习检测方法相较传统机器学习检测方法具有更高的准确性和稳定性。目前基于深度学习的行人检测方法大致可分为两类:一类是两步法,即先选择候选框区域再进行分类回归,如基于 R-CNN、Fast R-CNN、Faster R-CNN 等的方法;另一类是一步法,即直接进行分类回归,如基于 Single-shot detector(SSD)、YOLO 系列算法、RetinaNet 等的方法。

##### 3.2.1 两步法

在行人检测方法中,基于 R-CNN 的相关算法相较于基本的目标检测算法,其检测精度、稳定性有极大提升。Yang 等<sup>[46]</sup>提出了基于 R-CNN 的行人检测算法,通过建立人体各部位之间的从属关系,实现复杂场景下的有效检测。但是基于 R-CNN 的行人检测算法仍存在着重复计算、计算量大的缺点,因此研究人员提出了 Fast R-CNN、Faster R-CNN。如 Zhang 等<sup>[47]</sup>提出了基于 Fast R-CNN 的行人检测方法,将 K-means 聚类算法与 RPN 相结合,提升了检测效率,然而该方法提取的特征判别力不强。Cai 等<sup>[48]</sup>设计了 MS-CNN,能将浅层特征与深层特征相融合,提高提取特征的判别力;基于 MS-CNN 的行人检测方法取得了较好的检测效果,但其在背

景光照不均、行人色彩多样的情况下检测误差较大。Tian 等<sup>[49]</sup>针对该问题提出了基于自动色彩增强(Automatic color enhancement, ACE)、Faster R-CNN 的行人检测算法,利用 ACE 增强行人目标的对比度,从而提高了色彩复杂场景下的检测精度。针对遮挡情况下的行人检测,Zhang 等<sup>[50]</sup>在 Faster R-CNN 的网络结构中融入跨通道注意力机制,提高了行人目标的定位精度,降低了误检率。

### 3.2.2 一步法

Liu 等<sup>[51]</sup>提出了 SSD 算法用于目标检测,该算法直接采用了不同尺度的特征层进行分类预测,省略了 R-CNN、Faster R-CNN 等算法提取候选框区域的步骤,极大提升了检测效率,但是该算法的检测精度较低。裴伟等<sup>[52]</sup>针对该问题用表征能力更强的残差网络作为 SSD 的基础网络,利用残差学习降低模型训练难度,实验结果表明改进后的 SSD 算法与传统 SSD 算法相比,更具实时性和鲁棒性,但是改进后的算法对于存在遮挡的行人目标或小目标行人的检测召回率较低。Dong 等<sup>[53]</sup>针对地铁站等拥挤场景下的行人检测,提出了基于 SF-SSD 的检测方法。该方法基于传统 SSD 网络,通过合并低层特征图,准确定位小目标,降低了拥挤场景下行人检测的误检率和漏检率。

YOLO 系列算法也适用于行人目标的实时检测,YOLO 的网络结构一般分为 Backbone 层、Neck 层和 Head 层。基于 YOLO 系列的行人检测算法相较于其他 CNN,具有较强竞争力,但基于

YOLOv3 以前系列的检测算法对于行人目标定位的准确度不够高,且计算量过大。Zhang 等<sup>[54]</sup>对 YOLOv3 做了改进,在归一化层引入 L1 正则化,并对通道进行剪枝,减少了参数量,同时使用 CIoU 损失函数,加快网络收敛速度;该算法计算参数量少、检测效率高,可满足行人检测实时性的要求;然而对于拥挤场景及小目标行人检测的准确率仍有待提高。Guo 等<sup>[55]</sup>提出了基于改进 YOLOv5 的行人检测算法,通过放大数据集的训练网格提高了对遮挡目标的识别能力。Li 等<sup>[56]</sup>针对小目标行人检测问题,提出了一种高效的 YOLO-ACN 算法,通过添加注意力机制、CIoU 损失函数和深度可分离卷积进行了改进,增大感受野,有效提取了细节特征。

综上所述,深度学习检测方法在行人检测领域中得到了广泛应用。该方法能自动在大量的数据集中学习特征表达,相比于传统的行人检测方法,具有更优的检测准确率和模型泛化性能。然而该方法参数量大,计算速度远不及传统的目标检测方法,难以达到实时行人检测的要求。对于非刚性的行人目标以及复杂背景影响下的场景,基于深度学习的行人检测方法有待进一步高效化、精确化。

