

## 一种面向遮挡行人检测的改进 YOLOv3 算法

李翔<sup>1,2,3,4</sup>, 何淼<sup>1,2,3</sup>, 罗海波<sup>1,2,3\*</sup><sup>1</sup>中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;<sup>2</sup>中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;<sup>3</sup>中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110169;<sup>4</sup>中国科学院大学, 北京 100049

**摘要** 在密集行人检测场景中,目标间的相互遮挡重叠会造成 YOLOv3 模型的检测性能下降。针对造成 YOLOv3 性能下降的原因提出三点改进。一是提出了一种聚拢损失函数,通过优化预测框坐标的方差与均值,使得属于同一个目标的预测框更加紧致,进而降低假阳性率。二是提出了一种高分辨率特征金字塔,通过上采样提高每层金字塔特征的分辨率,并引入浅层特征以增强相邻子特征的差异,从而为高重叠目标生成具有区分度的深度特征。三是提出了一种基于空间注意力机制的检测头,用以降低冗余预测框的数量,减少非极大值抑制(NMS)过程的计算负担。在密集行人数据集 CrowdHuman 上进行的实验的结果显示,所提算法在使用传统 NMS 方法的情况下使得 YOLOv3 检测的平均精度和召回率分别提高了 2.91 个百分点和 3.20 个百分点,丢失率降低了 1.24 个百分点,有效提升了对遮挡行人的检测性能。

**关键词** 机器视觉; 目标检测; 神经网络; 行人检测; 注意力机制

中图分类号 TP183

文献标志码 A

DOI: 10.3788/AOS202242.1415003

## Occluded Pedestrian Detection Algorithm Based on Improved YOLOv3

Li Xiang<sup>1,2,3,4</sup>, He Miao<sup>1,2,3</sup>, Luo Haibo<sup>1,2,3\*</sup>

<sup>1</sup>Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China;

<sup>2</sup>Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China;

<sup>3</sup>Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, Liaoning, China;

<sup>4</sup>University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract** In crowded scenes, it is difficult for YOLOv3 to detect the objects that overlap each other heavily. Aiming at the reasons for the decline of YOLOv3 performance, three improvements are proposed. Firstly, a Tight Loss function is proposed, which optimizes the variance and mean of the coordinates of the prediction boxes to make the prediction boxes belonging to the same target more compact, thus reducing the false positive rate. Secondly, a high-resolution feature pyramid is proposed, in which the resolution of each pyramid feature is improved by upsampling, and shallow features are introduced to enhance the differences between adjacent sub-features, so as to generate distinguishing depth features for highly overlapped targets. Thirdly, a detection head based on spatial attention mechanism is proposed to reduce the number of redundant prediction boxes, so as to reduce the computational burden of the non-maximum suppression (NMS) process. The experimental results on the crowded dataset CrowdHuman show that the average accuracy and recall rate of YOLOv3 detection are improved by 2.91 percentage points and 3.20 percentage points, and the miss rate is reduced by 1.24 percentage points by using the proposed algorithms under the condition of using the traditional NMS method, which demonstrates the effectiveness of the proposed algorithms in boosting the performance in occluded pedestrian detection.

**Key words** machine vision; object detection; neural network; pedestrian detection; attention mechanism

## 1 引言

目标检测作为计算机视觉领域的一个重要课题,随着近年来深度卷积神经网络的发展,已经在智能交通<sup>[1-3]</sup>、自动驾驶<sup>[4-8]</sup>、遥感监测<sup>[9-10]</sup>和视频监控<sup>[11-13]</sup>等领域中得到了广泛的应用。行人检测作为目标检测的一个重要分支,往往被应用在商场、交通枢纽和道路交会处等行人目标密集出现的场景中。由于行人目标形状相似度高,故目标间的互相遮挡会给人检测带来挑战,主要表现在:1)传统的非极大值抑制(NMS)方法会错误地抑制高度重叠的预测框;2)目标间的严重遮挡会使得被遮挡目标的可视像素有限,且前景目标会占据被遮挡目标标注框的大部分区域,这会给后续生成可区分的高维深度特征带来困难。

针对传统 NMS 方法会错误抑制高重叠度目标的问题,文献[14]提出了 Soft NMS 方法,根据重叠程度适当降低目标置信度,然后通过指定一个置信度阈值完成对冗余目标框的筛选。文献[15]提出了一种自适应阈值的 NMS 方法——Adaptive NMS,将目标重叠度作为自适应调整不同预测框 NMS 阈值的依据,即抑制所有重叠度大于密度值的目标。上述算法可以有效提升模型对密集遮挡目标的检测能力,但当同属于一个目标的多个预测框不够紧凑一致时,检查结果的假阳率会增加。此外,还有一些工作研究了遮挡目标特征提取。文献[16]提出了一种多尺度注意力机制来提升对高重叠度目标可区分深度特征的提取能力。文献[17]将注意力机制引入特征提取网络中,提出了一种特征提取增强检测算法。以上算法虽然对被其他物体遮挡的目标的检测能力有所提升,但是仍然不能很好地应对目标之间相互严重重叠的情况。

