

文章编号:1008-1542(2019)06-0533-07

基于自适应池化的行人检测方法

余珮嘉¹, 张 靖², 谢晓尧³

(1. 贵州大学计算机科学与技术学院, 贵州贵阳 550025; 2. 贵州大学电气工程学院, 贵州贵阳 550025; 3. 贵州师范大学贵州省信息与计算科学重点实验室, 贵州贵阳 550001)

摘 要: 基于卷积神经网络的行人检测器普遍采用图像识别网络, 通常会引起多池化层导致小目标行人特征信息丢失、单一池化方法导致行人局部重要特征信息削弱甚至丢失等, 针对以上问题, 基于最大值池化和平均值池化方法, 提出了一种自适应池化方法, 结合通用目标检测器 Faster R-CNN, 形成了有效的行人检测器, 达到增强行人局部重要特征信息、保留小目标行人有效特征信息的目的。对多个公开的行人数据集进行大量实验, 结果表明, 与传统的卷积神经网络行人检测器相比, 所提方法将行人检测漏检率降低了 2%~3%, 验证了方法的有效性。新方法改进了卷积神经网络结构, 在无人驾驶领域具有一定的参考价值。

关键词: 神经网络; 卷积神经网络; 行人检测; 图像识别; 自适应池化; Faster R-CNN

中图分类号: TP183 文献标志码: A doi:10.7535/hbkd.2019yx06011

Pedestrian detection based on adaptive pooling method

YU Peijia¹, ZHANG Jing², XIE Xiaoyao³

(1. College of Computer Science and Technology, Guizhou University, Guiyang, Guizhou 550025, China; 2. College of Electrical Engineering, Guizhou University, Guiyang, Guizhou 550025, China; 3. Key Laboratory of Information and Computing Science Guizhou Province, Guizhou Normal University, Guiyang, Guizhou 550001, China)

Abstract: Pedestrian detectors based on convolutional neural networks generally adopt image recognition network, which usually causes the following problems: 1) multi-pool layers lead to the loss of feature information of small target pedestrian; 2) the single pool method leads to the weakening or even loss of the local important feature information of pedestrians. Therefore, based on the maximum pooling and average pooling methods, an adaptive pooling method is proposed, and combined with the Faster R-CNN, an effective pedestrian detector is formed, so as to enhance the local important feature information of pedestrians and retain the effective feature information of small target pedestrians. Through a large number of experiments on several public pedestrian datasets, the results show that compared with the traditional convolutional neural network pedestrian detector, the proposed method reduces the miss rate by about 2%~3%, which verifies the effectiveness of the method.

收稿日期: 2019-07-30; 修回日期: 2019-11-24; 责任编辑: 陈书欣

基金项目: 贵州省发改委重点项目(0502213V0002); 贵州省科学技术基金([2016]1036); 贵州省科技创新人才队伍项目([2018]5615)

第一作者简介: 余珮嘉(1981—), 女, 贵州贵阳人, 讲师, 博士研究生, 主要从事模式识别、深度学习方面的研究。

通信作者: 谢晓尧教授。E-mail: xiaoy_xie@outlook.com

余珮嘉, 张靖, 谢晓尧. 基于自适应池化的行人检测方法[J]. 河北科技大学学报, 2019, 40(6): 533-539.

YU Peijia, ZHANG Jing, XIE Xiaoyao. Pedestrian detection based on adaptive pooling method[J]. Journal of Hebei University of Science and Technology, 2019, 40(6): 533-539.

Keywords: computer neural network; convolution neural networks; pedestrian detection; image recognition; adaptive pooling; Faster R-CNN

在计算机视觉研究领域中,目标检测作为一类基本问题得到了深入的研究^[1-5]。近年来,在辅助驾驶、智能视频监控等应用背景下,行人检测作为目标检测中的一类重要应用,受到了国内外学术界和工业界的广泛关注^[6-9]。