## 4 基于分类与定位的行人检测方法

基于分类与定位的行人检测方法的原理是直接判断待测图像中是否存在行人目标且进行窗口定位。目前,行人检测领域中常用的分类定位方法包括滑动窗口法、超越滑动窗口法。

表4 基于分类与定位的行人检测方法的优缺点

方法类型	优点	缺点	相关文献
滑动窗口法	方便与级联结构、积分图等技术结合,检测速度快	窗口以固定长度步进,漏检率较高	文献[57-64]
超越滑动窗口法	能有效解决遮挡的问题,提高行人目标定位精度	计算过程复杂,速度慢	文献[65-67]

### 4.1 滑动窗口法

滑动窗口法的基本原理是在待测图像中选择一个固定宽高的窗口并依序滑动,然后通过训练后的分类器模型来判别窗口内是否存在行人目标。滑动窗

口法由整体法和部位法两种方法组成。整体法可直接提取窗口内行人目标的全局特征;部位法则是将行人目标拆分为各个部位进行特征提取,再建立各个部位之间的几何关系。上述两种方法的优缺点见表5。

表5 滑动窗口法的优缺点

常见方法	优点	缺点	相关文献
整体法	便捷,检测速度快	受遮挡物干扰影响较大	文献[57-62]
部位法	一定程度上改善了遮挡物带来的影响	检测速度慢	文献[63-64]

#### 4.1.1 整体法

整体法是目前行人检测中较为主流的分类方法,其中 SVM、Boosting 是目前应用最广泛的分类方法。SVM 最早在 1964 年被提出,这是一种二元分类模型,广泛用于行人识别、文本分类等研究中。

Oren 等<sup>[57]</sup>首次将 SVM 应用于行人检测中,并由比例模板扩展得到的小波模板方法来提取行人目标特征,该方法计算效率较高。核函数普遍应用于目标检测,Lei 等<sup>[58]</sup>设计了一种以直方图交叉核向量机为基础框架的分类器,该分类器的运算速度与 SVM



相差不大,但是其对行人检测准确率提升较大。

Freund 等<sup>[59]</sup>对 Boosting 算法进行优化,在 1995 年提出了检测效率更高的 AdaBoost 算法,但是其训练耗时较长,对异常样本较为敏感,会影响强学习分类器的预测准确性。Ma 等<sup>[42]</sup>将弱学习分类器组合成级联结构,显著提高了检测效率,但该方法不适用于复杂场景下的行人检测,易将行人目标与背景混淆。Kim 等<sup>[60]</sup>提出了用于图像和视觉特征共聚类的多分类器增强(Multiple classifier boosting, MCBost)算法;该算法能最大限度地区分待测目标与背景,从而适应行人检测中多类别且背景复杂的场景,但该算法存在过拟合的问题。Jiang 等<sup>[61]</sup>提出了以 HOG 特征和 Gentle AdaBoost 为基础框架的行人检测算法,在分类器模型训练时加入样本加权调节算法,可以有效解决过拟合的问题。Lin 等<sup>[62]</sup>提出了一种多示例学习框架,通过对形变部位的建模可以改善行人多姿态、非刚性造成的部分特征无法识别的问题;基于该框架的行人检测方法能自动匹配行人的局部特征,提高了行人检测的精度。

基于整体法的行人检测在训练数据集时只需要用矩形框标注出行人区域,检测效率高,然而容易忽视细节特征,在拥挤环境、遮挡影响下的行人检测精度会大幅降低,其模型泛化能力较弱。因此该方法不适用于小目标行人的检测。

#### 4.1.2 部位法

部位法的基本原理是将行人划分为若干个不同部位,再分别对检测不同部位的分类器进行训练,建立各个部位之间的几何关系。2010 年 Felzenszwalb 等<sup>[63]</sup>提出了形变部位模型(Deformable part model, DPM),基于 DPM 的方法能改善待测图像中行人目标非刚性的问题,但是该方法计算速度慢,影响检测效率。因此, Felzenszwalb 等<sup>[64]</sup>又结合级联结构分类器的思想,提出了一种级联结构的 DPM,兼顾了检测准确率和效率。