作为目标检测领域具有代表性的检测网络之一,YOLOv3<sup>[18]</sup>模型被广泛应用于小目标检测、遥感目标检测和红外目标检测等领域中。此外,基于 YOLOv3 的改进通用目标检测算法,如 YOLOv4<sup>[19]</sup>、Scaled-YOLOv4<sup>[20]</sup>和 YOLOX<sup>[21]</sup>等也相继被提出,表明 YOLO 算法仍然有很大的改进潜力。本文针对 YOLOv3 在密集遮挡场景中遇到的挑战提出了三点改进。对于回归结果不紧凑导致的假阳率上升的问题,提出了一种聚拢损失函数,以优化的方式促使对应于同一目标的多个预测框更加一致。对于高重叠目标深度特征区分度低的问题,提出了高分辨率金字塔(HRFP),利用上采样扩大特征金字塔的分辨率,并通过融合来自底层的细节信息来提高相邻子特征的可区分性。然而,高分辨率金字塔会增加目标的冗余预测框数量,从而增大后续 NMS 过程的计算量,降低模型的推理速度。因此,提出了一种基于空间注意力机制的检测头,用以降低目标相邻子特征的置信度,减少冗余框的数量。为验证上述改进的有效性,在当前最具有说服力的密集遮挡目标数据集 CrowdHuman<sup>[22]</sup>上进行了实验。实验结果表明,所提改进算法可以有效提升 YOLOv3 对于遮挡目标的检测能力。

## 2 高度重叠目标对 YOLOv3 性能的影响

造成 YOLOv3 对密集重叠目标检测性能下降的原因主要有两个。首先,在面对重叠严重的目标时,若对应同一个目标的多个预测框松散、不一致,则会使得检测结果的假阳率上升。具体来讲,为解决目标尺度变化的问题,YOLOv3 会在特征上铺设密集的锚框来捕获任意尺度、任意形状和任意位置的目标。不同于最原始的 YOLOv3 模型每个目标只分配一个锚框目标与锚框的匹配机制,其后的改进型如 YOLOv4 等,都选择采用宽松的匹配机制,使每个目标可能匹配到多个形状和尺度不同的锚框,而这被证明是一种更好的匹配策略。然而,宽松的匹配机制也意味着属于同一目标的预测框是分别从不同的起点开始回归的,且几乎每个目标的预测框都是冗余的,进而需要进行非极大值抑制。传统 NMS 方法简单地直接将重叠度大于一个固定阈值的预测框过滤掉,这会造成对密集遮挡目标的召回率大幅下降。Soft NMS 采用一种根据重叠程度来降低预测框置信度的方式缓解了这一问题。Adaptive NMS 采用了一种更加合理的自适应交并比(IoU)阈值的方法。然而,当属于同一目标的预测框松散且不够一致时,上述三种主流 NMS 方法的性能都会受到影响。以 Adaptive NMS 为例,如图 1(a)所示,粗实线框与细实线框分别为被遮挡者与遮挡者的标注框,他们之间的 IoU 值即为被遮挡目标的密度,也是自适应 NMS 过程中的 IoU 阈值。Adaptive NMS 认为所有与目标的重叠程度大于目标密度的预测框都是预测了同一个目标的冗余框,应当被抑制。由于自适应 IoU 阈值会对高重叠目标放宽抑制的力度,因此如果对应于同一目标的多个预测框不够紧凑一致的话,很有可能会造成假阳率升高。在图 1(a)中,假设此时图中三个框均为预测框,细实线与细虚线框属于遮挡者目标,粗实线框属于被遮挡者目标,显然细虚线框应当被抑制。此时,由于细实线框与细虚线框不够紧凑,故二者的 IoU 值明显小于细实线框与粗实线框检测的 IoU 值,即小于被遮挡目标的密度,则按照 Adaptive NMS 的机制,细虚线框将会被保留,进而形成假阳预测。因此,若为提升密集重叠目标的召回率而采用 Adaptive NMS,则需要提升预测框的紧凑度。

其次,YOLO 模型的目标与锚框的匹配机制也不利于高重叠度目标的检测。具体来讲,当两个目标高度重叠时,他们的边界框同样高度重叠,如图 1(b)所示,图中圆点与三角形点分别是前景目标与被遮挡目标的中心点。按照 YOLO 的匹配机制,若此时他们的中心点落在了同一个子特征内,会只有一个目标会被分配锚框并送入训练过程,另一个会则直接被忽略,这就使得部分高重叠目标无法作为样本被投入训练过程中。除了上述物理空间层面的问题,高重叠目标之间还存在特征的语义信息区分度低的问题。遮挡关系的

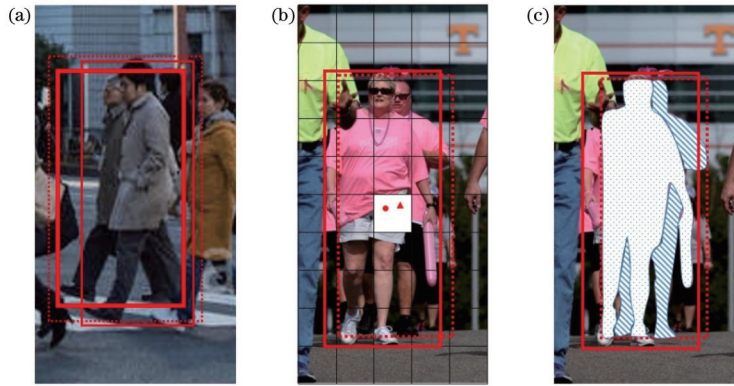


图 1 遮挡目标检测难点示意图。(a)高重叠目标的松散预测框;(b)高重叠目标预测框中心落在同一个特征网格内;(c)被遮挡目标框大部分区域被前景目标所占据

Fig. 1 Illustration of difficulties in occluded object detection. (a) Loose prediction boxes of heavily overlapped objects; (b) center points of prediction boxes of heavily overlapped objects locate in same feature grid; (c) most regions in occluded object box occupied by foreground object

存在会使得遮挡者与被遮挡者之间有大量的共享像素,这导致在特征层面上遮挡者与被遮挡者的区分度较低。如图 1(c)所示,实线框为前景目标的标注框,虚线框为被遮挡目标的标注框,斜线填充区域(以下简称斜线区域)为被遮挡目标的可见像素的集合,点阵填充区域(以下简称点阵区域)为被遮挡目标框中前景目标像素占据的区域。由图可见,与点阵区域相比,斜线区域占据被遮挡目标框的比例十分有限,即被遮挡目标的特征中掺杂了很多点阵区域的信息,进而难以与前景目标区分开来。以上两点都会影响 YOLOv3 对高重叠目标的检测性能。