过去 10 多年,出现了许多基于传统人工设计特征的行人检测方法。2005 年,文献[10]提出了一种描述图像局部特征的方法,通过对局部区域中每个像素点的梯度方向和幅值进行统计,获得了方向梯度直方图(histograms of oriented gradients, HOG),与后续分类器结合,构成了有效的行人检测器。文献[11]提出了利用 HOG 特征与 Boosted Cascad 算法^[12]级联结构相结合,实现了行人的快速检测。2009 年,文献[13]提出了基于积分通道特征(integrate channel features, ICF)的快速行人检测器,不同于 HOG 的单一特征,ICF 包含了图像的多种通道特征,特征信息更加丰富,同时采用 AdaBoost 分类器构成的检测器,从检测精度和运行速度上都要优于 HOG 检测器。文献[14]通过扩展 HOG 检测器提出了可形变部件模型(deformable parts model, DPM),它是一种基于组件的检测方法,对于遮挡类行人具有较好的检测效果。文献[15]对 DPM 方法进行了改进,增加了级联检测算法和分支定界算法,有效提高了检测速度。文献[16]提出了改进的颜色自相似性方法,能够降低颜色自相似性维度,提升特征提取速度和行人检测速度。

随着卷积神经网络(convolutional neural networks, CNNs)技术在计算机视觉领域的快速发展,基于 CNNs 的行人检测取得了巨大进展。卷积神经网络采用多层结构,目标物体的特征可以通过学习自动获取。文献[17]提出 TA-CNN 网络,通过加入语义信息,比如行人辅助信息(如背包),场景信息(如树)等,将多种信息融合用于处理行人检测中的难例问题。文献[18]提出将通道特征与基于卷积神经网络的行人检测器相融合,不光将通道特征加入,同时将卷积神经网络中逐层的特征图结合到最终的特征提取,丰富的特征显著提升了行人的检测精度。文献[19]提出利用提升森林策略替换通用目标检测器 Faster R-CNN^[2]的分类器,新的策略可以融合多种分辨率的特征图,提升难例挖掘能力。文献[20]提出对通用目标检测器 Faster R-CNN^[2]进行改进,使其适合于行人检测,其中一项重要改进是通过减少池化层数量达到调小降采样因子,目的是保留更多行人的特征信息,尤其对小目标行人的检测有较大提升。文献[21]中为目标检测建立专门的骨干特征提取网络,通过减少池化层数量达到减少感受野的大小,较小感受野能提升小目标的识别精度和大目标的定位精度。上述的行人检测器大多采用在图像分类任务(比如 ImagNet 数据集)中获得较好效果的分网络(比如 AlexNet^[3], VGG16^[4])作为行人检测器的特征提取网络。但是使用分类网络作为行人检测的特征提取网络会从以下方面导致特征信息丢失:1)分类网络中普遍采用较多池化层,虽然可以降低网络的计算量和增大感受野,对图像分类任务有利,但是多层池化形成的较大特征步长作用在宽度较小(比如 40 像素)的目标行人上,将导致小尺寸行人特征信息丢失,进而降低识别结果的置信度(如 Caltech 数据集中,大部分行人的高×宽=80 像素×40 像素);2)分类网络中使用的池化层通常采用单一池化方法,如最大值池化(max pooling)或者平均值池化(average pooling),行人作为一类特殊目标,具有尺度较小、外形特征较复杂的特点,但是在对行人特征提取过程中,因使用单一池化方法,不仅获取的特征信息粗糙,而且池化方法与局部特征不匹配会导致重要特征信息丢失。

针对以上问题,提出了基于自适应池化的行人检测方法,该方法能够根据行人身体各部位的特征,通过学习获得适合的池化层,改善网络特征提取能力。

1 理论基础

池化方法是指池化区域中用总体统计特征来代替网络在该区域的输出。卷积神经网络中使用池化方法可使网络具有特征不变性、特征降维和防止过拟合等功能。

1.1 多次池化

二维图像经过单次池化后输出的特征图空间分辨率如式(1)所示:

$$o_j^i = \left(\frac{A_j^{i-1} - k_j^i}{s_j^i} \right) + 1, \quad j = 1, 2; i = 1, 2, \dots, n, \quad (1)$$

式中: A_j^{i-1} 表示第 i 次池化的输入特征图沿 j 维度方向的大小; k_j^i 表示第 i 次池化沿 j 维度方向的核大小; s_j^i 表示第 i 次池化沿 j 维度方向的步长; o_j^i 表示第 i 次池化的输出特征图沿 j 维度方向的大小。

在典型的图像分类网络中,通常令 $k_j^i = s_j^i$,因此式(1)可以变换为式(2):

$$o_j^i = \frac{A_j^{i-1}}{s_j^i}, \quad j = 1, 2; i = 1, 2, \dots, n. \quad (2)$$