基于部位法的行人检测重点在于如何划分行人目标的各个部位以及构造各部位的分类器模型。在基于级联结构分类器的行人检测中,各部位分类器的检测性能起到了决定性作用,其作用大于各个部位之间的几何关系。部位法通常对遮挡较多、行人重叠等场景具有较强的鲁棒性,然而各部位与整体图像的匹配需要大量时间,在行人检测实时性方面达不到应用要求。

#### 4.2 超越滑动窗口法

随着基于词袋(Bag of words, BOW)的场景分

类技术的发展,超越滑动窗口法应运而生,但是该方法在行人检测技术中应用相对较少。Lampert 等<sup>[65]</sup>提出了高效子窗口搜索法(Efficient subwindow search, ESS),这是一种能快速定位目标的方法,但计算过程较为复杂。An 等<sup>[66]</sup>对其搜索策略做出改进,提出了一种基于交替搜索的近似算法,其速度比 ESS 方法快约 900 倍,但是其对拥挤环境下的遮挡问题并没有有效解决。隐式形状模型是广义霍夫变换(Generalized hough transform, GHT)和特征袋(Bag of features, BOF)方法的结合,在行人检测中能有效解决遮挡的问题。Thi 等<sup>[67]</sup>将隐式形状模型运用到目标定位中,该模型不依赖图像预处理来区别复杂背景,提高了目标定位精度。超越滑动窗口法可改善行人检测中遮挡较多、行人多姿态的问题,但无法兼顾检测实时性,局部特征的判别力较差,且误检率也较高。

综上所述,不同的方法对不同场景的行人检测,其检测性能、检测精度以及鲁棒性存在较大差异。常见行人检测方法的类型、优缺点和相关文献见表 6。

### 5 评价指标及各算法性能比较

#### 5.1 评价指标

本文总结了目前常见的几种行人检测算法评价指标。 $n_{TP}$ ,表示预测为正样本且预测结果正确(True positive, TP)的数量; $n_{FP}$ ,表示预测为正样本且预测结果错误(False positive, FP)的数量; $n_{TN}$ ,表示预测为负样本且预测结果正确(True negative, TN)的数量; $n_{FN}$ ,表示预测结果为负样本且预测结果错误(False negative, FN)的数量。

a) 准确率(Precision):  $p$ ,表示  $n_{TP}$  占有所有被预测为正样本数量的比值,  $p = n_{TP} / (n_{TP} + n_{FP})$ 。

b) 召回率(Recall):  $r$ ,表示  $n_{TP}$  占有所有正样本数量的比值,  $r = n_{TP} / (n_{TP} + n_{FN})$ 。

c) 漏检率(Miss rate, MR):  $r_M$ ,表示未被检测出的正样本数量占有所有正样本数量的比值,  $r_M = 1 - r$ 。该指标越小,模型检测召回率越高,检测性能越好。

d) 平均误检率(False positive per image, FPPI):  $r_F$ ,表示每张图像的平均误检率。

e) 平均对数漏检率(Log-average miss rate,  $MR^{-2}$ ):  $r_{LM}$ ,该指标常用于行人检测器的性能评价,  $r_M - r_F$  曲线上在  $[0.01, 1]$  区间均匀采样 9 个  $r_M$  值,并求其平均数  $r_{LM}$ 。  $r_{LM}$  值越小,表示模型漏检率越低,性能越好。