### 3 对 YOLOv3 检测器的改进算法

本文针对 YOLOv3 模型在重叠目标检测任务中所面临的挑战提出三点改进。首先,针对预测框不紧凑容易产生假阳预测的问题,提出了聚拢损失函数,在优化的层面提升同一目标多个预测框的紧凑程度,降低 NMS 后预测结果的假阳率。其次,对于低分辨率金字塔难以为高重叠目标生成可区分特征的问题,提出了高分辨率金字塔,并引入低层次细节信息来提升

相邻子特征的区分度。最后,针对扩大分辨率后相邻子特征相似造成的稀疏目标冗余预测框增多的情况,提出了一种基于空间注意力机制的检测头。下面分别详细介绍三种改进算法。

#### 3.1 聚拢损失函数

在行人检测中,目标往往具有不同的尺度和任意的位。基于锚框的 YOLOv3 检测器为了不遗漏任何一个拥有随机尺度与位置的目标,会在视野范围内铺设密集且具有不同尺度的锚框。同时,锚框与目标之间的匹配条件会被设置得相对宽松,这就导致一个目标往往会被分配多个锚框,且这些锚框在位置或者尺度上均有所差异。在优化过程中,预测框回归是以这些锚框为起点的,锚框的差异意味着回归到同一目标的难易程度也具有差异。如图 2(a)所示,以不同锚框为起点向同一目标进行回归优化,最终的结果有可能会因为回归的难易程度不一致而有所差异,造成属于同一目标的多个预测框不紧凑的现象的出现。本节提出一种针对预测框坐标的聚拢损失函数来解决上述问题。

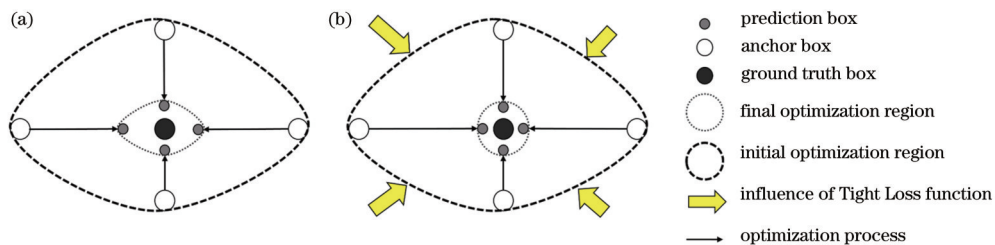


图 2 引入聚拢损失函数前后的收敛结果。(a)无聚拢损失函数时各预测框收敛结果的方差较大;(b)引入聚拢损失函数后以不同锚框为起点的预测框最终收敛于一致

Fig. 2 Convergence results before and after introducing Tight Loss function. (a) Variance of convergence result of prediction box is relatively larger without Tight Loss function; (b) prediction boxes with different anchor frames as starting points tend to be consistent after introducing Tight Loss function



假设输入图像中包含  $M$  个目标, 其中第  $i$  个目标  $t^{(i)}$  匹配了  $N$  个锚框。令  $\varphi_i = \{b_1^{(i)}, b_2^{(i)}, \dots, b_j^{(i)}, \dots, b_N^{(i)}\}$  表示基于以上  $N$  个锚框所产生的对于目标  $t^{(i)}$  的  $N$  个预测框的集合, 其中  $b_j^{(i)} = \{b_{j,x_1}^{(i)}, b_{j,y_1}^{(i)}, b_{j,x_2}^{(i)}, b_{j,y_2}^{(i)}\}$  表示框的左上顶点和右下顶点在原图像坐标系下的坐标值,  $j=1, \dots, N$ 。根据 4 个坐标, 将  $\varphi_i$  拆分成对应于不同坐标的集合  $\varphi_{i,k} = \{b_{1,k}^{(i)}, b_{2,k}^{(i)}, \dots, b_{N,k}^{(i)}\}$ , 其中  $k \in \{x_1, y_1, x_2, y_2\}$ 。因此, 聚拢损失函数可表示为

$$L_{\text{tight}} = \frac{1}{M} \sum_{i=1}^M [L_{\text{mean}}^{(i)} + L_{\text{std}}^{(i)}], \quad (1)$$

其中,  $L_{\text{mean}}^{(i)}$  和  $L_{\text{std}}^{(i)}$  的表达式为

$$L_{\text{std}}^{(i)} = \sqrt{\frac{1}{4(N-1)} \sum_{k \in \{x_1, y_1, x_2, y_2\}} \sum_{j=1}^N \|b_{j,k}^{(i)} - M_{\text{mean}}(\varphi_{i,k})\|^2}, \quad (2)$$

$$L_{\text{mean}}^{(i)} = \frac{1}{4} \sum_{k \in \{x_1, y_1, x_2, y_2\}} [t_k^{(i)} - M_{\text{mean}}(\varphi_{i,k})]^2, \quad (3)$$

$$M_{\text{mean}}(\varphi_{i,k}) = \frac{1}{N} \sum_{j=1}^N b_{j,k}^{(i)}, \quad (4)$$

式中:  $\|\cdot\|$  为  $\text{Smooth}_{L_1}$  损失函数, 其表达式为

$$\text{Smooth}_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| \geq 1 \end{cases}. \quad (5)$$