根据式(2)可以推导得出,经过 n 次池化,输出特征图沿 j 维度方向的大小 o_j^n 与原始输入图沿 j 维度方向的大小 A_j^0 的关系,如式(3)所示:

$$o_j^n = \frac{A_j^0}{\prod_{i=1}^n s_j^i}, \quad j = 1, 2. \quad (3)$$

从式(3)可以得到,经过 n 次池化,输出特征图的空间分辨率将逐级降低,池化次数越多,池化步长累积越大,丢失的信息也就越多,输出特征图的空间分辨率较原始图像降低。对于行人这类小目标,原始图像中行人的分辨率较小,比如 Caltech 数据集中,大部分行人的高 \times 宽 = 80 像素 \times 40 像素,池化步长越大,则丢失的行人特征信息越多。该多次池化特征提取示意图如图 1 所示。

1.2 自适应池化方法

基于卷积神经网络的行人检测器通常采用的池化方法为式(4)所示的最大值池化和式(5)所示的平均值池化。

最大值池化是指在每个池化区域中,选择最大值的元素来表示池化区域的输出,池化操作可以表示为

$$y_{ij_max}^{l+1} = \max_{(p,q) \in R_{ij}} x_{pq}^l, \quad (4)$$

式中: y_{ij}^{l+1} 表示池化操作作用在区域 R_{ij} 后的输出; l 表示卷积神经网络的第 l 层特征图; x_{pq}^l 表示在池化区域 R_{ij} 中,位置为 (p, q) 的元素。

平均值池化是指采用池化区域中所有元素的数学平均值来表示池化区域的输出,可以表示为

$$y_{ij_ave}^{l+1} = \frac{1}{N} \sum_{(p,q) \in R_{ij}} x_{pq}^l, \quad (5)$$

式中, N 表示池化区域 R_{ij} 中元素的个数。

虽然最大值池化和平均值池化两种方法分别在很多研究中表现良好,但是这两种池化方法对不同问题的处理效果差异较大,均不能同时兼顾。如图 2 a) 所示的灰度图像中,大部分像素点具有高灰度值,少部分像素点具有低灰度值,两类像素点的值差异明显,在图像中的特征对比明显。经过最大值池化后,具有低灰度值的像素点消失,只有高灰度值的像素点被保留,原始图像中高低灰度值对比明显的特征消失。如果采用平均值池化方法,仍保留了高低灰度值对比的特征。如图 2 b) 所示的灰度图像中,大部分像素点具有低灰度值,少部分像素点具有高灰度值,两类像素点的值差异明显,在图像中的特征对比明显。经过平均值池化

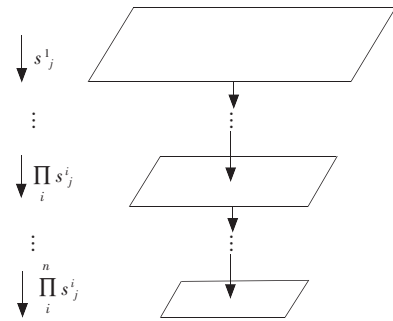
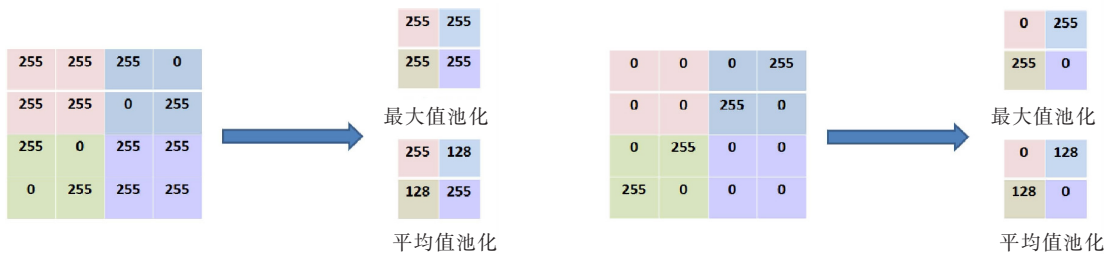