表 6 常见行人检测方法的类型和优缺点

方法种类	方法类型	常见方法	优点	缺点	相关文献
基于图像分割的检测方法	阈值分割检测方法	单高斯模型阈值分割	适用于稳定环境	易受背景干扰	文献[14]
		最大类间方差法	抗噪能力强,适用于背景复杂环境	稳定性不高,计算量大	文献[16]
		自适应阈值分割	计算简单,鲁棒性好	—	文献[8,17]
基于图像分割的检测方法	边缘分割检测方法	Sobel 算法	计算量小,速度快	边缘定位精度低、易受噪声干扰	文献[18-22]
		Prewitt 算法	抗噪能力强	—	文献[23]
		语义分割树算法	精确度较高,鲁棒性好	相邻像素间易混淆、标签不一致	文献[24]
基于特征提取的检测方法	语义分割检测方法	U-Net	检测精度高,泛化能力强	参数量大,计算速度慢	文献[28]
		MS-PPNet	计算速度快,实时性好	—	文献[29]
		HOG	鲁棒性好	维度高,计算速度慢	文献[30-33]
基于特征提取的检测方法	底层特征检测方法	LBP	计算速度快	抗噪能力弱	文献[34,36]
		CENTRIST	能有效处理细节特征,抗噪性好	—	文献[8,35]
		Edgelet	适用于遮挡物较多的环境	计算量大	文献[45]
基于机器学习的检测方法	混合特征检测方法	—	特征描述较全面,检测准确率高	对特征的选择要求较高	文献[36-41]
		AdaBoost	灵活度高,检测精度高	对异常样本较敏感	文献[42,45,59]
		GBDT	抗噪能力强	—	文献[44]
基于机器学习的检测方法	深度学习检测方法	R-CNN	鲁棒性好,泛化能力好	检测效率低,实时性较差	文献[46-50]
		SSD	省略了获取候选框的步骤,效率高	不适用于小目标行人检测	文献[51-53]
		轻量化 YOLOv3	实时性好	拥挤场景下的精度较低	文献[54]
基于分类与定位的检测方法	滑动窗口法	YOLO-ACN	易关注细节特征,适用于小目标行人检测	计算量较大	文献[56]
		整体法	计算效率高	受遮挡干扰影响较大	文献[57-62]
		部位法	改善行人目标非刚性问题	计算量大,速度慢	文献[63-64]
基于分类与定位的检测方法	超越滑动窗口法	高效子窗口搜索法	快速定位目标	检测速度慢	文献[65]
		隐式形状模型	有效解决遮挡问题	—	文献[67]



f)平均检测精度(Average precision, AP): $r_{AP}$ ,该指标通过计算  $p$  和  $r$  来判定行人检测器的性能。构造  $p-r$  曲线,计算线下面积即可得到  $r_{AP}$ ;  $r_{AP}$  越大,模型检测精度越高,其性能越好。

g)FPS:检测速度评价指标,表示在统一的硬件环境下每秒处理的图像数量。FPS 越大,检测速度越快。

## 5.2 各算法性能比较

由于行人检测算法种类繁多,彼此采用的实验数据集和评价指标各有差异,故难以在同一维度进行各算法性能的分析对比。以下以  $r_{LM}$  为评价指标,总结了常见的行人检测算法在 Caltech 数据集上的性能,见表 7。

表 7 以  $r_{LM}$  为评价指标,常见的行人检测算法在 Caltech 数据集上的性能

行人检测算法	$r_{LM}$			
	R	HO	R+HO	A
ACF <sup>[68]</sup>	44.2	90.2	54.6	79.6
LDCF <sup>[69]</sup>	24.8	81.3	37.7	71.2
Faster-RCNN <sup>[47]</sup>	9.2	57.6	20.0	—
RPN+BF <sup>[70]</sup>	9.6	74.3	24.0	64.7
MS-CNN <sup>[48]</sup>	9.9	59.5	21.5	60.9
Faster-RCNN+ATT <sup>[50]</sup>	10.3	45.2	18.2	54.5

表 7 是常见行人检测算法在不同遮挡情况下的平均对数漏检率,其中:R 表示 Reasonable,少量遮挡及无遮挡;HO 表示 Heavily occlusion,严重遮挡;R+HO 表示综合结果;A 表示完整数据集。从表 7 可以看出,Dollár 等<sup>[68]</sup>提出的 ACF 算法和 Nam 等<sup>[69]</sup>提出的 LDCF 算法均基于手工特征、多种特征融合的方式进行行人检测,其性能相较于深度神经网络的行人检测算法仍有较大差距。Zhang 等<sup>[50]</sup>采用基于锚点框的方法,通过加入注意力机制降低漏检和误检率。此外,相比于手工特征,基于深度神经网络的行人检测算法在检测准确率上有大幅提升,但是距离高精度、高效率的行人检测还有一定的差距。