$L_{\text{std}}$  计算了属于同一个目标的所有预测框的各维坐标的标准方差, 通过优化的方式使得该标准方差趋近于 0, 以降低预测同一目标的框坐标之间的差异。标准方差的优化使得同属于一个目标的所有预测框的回归过程之间产生相互联系, 而不是各自独立的回归过程。如图 2(a) 所示, 起点的差异会造成优化难度的不同, 当难度差异过大时, 回归的结果也会不尽相同。在加入方差损失后, 各框的回归过程之间产生联系。

方差的减小使他们的坐标趋于一致, 进而保证了预测框之间的紧凑性与一致性。 $L_{\text{mean}}$  计算了同属于一个目标的所有预测框坐标的均值与目标框坐标之间的欧氏距离。由于  $L_{\text{std}}$  函数的作用是使得预测同一个目标的框趋于一致, 故易于回归的样本会给难以回归的样本一个朝向回归目标的推动作用, 且难以回归的目标同样会给易于回归的样本一个背离回归目标的拉动作用, 这意味着最终得到的预测框可能存在误差。因此,  $L_{\text{mean}}$  的目的是在保证预测框紧凑性的基础上, 使预测框整体向目标框方向回归, 进而保证预测框的精确性, 防止预测框的精度下降。在加入聚拢损失函数后, 总的损失函数为

$$L_{\text{Loss}} = \alpha L_{\text{obj}} + \beta L_{\text{reg}} + \gamma L_{\text{tight}}, \quad (6)$$

式中,  $L_{\text{obj}}$  为置信度损失, 采用二元交叉熵函数计算;  $L_{\text{reg}}$  为目标框回归损失, 采用 CIoU 函数<sup>[23]</sup> 计算;  $\alpha, \beta, \gamma$  分别为权重系数, 本文设置  $\alpha = 1.0, \beta = 0.5, \gamma = 5^{-4}$ 。

### 3.2 高分辨率特征金字塔

针对目标的重叠程度大造成的彼此难以区分的问题, 提出了高分辨率金字塔如图 3(a) 所示, 其中 SAPH 为空间注意力检测头。高分辨率金字塔的改进针对物理空间与语义信息两个层面。在物理空间层面, 在图片送入主干网络经过若干次下采样后, 分辨率不断减小, 使得高重叠目标在最终的特征金字塔内只能由同一个子特征表示。如图 3(c) 所示, 实线框和虚线框分别属于两个高度重叠的目标, 圆点与三角形点分别是他们的中心点。当两个目标高度重叠时, 他们的中心点也彼此靠近, 若此时特征金字塔分辨率过低, 则两目标会被同一深度特征表示, 难以彼此区分。在语义层面, 重叠程度高意味着目标之间共享像素比例高, 造成被遮挡的目标的深度特征与遮挡者相似度高, 难以彼此区分开来。

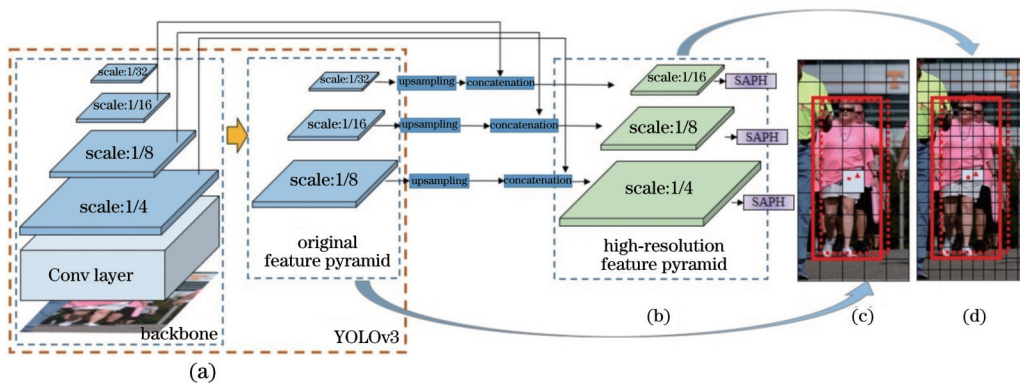


图 3 高分辨率特征金字塔与空间注意检测头在网络中的插入位置示意图。(a) YOLOv3 模型; (b) 高分辨率金字塔; (c) 原始特征金字塔中高重叠目标中心点落在同一个网格内; (d) 高分辨率金字塔中高重叠目标中心点落在不同的网格内

Fig. 3 Schematic diagrams of high-resolution feature pyramid and insertion position of spatial attention prediction head in network. (a) YOLOv3 network; (b) high-resolution feature pyramid; (c) center points of heavily overlapped objects locate in same grid in original feature pyramid; (d) center points of heavily overlapped objects locate in different grids in high resolution feature pyramid

首先, 对原始金字塔特征进行上采样以扩大其分辨率, 在物理空间的层面使得高重叠目标框的中心点

尽量落在不同的子特征区域内。如图 3(d) 所示, 虽然两个目标高度重叠且中心点彼此接近, 但是分辨率的

提升依然可以使得他们由不同的子特征来表示。然而,普通上采样机制是通过简单复制或双线性插值来实现分辨率的扩充的,相邻特征是完全相同或者高度相似的,这不利于高重叠目标在语义层面上的区分。为提升相邻特征之间的差异性,由主干网络引入相同分辨率的浅层特征与上采样后的特征金字塔进行串联融合。虽然与深层特征相比,浅层特征包含较少的目标实例的高级语义信息,但是浅层特征的引入对于高重叠目标的特征提取有两方面的帮助作用。首先,浅层特征的感受野较小,其相邻子特征相较同分辨率的深层特征而言具有更好的区分度。在浅层特征与上采样后的特征金字塔串联之后,相邻的子特征中融入了来自低层特征的具有区分度的细节信息,在增强特征金字塔中相邻子特征间的差异性的同时,保留了深层特征中目标的高级语义信息。其次,因为被遮挡目标的特异性主要来源于其有限的可视像素,而浅层特征由于感受野较小,相较深层特征而言包含更多的被遮挡目标的细节信息,故浅层特征的引入使得特征金字塔直接融入了被遮挡目标的语义信息,这对于提取被遮挡目标的特征是有利的。因此,上述操作不仅可以分别从物理空间层面与语义信息层面上解决高重叠度目标深度特征难以区分的问题,还增强了网络对被遮挡目标的信息提取能力。