图 1 多次池化特征提取示意图

Fig.1 Feature extraction based on multi-pooling



a) 高灰度值图像的池化方法比较

b) 低灰度值图像的池化方法比较

图 2 池化方法对特征信息影响示意图

Fig.2 Impact of pooling methods to the features

后,高低灰度值对比的特征信息存在,但是被弱化。如果采用最大值池化方法,高低灰度值对比明显的特征信息仍然被保留^[22]。

因此,无论采用以上哪种池化方法,考虑到行人身体结构的复杂性,在特征提取过程中,将引起特征信息的削弱甚至丢失。为此,文献^[22]提出了随机混合池化方法,将最大值池化和平均值池化进行简单线性连接,但是其线性连接系数为随机数,不能可靠地学习对象的不同特征以获得适合的池化层。为解决该问题,本文提出自适应池化方法,根据行人身体各部位的特征,通过学习获得自适应池化层。该方法表示如下:

$$y_{ij}^l = f_{act}(f_{con}(y_{ij_max}^l, y_{ij_ave}^l)), \tag{6}$$

式中: $y_{ij_max}^l$ 和 $y_{ij_ave}^l$ 分别表示第 l 层中最大值池化和平均值池化在位置 (i, j) 的输出值; f_{con} 表示自适应连接函数,连接图像中相同区域的最大值池化和平均值池化,本文采用卷积函数来实现 f_{con} ; f_{act} 为非线性激活函数。

为了实现对图像中每个局部区域最大值池化和平均值池化的自适应连接,使用宽、高分别为1,通道为2的卷积核。非线性激活函数 f_{act} 用于增强 f_{con} 的表达能力,文中采用线性整流(rectified linear unit, ReLU)函数作为非线性激活函数。图3描述了自适应池化方法与普通池化方法结构的区别。

2 实验结果与讨论

Caltech^[8], CityPersons^[20]和ETH^[23]是目前研究中通常采用的数据集,每个数据集的图片大小、分辨率等参数都存在明显差异性,因此,为验证本文所提方法的通用性和有效性,对这3个典型行人数据集进行测试,并分别与采用最大值池化方法和平均值池化方法的卷积神经网络行人检测器进行对比。

实验采用基于CNNs的通用目标检测器Faster R-CNN作为行人检测器。在网络训练过程中,使用水平翻转方法进行数据量的增广,采用基于动量(Momentum)的随机梯度下降(stochastic gradient descent, SGD)算法进行训练,动量值设置为0.9,初始学习率设置为0.001。训练中用于判定候选框中为正/负样例的交并比(intersection-over-union, IoU)设置为0.5。

2.1 评价指标

采用文献^[8]所示的评价指标来评估行人检测方法的性能。该评价指标描述了平均每张图片不同错误正样本数量(false positives per-image, FPPI)下对应的漏检率(miss rate, MR)。FPPI具体指测试集中平均每张图片错误正样本数量。MR是漏检率,指测试集图片中没有被框中的行人在整个测试集中的比率。FPPI与MR具有较强的相关性,FPPI值越大,MR值越小,为了比较不同行人检测方法的性能,统一采用 $FPPI=10^{-1}$ 时不同方法对应的MR值为衡量标准,其中MR值越小,表示检测方法越好。MR和FPPI如式(7)和式(8)所示:

$$MR = \frac{\sum_{i=1}^{N_{gt}} GT(i)}{N_{gt}} \times 100\%, \tag{7}$$

式中: $GT(i)$ 是判断标注框是否被匹配,未匹配则置为1,匹配则置为0; N_{gt} 是测试数据集中标注框(ground truth)的总数。

$$FPPI = \frac{\sum_{i=1}^{N_{fb}} FP(i)}{N}, \tag{8}$$

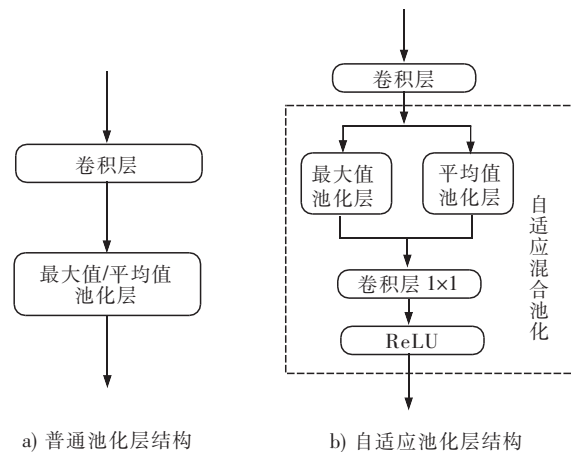


图3 通用池化层与自适应池化层结构对比
Fig.3 Structures comparison of general pooling layer and adaptive pooling layer