## 6 总结与展望

本文介绍了行人检测方法的研究进展,综述了基于图像分割、特征提取、机器学习、分类与定位的行人检测方法,归纳了这些方法的基本原理,分析了优缺点及其适用性。目前,行人检测方法已改善的有如下几点:

a)关于行人检测目标重叠、定位精度低的问题,基于单视角的行人检测技术很难从本质上解决该问

题。为了改善该问题,一种方式是采用无人机、雷达无线电定位等技术多视角摄像机来检测行人目标。在多视角下,行人目标更为立体,姿态和空间定位的获取更为准确。另一种方式是利用深度神经网络检测行人,增大感受野,获得多尺度行人目标特征。该方法对于图像序列中小目标行人的检测效果提升较明显,是未来行人检测领域的主要研究方向之一。

b)关于行人检测系统实际应用中面临恶劣环境的问题,如在雨天、大雾等能见度低、分辨率低的环境下对行人目标进行准确检测,可建立恶劣环境下的行人检测数据集,提高模型的泛化能力,提升检测系统的鲁棒性。

c)关于小目标行人定位准确度较低的问题,可利用热成像技术、CT 图像及高光谱技术更有效地获取行人目标信息,改善了传统的 CCD 相机仅能得到目标表面信息的不足,提高了行人空间定位的准确性。该类技术可与行人检测算法相结合,有望成为行人检测领域的热点方向。

d)关于行人多姿态、非刚性的问题,可通过构建自适应检测器,对行人目标进行更鲁棒、精确的检测。该方法对行人目标检测的准确率有较大提升,对不同背景环境变化和待测目标灰度变化有较好的鲁棒性。通过自学习提升行人检测系统的性能将是重要发展方向。

目前,行人检测已经是计算机视觉应用领域的一大研究热点,然而在实际应用中行人检测方法仍有待发展,算法结构改进、分类器优化、复杂场景下的行人检测等问题仍是未来的研究重点。根据本文对行人检测方法原理和优缺点的分析,行人检测方法还需在以下几个方面加快发展:

a)基于图像分割的行人检测方法较为简单便捷且节约硬件资源,但该方法在客流量较大、多行人遮挡的场景中,尚未改善漏检、误检等问题,因此需提升该类方法的检测准确率和召回率。

b)基于传统底层特征提取的行人检测方法相较于基于图像分割的方法,其检测效率及准确率略有提升,但仍受复杂背景或遮挡的影响,对于恶劣环境下的行人检测,其鲁棒性仍待提升,因此采用基于混合特征的行人检测方法在一定程度上可提高该类方法的性能。

c)基于机器学习的行人检测方法鲁棒性较优、适用性广,对于小目标或遮挡较严重等场景下的行人检测效果尤为突出;基于深度神经网络的方法能提取判别力较强的特征,极大地提高了模型的泛化

能力,降低了复杂背景对行人检测结果的影响;然而该类方法计算参数量大,硬件成本相对较高,如何兼顾检测精度和检测效率还需进一步研究。

d)基于分类与定位的行人检测方法中,基于滑动窗口法的行人检测通常存在检测速度慢、误检率较高等问题,如何提高其检测精度和实时性还待进一步探究;基于超越滑动窗口法的行人检测性能较为优越,但该类方法硬件成本高,不适用于一般行人检测任务,如何设计与该类技术相结合的行人检测方法仍需进一步研究。