### 3.3 基于空间注意力机制的检测头

在金字塔分辨率被扩大之后,特征被分隔成更加密集的网格,而图像中目标的数量是固定的,因此特征上包含目标中心点的子特征与背景占据的子特征的比例被缩小,这对目标置信度的预测提出了更高的要求。虽然 3.2 节提出的高分辨率金字塔在语义信息层面和物理空间层面可以为高重叠度目标提供可分辨的相邻子特征,但是由于其扩大分辨率的方式是不可学习的上采样机制,故相邻子特征仍然具有较高的相似性。对于包含目标中心点的子特征,若其周围子特征与之相似性高,则可能输出相似的置信度预测,因而会产生大量高置信度(大于置信度阈值)的冗余预测框。如图 4 所示,对于 4(a) 中的稀疏目标,在原先低分辨率特征中只会产生一个高置信度的预测框。在经过上采样后,相邻子特征具有相似性,如图 4(a) 图中深色阴影区域所示,则原先的一个高置信度预测最多可能变成四个,如图 4(b) 中深色阴影所示。显然,这些冗余的预测框均属于同一稀疏目标,而大量冗余的预测框会增加后续 NMS 过程的计算量,影响模型的推理时间。针对上述问题,本节提出一种基于空间注意力机制的预测头,利用空间注意力的方式使模型更加专注于目标最有可能出现的子特征,并降低对稀疏目标相邻子特征的关注度,从而过滤掉冗余框,如图 4(c) 中深色阴影区域所示。

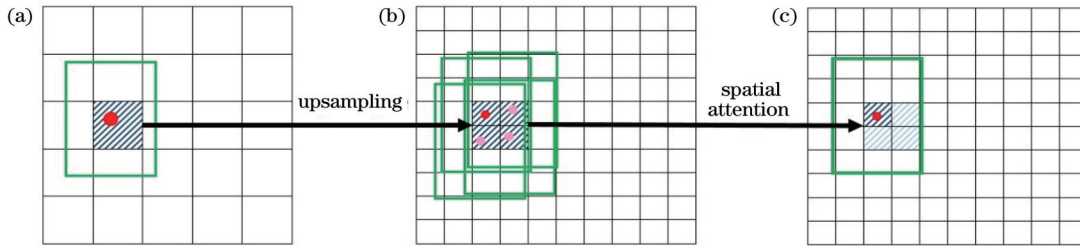


图 4 高分辨率特征金字塔中高置信度冗余框示意图。(a)原始特征金字塔中目标框及其置信度预测;(b)上采样机制产生的置信度预测和冗余预测框;(c)空间注意力机制过滤后的置信度和预测框

Fig. 4 Schematic diagram of redundant bounding boxes with high confidence in high-resolution feature pyramid. (a) Target box in original feature pyramid and its confidence prediction; (b) confidence prediction and redundant prediction boxes generated by upsampling mechanism; (c) confidence and prediction boxes filtered by spatial attention mechanism

文献[24]提出一种通道注意力与空间注意力混合的注意力机制来提升模型提取特征的能力。本文采用其中的空间注意力模块(SAM),如图 5 所示。

该注意力模块首先对输入特征  $F$  在通道方向上分

别进行最大池化(AvgPool,  $M_{AvgPool}$ )与平均池化(MaxPool,  $M_{MaxPool}$ ),并将两次池化的结果在通道方向上进行串联。然后,将串联后的特征进行一次卷积(Conv)操作,并利用 sigmoid 函数  $\sigma(x) = 1/[1 + \exp(-x)]$  将输出映射到 0-1 空间,从而生成空间注意力权重图  $M_s$ 。 $M_s$  的表达式为

$$M_s(F) = \sigma \left\{ f^{(k \times k)} \left\{ [M_{AvgPool}(F); M_{MaxPool}(F)] \right\} \right\} = \sigma \left\{ f^{(k \times k)} \left\{ [F_{avg}^s; F_{max}^s] \right\} \right\}, \quad (7)$$

式中:  $f^{(k \times k)}$  为卷积核大小为  $k \times k$  的卷积层,本文设置  $k=7$ 。 $F_{avg}^s$  为通道方向平均池化后的特征;  $F_{max}^s$  为通道方向最大池化后的特征。该权重图  $M_s$  会给感兴趣区域分配较高的权重,降低无关区域的权重,从而提高网

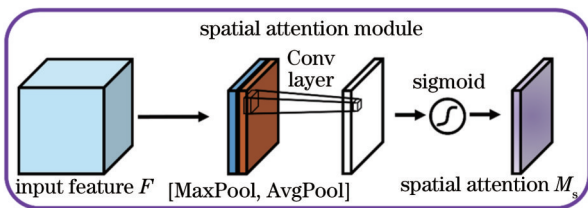


图 5 空间注意力模块

Fig. 5 Spatial attention module



络的专注度,过滤无效信息。

基于上述空间注意力机制,提出了一种 SAPH,如图 3(a)所示,SAPH 外接在高分辨率金字塔特征上,用于减少由高相似度的相邻子特征产生的高置信度的冗余预测框。如图 6(a)所示,SAPH 模块由两个空间注意力残差模块(SARB)与两个卷积层组成,其中 SARB 由两个卷积层与一个空间注意力模块组成,如图 6(b)所示。SARB 模块首先对特征分别进行  $1 \times 1$  和  $3 \times 3$  的两次卷积操作,再将输出特征送入如图 5 所示的空间注意力模块中,生成空间注意力权重图  $M_s$ 。然后,将权重图  $M_s$  与卷积核大小为  $3 \times 3$  的卷积层的输出特征进行逐元素相乘,并将该赋予了不同权重的特征与之前卷积核大小为  $1 \times 1$  的卷积层的输出特征逐元素相加。输出特征  $F'$  的表达式为

$$F' = f^{(1 \times 1)}(F) \oplus \left\{ f^{(3 \times 3)} \left[ f^{(1 \times 1)}(F) \right] \otimes M_s \left\{ f^{(3 \times 3)} \left[ f^{(1 \times 1)}(F) \right] \right\} \right\}, \quad (8)$$