式中:FP(i)是检测器的检测结果是否框住行人,未框住则置为 1,框住则置为 0; N 是测试数据集中图片的总数。

2.2 Caltech 数据集上的实验结果与分析

Caltech 数据集是由从大约 10 h 的视频中提取出的图片构成的。整个数据集分为训练集和测试集 2 部分,训练集共 6 个子集(set00-set05),测试集共 5 个子集(set06-set10)。数据集中的所有图片大小均为 640 像素×480 像素。表 1 列出了 Caltech 数据集的详细信息。

表 1 Caltech 数据集的详细信息

Tab.1 Details of Caltech dataset

图片总数/张	训练集/张	测试集/张	标注框数量/个
250 000	245 976	4 024	350 000

表 2 中列举了基于 Faster R-CNN 检测网络在不同数量池化层对应的行人检测结果。从表 2 中可以得到,随着池化层数量的增加,对应的池化步长相应增大,特征信息的丢失增多,表征行人检测效果的漏检率(MR)逐步增大,检测性能降低。池化层数量过少,比如池化数目为 2,池化步长为 4,并不是提升检测性能的有效方法,反而会使网络可能出现过拟合情况。因为,当池化层数量为 3、池化步长为 8 时,既能使网络避免过拟合,行人特征具有不变性,又能有效控制特征信息丢失过多等问题。

为了验证本文所提基于自适应池化的行人检测方法,将其与基于最大值池化的行人检测方法和基于平均值池化的行人检测方法进行比较。实验中,将 3 种池化方法分别应用于包含 3 个池化层的 Faster R-CNN 检测网络,使用相同的 Caltech 数据集。从表 3 中可以看出,最大值池化方法和平均值池化方法的漏检率分别为 16.93%和 17.07%,而本文所提方法漏检率为 14.91%。可以看出,本文所提方法漏检率明显降低,对行人检测的性能提升效果较好。

本文基于自适应池化方法改进 Faster R-CNN 目标检测器,使其适用于行人检测。改进后的行人检测器与多个行人检测器在 Caltech 数据集上进行比较。如图 4 所示,进行比较的检测器大体分为 3 类。1) 基于传统人工特征的检测器, VJ, HOG, ACF + SDt, InformedHaar, ACF-Caltech+, LDCF 检测器; 2) 基于 DPM 的检测器, MT-DPM, Joint-Deep, MT-DPM+Context; 3) 基于 CNN 的检测器, SpatialPooling, TA-CNN。本文提出的基于自适应池化(AdaptPooling)方法改进的行人检测取得了较好的检测效果, MR 为 14.91%。