### 参考文献:

- [1] Zheng G, Chen Y B. A review on vision-based pedestrian detection[C]//2012 IEEE Global High Tech Congress on Electronics. Shenzhen, China. IEEE, 2012: 49-54.
- [2] Ahmed Z, Iniyavan R, Madhan M P. Enhanced vulnerable pedestrian detection using deep learning[C]//2019 International Conference on Communication and Signal Processing (ICCSP). Chennai, India. IEEE, 2019: 971-974.
- [3] Song Y, Li M, Qiu X H, et al. Full-time infrared feature pedestrian detection based on CSP network [C] // 2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS). Vientiane, Laos. IEEE, 2020: 516-518.
- [4] 雷诗瑶. 改进的候选区域生成网络应用于半监督行人检测[D]. 广州: 华南理工大学, 2019: 1-17.
- [5] Renu Chebrolu K N, Kumar P N. Deep learning based pedestrian detection at all light conditions [C] // 2019 International Conference on Communication and Signal Processing (ICCSP). Chennai, India. IEEE, 2019: 838-842.
- [6] Cao J L, Pang Y W, Xie J, et al. From handcrafted to deep features for pedestrian detection: a survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(9): 4913-4934.
- [7] Feng T T, Ge H Y. Pedestrian detection based on attention mechanism and feature enhancement with SSD [C] // 2020 5th International Conference on Communication, Image and Signal Processing (CCISP). Chengdu, China. IEEE, 2020: 145-148.
- [8] Li F, Li X Y, Liu Q, et al. Occlusion handling and multi-scale pedestrian detection based on deep learning: a review [J]. IEEE Access, 2022, 10: 19937-19957.
- [9] Cheng Y, Li B J. Image segmentation technology and its application in digital image processing [C] // 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC). Dalian, China. IEEE, 2021: 1174-1177.
- [10] Guo X Q, Yang C, Liu Y J, et al. Learn to threshold: ThresholdNet with confidence-guided manifold mixup for polyp segmentation [J]. IEEE Transactions on Medical Imaging, 2021, 40(4): 1134-1146.
- [11] Li Y, Zhang Y, Cui W G, et al. Dual encoder-based dynamic-channel graph convolutional network with edge enhancement for retinal vessel segmentation [J]. IEEE Transactions on Medical Imaging, 2022, 41(8): 1975-1989.
- [12] 罗艳, 张重阳, 田永鸿, 等. 深度学习行人检测方法综述 [J]. 中国图象图形学报, 2022, 27(7): 2094-2111.
- [13] Song Y H, Yan H. Image segmentation techniques overview [C] // 2017 Asia Modelling Symposium (AMS). Kota Kinabalu, Malaysia. IEEE, 2018: 103-107.
- [14] 唐清. 阈值分割及红外图像行人检测研究 [D]. 广州: 华南理工大学, 2010: 10-24.
- [15] Swetha V, Sushma K, Divya Praneetha N, et al. Application to pedestrian detection and object detection [C] // 2022 6th International Conference on Computing Methodologies and Communication (ICCMC). Erode, India. IEEE, 2022: 1462-1466.
- [16] Su H Q, Wen C J. A new algorithm based on super-green features for ostu's method using image segmentation [C] // 2012 World Automation Congress. Puerto Vallarta, Mexico. IEEE, 2012: 1-4.
- [17] Trivedi S, Kumar Khunteta D, Narayan S. Hand segmentation using modified K-means clustering with depth information and adaptive thresholding by histogram analysis [C] // 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI). Chennai, India. IEEE, 2018: 1607-1609.
- [18] Liu Y, Xie Z W, Liu H. An adaptive and robust edge detection method based on edge proportion statistics [J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2020, 29: 5206-5215.
- [19] Fernandez J J, Nithyanandam P. Fingerprint core point detection using connected component approach and orientation map edge tracing approach [J]. International Journal of Biometrics, 2021, 13(4): 409-431.
- [20] Ma Y L, Ma H Y, Chu P C. Demonstration of quantum image edge extraction enhancement through improved Sobel operator [J]. IEEE Access, 2020, 8:

- 210277-210285.
- [21] Wu T, Wang L W, Zhu J C. Image edge detection based on Sobel with morphology[C]//2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Xi'an, China. IEEE, 2021: 1216-1220.
- [22] Zhang Y, Han X Y, Zhang H, et al. Edge detection algorithm of image fusion based on improved Sobel operator[C]//2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC). Chongqing, China. IEEE, 2017: 457-461.
- [23] Shi J T, Zhu Y, Chu X L, et al. A novel approach of edge detection based on gray weighted absolute correlation degree and Prewitt operator [C] // 2010 International Conference on Intelligent Computing and Integrated Systems. Guilin. IEEE, 2010: 232-234.
- [24] Hu H, Cai H, Ma Z Y, et al. Semantic segmentation based on semantic edge optimization [C] // 2021 International Conference on Electronic Information Engineering and Computer Science (EIECS). Changchun, China. IEEE, 2021: 612-615.
- [25] Gao C, Zhang X, Wang H. A combined method for multi-class image semantic segmentation[J]. IEEE Transactions on Consumer Electronics, 2012, 58(2): 596-604.
- [26] Jiang H, Zhang C, Wu M. Pedestrian detection based on multi-scale fusion features[C]//2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC). Guiyang, China. IEEE, 2018: 329-333.
- [27] 田莹,王亮,丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2): 440-468.
- [28] Li P H, Zhang L J, Qiao J L, et al. A semantic segmentation method based on improved U-net network [C] // 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE). Changsha, China. IEEE, 2021: 600-603.
- [29] 姜宏宇. 面向道路场景的轻量化语义分割研究[D]. 北京: 中国地质大学(北京), 2021: 7-33.
- [30] Ma N, Chen L, Hu J C, et al. Pedestrian detection based on HOG features and SVM realizes vehicle-human-environment interaction [C] // 2019 15th International Conference on Computational Intelligence and Security (CIS). Macao, China. IEEE, 2020: 287-291.
- [31] Liu Z H. Improvement of feature extraction based on HOG [C] // 2021 2nd Asia Symposium on Signal Processing (ASSP). Beijing, China. IEEE, 2022: 55-60.
- [32] Bilal M, Hanif M S. Benchmark revision for HOG-SVM pedestrian detector through reinvigorated training and evaluation methodologies[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(3): 1277-1287.
- [33] Fan G J, Li B, Mu W Q, et al. HOGG: Gabor and HOG-based human detection [C] // 2016 8th International Conference on Information Technology in Medicine and Education (ITME). Fuzhou, China. IEEE, 2017: 562-566.
- [34] Peng W L, Yu X, Rong X. Human detection with log-polar transform and HOG-LBP features [J]. ICIC Express Letters, 2018, 12(7): 637-644.
- [35] 林才纺. 基于 CENTRIST 特征行人检测算法的 GPU 实现[D]. 哈尔滨: 哈尔滨工业大学, 2017: 6-39.
- [36] Park W J, Kim D H, Suryanto, et al. Fast human detection using selective block-based HOG-LBP[C]//2012 19th IEEE International Conference on Image Processing. Orlando, USA. IEEE, 2013: 601-604.
- [37] Watanabe T, Ito S, Yokoi K. Image feature descriptor using co-occurrence histograms of oriented gradients for human detection[J]. The Journal of the Institute of Image Information and Television Engineers, 2017, 71(1): J28-J34.
- [38] 种衍文, 匡湖林, 李清泉. 一种基于多特征和机器学习的分级行人检测方法[J]. 自动化学报, 2012, 38(3): 375-381.
- [39] Schwartz W R, Kembhavi A, Harwood D, et al. Human detection using partial least squares analysis [C]//IEEE 12th International Conference on Computer Vision. Kyoto, Japan. IEEE, 2010: 24-31.
- [40] 储珺, 束雯, 周子博, 等. 结合语义和多层特征融合的行人检测[J]. 自动化学报, 2022, 48(1): 282-291.
- [41] Dollár P, Tu Z W, Perona P, et al. Integral channel features[C]//Proceedings of the British Machine Vision Conference, BMVC. London, UK. British Machine Vision Association, 2009: 1-11.
- [42] Ma S Y, Bai L. A face detection algorithm based on Adaboost and new Haar-like feature [C] // 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS). Beijing, China. IEEE, 2017: 651-654.
- [43] Yang Y S, Lin L. Automatic pedestrians segmentation based on machine learning in surveillance video[C]//2019 IEEE International Conference on Computational Electromagnetics (ICCEM). Shanghai, China. IEEE, 2019: 1-3.