式中: $\otimes$  与  $\oplus$  分别为逐元素相乘符号与逐元素相加符号。

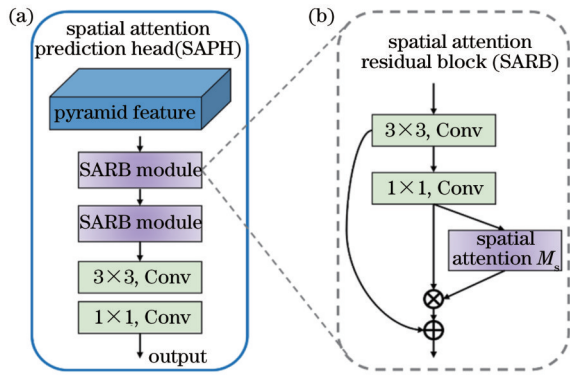


图 6 空间注意力预测头和空间注意力残差块示意图。(a)空间注意力预测头;(b)空间注意力残差块

Fig. 6 Schematic diagrams of spatial attention prediction head and spatial attention residual block. (a) Spatial attention prediction head; (b) spatial attention residual block

空间注意力的思想在于让网络为空间方向的每一个子特征自主学习出一个归一化的权重,从而达到突出有效信息,过滤无效或者干扰信息的目的。在提高了分辨率的金字塔中,由于相邻子特征具有一定的相似性,故不可避免地会造成相邻特征预测框置信度相似的情况,这就使得目标无论是稀疏还是密集,都会产生大量高置信度的冗余预测框,给后续 NMS 过程带来计算负担。为过滤掉稀疏目标的高置信度冗余预测框,同时不影响高重叠目标的预测,引入了空间注意力机制,通过自主学习相应的权重,突出目标本身所处位置的信息,降低目标特征周围不包含目标的相邻子特征的影响,以达到保留正确预测框,过滤冗余预测框的目的。此外,得益于图 5 所示的空间注意力模块的轻量化设计,虽然 SAPH 模块中仅引入了可以忽略不计的额外参数量,但是后续实验证明其带来了可观的性

能提升。

## 4 实验与分析

### 4.1 数据集和评价指标

本文在 CrowdHuman 数据集上对所提算法的有效性进行实验验证。CrowdHuman 是一个大型的公开行人检测数据集。相比其他的行人检测数据集,CrowdHuman 包含的行人实例密度更大,平均每张图片中 IoU 大于 0.5 的框有 2.4 个。因此,基于该数据集的实验结果在遮挡目标检测问题上更加有说服力。本文所有的实验结果都是在验证集上测试得来的。

该数据集提供三个评价指标,分别是平均精度( $M_{AP}$ )、丢失率( $M_{MR}$ )和召回率( $M_{Recall}$ )。平均精度是目标检测领域中最主流的评价指标,它同时考虑了召回率和精准度,可以全面地评价一个检测模型的性能,平均精度的值越大代表性能越好。丢失率是行人检测任务中常用的评价指标,表征了在不同假阳率阈值下漏检率的大小。丢失率对于假阳率非常敏感,丢失率越低代表性能越好。召回率是所有被正确预测的目标数量占目标总数的比例。召回率越大代表性能越好,其表达式为

$$M_{Recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (9)$$

式中, $N_{TP}$  为真阳样本数量; $N_{FN}$  为假阴样本数量。

### 4.2 实验设置

原始 YOLOv3 的锚框匹配机制是每个目标匹配一个锚框,这被其后续改进模型 YOLOv4 等证明并不是最优选择,因此将 YOLOv3 的锚框匹配机制改为一个目标可以匹配多个锚框,可表示为

$$M_{anchor} = \begin{cases} M_{positive}, & \max\left(\frac{w_a}{w_t}, \frac{w_t}{w_a}, \frac{h_a}{h_t}, \frac{h_t}{h_a}\right) < 4.0, \\ M_{negative}, & otherwise \end{cases}, \quad (10)$$

式中, $w_a$  和  $h_a$  分别是当前锚框  $M_{anchor}$  的宽与高; $w_t$  和  $h_t$  分别为当前目标框的宽与高; $M_{positive}$  为正样本; $M_{negative}$  为负样本。式(10)的含义为:当某个锚框与当前目标的宽之比或高之比小于 4.0 时,认为该锚框与当前目标相匹配,判别其为正样本;否则,判别该锚框为负样本。

设置训练图片尺寸为  $640 \text{ pixel} \times 640 \text{ pixel}$ ,测试图片尺寸的最长边为  $672 \text{ pixel}$ ,训练时每批数据包含 96 张图片,总训练轮次数(epoch)为 300。采用随机梯度下降(SGD)优化器对网络进行迭代优化,用余弦学习率<sup>[25]</sup>调整策略改变学习率,设置初始学习率为 0.010,终止学习率为 0.002,动量设置为 0.937,权值衰减系数为 0.0005。以 Mosaic 和 Mixup<sup>[19]</sup>为图像增强方式,不采用多尺度训练与多尺度测试。实验模型以 PyTorch1.7.1 为框架搭建,并采用 4 张 NVIDIA RTX 3090 GPU 以并行的方式进行训练。