2.3 其他数据集上的实验结果与分析

为了验证所提方法在其他行人数据集上的通用性,采用与 Caltech 数据集有明显差别的 CityPersons 数据集和 ETH 数据集进行实验分析。

表 2 基于不同池化步长的行人检测效果对比

Tab.2 Comparison of different pooling strides

池化数目	池化步长	漏检率/%
2	4	41.04
3	8	19.11
4	16	21.08
5	32	50.98

表 3 在 Caltech 数据集上对比不同的池化方法

Tab.3 Comparison of different pooling methods on the Caltech dataset

池化方法	漏检率/%
最大值池化方法	16.93
平均值池化方法	17.07
自适应池化方法	14.91

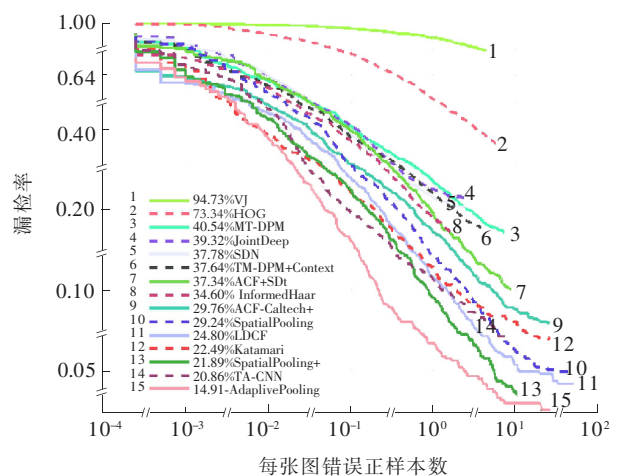


图 4 Caltech 数据集上多种行人检测方法比较

Fig.4 Comparison of pedestrian detection methods on the Caltech set

CityPersons 数据集包含来自德国和邻国不同城市的大量多样化视频,表 4 给出了 CityPersons 数据集的详细信息。可以看出, CityPersons 数据集不论是总体图片数量还是各个子集图片数量均少于 Caltech 数据集,但是 CityPersons 数据集的图片分辨率较高,标注框质量较好。

采用相同方法在 CityPersons 数据集上进行训练与测试。实验结果如表 5 所示,最大值池化方法和平均值池化方法的漏检率分别为 19.26% 和 20.61%,而本文所提方法漏检率为 15.67%。可以看出,本文所提方法漏检率明显降低,对行人检测的性能提升效果较好。

ETH 数据集是由一对装载于移动平台上的双目摄像头拍摄得到的视频。训练数据集由 490 张图片组成,测试数据集有 1 803 张图片。训练集的数据量很少,无法训练具有较多参数的 Faster R-CNN 检测器。本文在 Caltech+CityPersons 数据集上训练检测器,而使用 ETH 数据集进行测试。实验结果如表 6 所示,最大值池化方法和平均值池化方法的漏检率分别为 22.01% 和 25.81%,而本文所提方法漏检率为 19.77%,漏检率明显降低,对行人检测的性能提升效果较好。

对比表 3、表 5 和表 6 的结果可以看出,将在 Caltech+CityPersons 数据集上训练好的行人检测器用于测试 ETH 数据集,虽然各种方法的漏检率略微提高,但仍然能得到很好的效果,有较好的通用性。

2.4 运行时间

采用英伟达 1080Ti 显卡对 3 种方法进行了测试,结果如表 7 所示。基于最大值池化和平均值池化的行人检测器的运行时间都约为 50 ms,所提出的自适应池化的行人检测器的运行时间稍多一些,为 58 ms,但与前两种方法仍处于同一数量级。虽然注意到这是以略微增加的计算时间成本为代价(见表 3),但是本文提出的基于自适应池化的行人检测方法在 Caltech 数据集上达到了较好的检测效果,漏检率提高到 14.91%。

通过在主流的公开数据集 Caltech, CityPersons 以及 ETH 上进行大量实验,计算了漏检率、运行时间,并分别与采用最大值池化方法和平均值池化方法的卷积神经网络行人检测器相比,本文所提出的方法均能降低漏检率 2%~3%,表明了本文所提方法是通用的和有效的。

3 结 语

提出了基于自适应池化的行人检测方法,自适应池化方法考虑了行人特征的复杂性,在行人特征提取过程中,能根据行人身体的不同部位特征,将最大值池化和平均值池化进行自适应混合处理。基于自适应池化的行人检测器,能获得更加丰富的行人特征,对低分辨率的小目标行人的特征提取能力要明显优于采用单一池化方法的行人检测器。通过对多个公开的标准行人数据集进行大量实验,结果验证了所提出的基于自适应池化的行人检测方法能明显提高行人检测的精度,降低漏检率,也可灵活应用于其他目标检测器。未来的研究重点是进一步提升检测精度,并对整个检测系统进行算法优化,降低检测运行时间,更好地实现检测系

表 4 CityPersons 数据集的详细信息

图片总数/张	训练集/张	验证集/张	测试集/张
5 000	2 975	500	1 575

表 5 在 CityPersons 数据集上对比不同的池化方法

池化方法	漏检率/%
最大值池化方法	19.26
平均值池化方法	20.61
自适应池化方法	15.67

表 6 在 ETH 数据集上对比不同的池化方法

池化方法	漏检率/%
最大值池化方法	22.01
平均值池化方法	25.81
自适应池化方法	19.77

表 7 Caltech 数据集上运行时间的比较

池化方法	运行时间/ms
最大值池化方法	50
平均值池化方法	50
自适应池化方法	58

统的实时性。

参考文献/References:

- [1] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Boston:IEEE,2015:1440-1448.
- [2] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,39(6):1137-1149.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012,25(2):3065-3086.
- [4] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014: 1556v3.
- [5] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston: IEEE,2015:7298-594.
- [6] 陈政宏, 李爱娟, 邱绪云, 等. 智能车环境视觉感知及其关键技术研究现状[J]. 河北科技大学学报, 2019,40(1):15-23.
CHEN Zhenghong, LI Aijuan, QIU Xuyun, et al. Survey of environment visual perception for intelligent vehicle and its supporting key technologies[J]. Journal of Hebei University of Science and Technology, 2019,40(1):15-23.
- [7] 徐超, 闫胜业. 改进的卷积神经网络行人检测方法[J]. 计算机应用, 2017,37(6):1708-1715.
XU Chao, YAN Shengye. Improved pedestrian detection method based on convolutional neural network[J]. Journal of Computer Applications, 2017,37(6):1708-1715.
- [8] BENENSON R, OMRAN M, HOSANG J, et al. Ten years of pedestrian detection, what have we learned? [C]//European Conference on Computer Vision. Cham: Springer,2014: 613-627.
- [9] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: A benchmark[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S.l.]:[s.n.],2009:304-311.
- [10] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'05). San Diego:IEEE,2005:886-893.
- [11] 叶林, 陈岳林, 林景亮. 基于 HOG 的行人快速检测[J]. 计算机工程, 2010,36(22):206-207.
YE Lin, CHEN Yuelin, LIN Jingliang. Pedestrian fast detection based on histograms of oriented gradient[J]. Computer Engineering, 2010, 36(22):206-207.
- [12] VIOLA P, JONES M J. Robust real-time face detection[J]. International Journal of Computer Vision, 2004,57(2):137-154.
- [13] DOLLÁR P, TU Z W, PERONA P, et al. Integral channel features[C]//Proceedings of the British Machine Vision Conference. London: [s.n.], 2009: 10.5244/C.23.91.
- [14] FELZENSZWALB P F, MCALLESTER D A, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]//2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Alaska: IEEE, 2008:4587-597.
- [15] 柴恩惠, 智敏. 融合分支定界的可变形部件模型的行人检测[J]. 计算机应用, 2017,37(7): 2003-2007.
CAI Enhui, ZHI Min. Pedestrian detection based on deformable part model with branch and bound[J]. Journal of Computer Applications, 2017,37(7):2003-2007.
- [16] 顾会建, 陈俊周. 基于改进颜色自相似特征的行人检测方法[J]. 计算机应用, 2014,34(7):2033-2035.
GU Huijian, CHEN Junzhou. Pedestrian detection based on improved color self-similarity feature[J]. Journal of Computer Applications, 2014,34(7):2033-2035.
- [17] TIAN Yonglong, LUO Ping, WANG Xiaogang, et al. Pedestrian detection aided by deep learning semantic tasks[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston:[s.n.],2015:5079-5087.
- [18] MAO Jiayuan, XIAO Tete, JIANG Yuning, et al. What can help pedestrian detection? [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). [S.l.]:[s.n.],2017:3127-3136.
- [19] ZHANG Liliang, LIN Liang, LIANG Xiaodan, et al. Is Faster R-CNN doing well for pedestrian detection? [C]//European Conference on Computer Vision. Cham: Springer,2016:443-457.
- [20] ZHANG Shanshan, BENENSON R, SCHIELE B. CityPersons: A diverse dataset for pedestrian detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]:[s.n.],2017:4457-4465.
- [21] LI Z, PENG C, YU G, et al. Detnet: Design backbone for object detection[C]//Proceedings of the European Conference on Computer Vision (ECCV). [S.l.]:[s.n.], 2018:334-350.
- [22] YU Dingjun, WANG Hanli, CHEN Peiqiu, et al. Mixed pooling for convolutional neural networks[C]//The 9th International Conference on Rough Sets and Knowledge Technology(RSKT 2014). Shanghai:[s.n.], 2014:364-375.
- [23] ESS A, LEIBE B, van GOOL L, et al. Depth and appearance for mobile scene analysis[C]//2007 IEEE 11th International Conference on Computer Vision. Rio de Janeiro:IEEE,2007:4409-092.