- [44] 王焱. 基于随机梯度提升决策树的行人检测算法设计与实现[D]. 杭州: 浙江大学, 2017: 33-66.
- [45] Shridhar H, Neelgar B, Premananda R. An improved image fusion technique based on wavelet and edge gradients to augment spatial prominence of image features [C] // 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT). Mysuru, India. IEEE, 2020: 416-421.
- [46] Yang L, Song Q, Wang Z H, et al. Hier R-CNN: instance-level human parts detection and a new benchmark [J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2021, 30: 39-54.
- [47] Zhang H, Du Y, Ning S R, et al. Pedestrian detection method based on Faster R-CNN [C] // 2017 13th International Conference on Computational Intelligence and Security (CIS). Hong Kong, China. IEEE, 2018: 427-430.
- [48] Cai Z W, Fan Q F, Feris R S, et al. A unified multi-scale deep convolutional neural network for fast object detection [C] // Proceedings of the 14th European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 354-370.
- [49] Tian Q, Wang M H, Zhang Y, et al. A research for automatic pedestrian detection with ACE enhancement on Fasters R-CNN [C] // 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Beijing, China. IEEE, 2019: 1-9.
- [50] Zhang S S, Yang J, Schiele B. Occluded pedestrian detection through guided attention in CNNs[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 6995-7003.
- [51] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multiBox detector [C] // Proceedings of the 2016 European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 21-37.
- [52] 裴伟, 许晏铭, 朱永英, 等. 改进的SSD航拍目标检测方法[J]. 软件学报, 2019, 30(3): 738-758.
- [53] Dong X W, Han Y, Li W, et al. Pedestrian detection in metro station based on improved SSD[C]//2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Dalian, China. IEEE, 2020: 936-939.
- [54] Zhang N, Fan J H. A lightweight object detection algorithm based on YOLOv3 for vehicle and pedestrian detection[C]//2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC). Dalian, China. IEEE, 2021: 742-745.
- [55] Guo W J, Shen N B, Zhang T Z. Overlapped pedestrian detection based on YOLOv5 in crowded scenes[C] // 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA). Changchun, China. IEEE, 2022: 412-416.
- [56] Li Y J, Li S S, Du H H, et al. YOLO-ACN: focusing on small target and occluded object detection[J]. IEEE Access, 2020, 8: 227288-227303.
- [57] Oren M, Papageorgiou C, Sinha P, et al. Pedestrian detection using wavelet templates[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Juan, PR, USA. IEEE, 2002: 193-199.
- [58] Lei Y, Huang Z J. Research on pedestrian detection algorithm based on monocular vision [C] // 2018 International Conference on Robots & Intelligent System (ICRIS). Changsha, China. IEEE, 2018: 161-163.
- [59] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting [J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [60] Kim T K, Cipolla R. MCBoost: multiple classifier boosting for perceptual co-clustering of images and visual features[C]//Conference on Neural Information Processing Systems. New York: ACM, 2008: 841-848.
- [61] Jiang J F, Xiong H. Fast pedestrian detection based on HOG-PCA and gentle AdaBoost [C] // 2012 International Conference on Computer Science and Service System. Nanjing, China. IEEE, 2012: 1819-1822.
- [62] Lin Z, Hua G, Davis L S. Multiple instance fFeature for robust part-based object detection[C]//2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009). Miami, Florida, USA. IEEE, 2009: 405-412.
- [63] Felzenszwalb P F, Girshick R B, McAllester D, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [64] Felzenszwalb P F, Girshick R B, McAllester D.

- Cascade object detection with deformable part models [C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA. IEEE, 2010: 2241-2248.
- [65] Lampert C H, Blaschko M B, Hofmann T. Efficient subwindow search: a branch and bound framework for object localization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31 (12): 2129-2142.
- [66] An S J, Peursum P, Liu W Q, et al. Efficient algorithms for subwindow search in object detection and localization[C]//IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA. IEEE, 2009: 264-271.
- [67] Thi T H, Cheng L, Zhang J, et al. Weakly supervised action recognition using implicit shape models[C]//2010 20th International Conference on Pattern Recognition. Istanbul, Turkey. IEEE, 2010: 3517-3520.
- [68] Dollár P, Appel R, Belongie S, et al. Fast feature pyramids for object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36 (8): 1532-1545.
- [69] Nam W, Dollár P, Han J H. Local decorrelation for improved pedestrian detection[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems-Volume 1. New York: ACM, 2014: 424-432.
- [70] Zhang L L, Lin L, Liang X D, et al. Is faster R-CNN doing well for pedestrian detection? [C]//Proceedings of the 14th European Conference on Computer Vision (ECCV). Cham: Springer International Publishing, 2016: 443-457.

(责任编辑:康 锋)