### 4.3 消融实验

以 YOLOv3 为基线模型对所提出的改进算法进

行消融实验。在实验中,原始 YOLOv3 模型、YOLOv3+HRFP 模型和 YOLOv3+HRFP+SAPH 模型均采用 4.2 节中的实验设置进行训练。分别以上述三个模型训练完成的权重为基础,加入 Tight Loss 函数后进行微调,微调的轮次数为 40。为有效验证聚拢损失对假阳率的抑制作用,采用 Adaptive NMS 抑制算法来处理冗余预测框,设置自适应 NMS 阈值的下限为 0.5,即当两个候选框之间的 IoU 值小于 0.5 时,

按照传统的 NMS 方法来抑制冗余框。否则,根据预测框的密度来自适应调整 NMS 阈值。需要注意的是,Adaptive NMS 算法依赖由网络预测得出的目标的密度信息来调整 NMS 阈值,故统一采用由标签信息计算得到的目标密度值作为自适应 IoU 阈值,以防止不同模型在密度预测能力上的偏差对消融实验的公平性造成影响。实验结果如表 1 所示,表中加粗数据为对应项中的最佳值。

表 1 基于 YOLOv3 的消融实验结果  
Table 1 Results of ablation experiments based on YOLOv3

Tight loss	HRFP	SAPH	$M_{AP} / \%$	$M_{MR} / \%$	$M_{Recall} / \%$	Reasoning speed / (frame·s <sup>-1</sup> )
			86.84	49.52	90.68	38
✓			86.86	49.38	90.65	38
	✓		89.53	48.30	93.80	26
✓	✓		89.57	<b>48.14</b>	93.72	26
	✓	✓	89.72	48.34	93.81	32
✓	✓	✓	<b>89.75</b>	48.28	<b>93.88</b>	32

由表 1 结果可以看出,HRFP 模块可以有效提升 YOLOv3 模块对于密集遮挡目标的检测性能。提高的分辨率可以为高重叠目标生成相邻且可分辨的深度特征,进而提升目标的召回率。相较原始 YOLOv3 模型,HRFP 模块将召回率提升了 3.12 个百分点,由于平均精度是与召回率高度相关的,因此平均精度也提升了 2.69 个百分点。在加入了 Tight Loss 后,可以看到,无论是原始 YOLOv3 模型还是加入了 HRFP 后的模型,丢失率都有一定程度的降低。其中,相较原始 YOLOv3,加入 Tight Loss 后,丢失率降低了 0.14 个百分点,相较 YOLOv3+HRFP 模型,Tight Loss 令丢失率降低了 0.16 个百分点。丢失率改善的同时,平均精度与召回率基本不变,这说明此时丢失率的改善是得益于 Tight Loss 使得预测框更为紧凑而带来的假阳率下降。

虽然 HRFP 模块可以显著提升模型的检测指标,但是由于其提高了特征的分辨率,故高置信度(大于置信度阈值)的冗余预测框会数量增加,增大了 NMS 过程的计算量。由表 1 可知,相较原始 YOLOv3 模型,HRFP 模块使得推理速度下降了 12 frame/s。在加入 SAPH 后,不仅平均精度提升了 0.19 个百分点,模型

的推理速度还提升了 6 frame/s。这一提升主要得益于注意力机制使得检测头更加关注目标本身,不仅增加了对被遮挡目标的关注度,还过滤了稀疏目标的冗余预测框,进而降低了 NMS 过程的计算成本。相较原始 YOLOv3 模型,虽然加入 SAPH 和 HRFP 后最终推理速度下降了 6 frame/s,但是平均精度提升了 2.91 个百分点,召回率提升了 3.20 个百分点,丢失率下降了 1.24 个百分点。为证明推理速度的提升是因为减少了冗余预测框,如表 2 所示,统计了在整个验证集中,不同置信度阈值( $M_{Thres}$ )下预测框的总量。其中,H 表示 HRFP 模块,S 表示 SAPH 模块。表 2 中最后一行 Variation 显示的是在不同阈值下 SAPH 模块对预测框数量的影响。由第一列可以看到,置信度大于 0 的预测框数量(输出的预测框总量)大幅下降,表明 SAPH 模块可以有效减少冗余预测框的数量,而置信度大于 0.2 的情况下框的数量却有少量的增加。以上现象证明 SAPH 模块提升了网络对部分目标的判别能力,即降低了非目标区域的置信度,提升了目标区域的置信度,这主要得益于注意力机制对有用信息的凸显作用和对干扰信息的过滤作用。

表 2 空间注意力机制对预测框总数的影响  
Table 2 Influence of spatial attention mechanism on number of predicted boxes

Mode	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
H	374911	141925	107315	90477	77812	65193	51491	36305	21742	8725
H+S	358712	141581	109261	92790	79855	66961	53084	37489	22575	9154
Variation	-16199	-344	+1946	+2313	+2043	+1768	+1593	+1184	+883	+429

图 7 展示了 SAPH 模块对目标置信度的影响。图 7 中圆圈圈出部分表示目标置信度较高的区域,即此区域出现目标中心点的概率较大。如 7 图中箭头所指区域所示,在加入了 SAPH 模块后,目标中心高置信度区域面积减小,并且明显向着目标框中心点收缩,这正是高置信度(大于置信度阈值)冗余框数量会减少的

原因。

#### 4.4 对比实验

为综合验证所提算法在不同 NMS 算法下对密集遮挡目标检测性能的提升效果,对比了在另外两种 NMS 算法下模型的检测性能,分别是传统的 NMS 和 Soft NMS,以证明所提算法对不同 NMS 方法的鲁棒

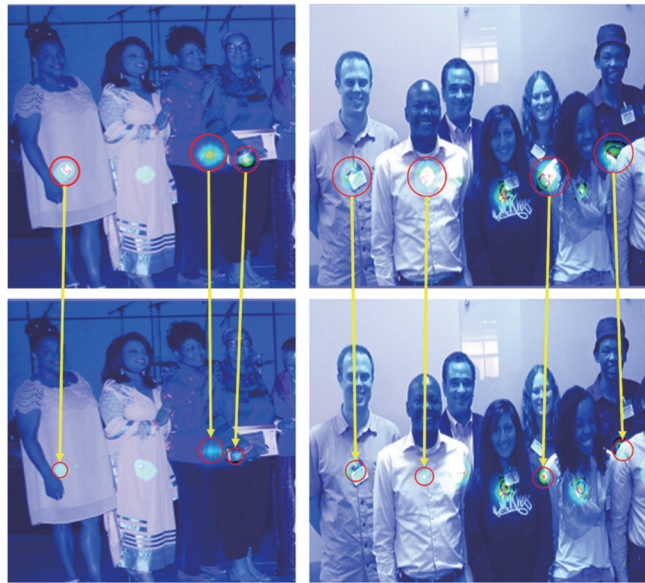


图 7 目标置信度热力图

Fig. 7 Heat map of target confidence

性。测试结果如表 3 所示,其中 T 表示 Tight Loss 函数,粗体表示对应项中的最佳数据。

表 3 不同 NMS 方法下模型的性能对比

Table 3 Performance comparison of models under different NMS methods unit: %

Index	NMS methods					
	Original NMS		Soft NMS		Adaptive NMS	
	YOLOv3	H+T+S	YOLOv3	H+T+S	YOLOv3	H+T+S
$M_{AP}$	84.97	88.38	89.04	<b>91.41</b>	86.84	89.75
$M_{MR}$	50.39	49.14	50.26	49.07	49.52	<b>48.28</b>
$M_{Recall}$	88.85	92.66	94.97	<b>97.32</b>	90.68	93.88

由表 3 可知,在同时应用所提的三种改进算法后:改进后的 YOLOv3 在传统 NMS 方法下使得平均精度和召回率分别提升了 3.41 个百分点和 3.81 个百分点,使得丢失率降低了 1.25 个百分点;在 Soft NMS 方法下使得平均精度和召回率分别提升了 2.37 个百分点和 2.35 个百分点,使得丢失率降低了 1.19 个百分点;在 Adaptive NMS 方法下使得平均精度和召回率分别提升了 2.91 个百分点和 3.20 个百分点,使得丢失率下降了 1.24 个百分点。以上结果表明所提算法对不同 NMS 方法具有鲁棒性。

此外,将所提算法与当前遮挡目标检测领域中的先进算法进行对比,以证明改进后的 YOLOv3 模型在密集遮挡目标检测上的先进性。对比结果如表 4 所示,表中加粗的结果为各项中最佳结果。由表 4 可知,将所提的改进算法应用于 YOLOv3 检测器后,在 Adaptive NMS 和 Soft NMS 情况下,三项评价指标均达到最优,全面超过表中所列现有先进算法。在 Original NMS 情况下,其在平均精度与丢失率两项指标上均优于表中其他先进算法。以上结果表明,改进

后的 YOLOv3 模型在遮挡行人检测上具有先进性。

表 4 所提算法与当前先进遮挡行人检测算法的对比

Table 4 Comparison between proposed algorithm and current advanced occluded pedestrian detection algorithms unit: %

Algorithm	NMS Method	$M_{AP}$	$M_{MR}$	$M_{Recall}$
RetinaNet <sup>[15]</sup>	Original NMS	78.33	65.22	94.13
IterDet (RetinaNet) <sup>[26]</sup>	Original NMS	84.77	56.21	91.49
Faster RCNN <sup>[15]</sup>	Original NMS	83.07	52.35	90.57
PS-RCNN <sup>[27]</sup>	Original NMS	86.05		93.77
IterDet (Faster RCNN) <sup>[26]</sup>	Original NMS	88.08	49.44	<b>95.80</b>
YOLOv3+H+S+T	Original NMS	<b>88.38</b>	<b>49.14</b>	92.66
RetinaNet <sup>[15]</sup>	Soft NMS	78.10	66.34	95.37
Faster RCNN <sup>[15]</sup>	Soft NMS	83.92	51.97	91.73
YOLOv3+H+S+T	Soft NMS	<b>91.41</b>	<b>49.07</b>	<b>97.32</b>
RetinaNet <sup>[15]</sup>	Adaptive NMS	79.67	63.03	94.77
Faster RCNN <sup>[15]</sup>	Adaptive NMS	84.71	49.73	91.27
YOLOv3+H+S+T	Adaptive NMS	<b>89.75</b>	<b>48.28</b>	<b>93.88</b>

#### 4.5 定性结果

本节展示了所提算法对检测性能提升的定性结果。图 8 所示为 Tight Loss 函数对检测框紧凑性提升的定性结果。图 8 中展示的是未经 NMS 后处理的结果,即所有置信度大于阈值的框都被保留了下来,以展示对应于同一目标的不同预测框的紧凑性。其中,图 8(b)、(d) 中的框为原始 YOLOv3 的检测结果,图 8(a)、(c) 中的框为使用 Tight Loss 进行微调之后的结果。可以看到,经过 Tight Loss 函数微调后,目标的预



测框由原来的松散变得紧凑且一致。

图 9 展示了所提算法对模型检测性能的提升效果。其中,图 9(a)、(c)中的框为改进算法的预测结果,图 9(b)、(d)为原始 YOLOv3 的检测结果。图 9 中所有预测框都是经过 Adaptive NMS 处理后保留下来的。

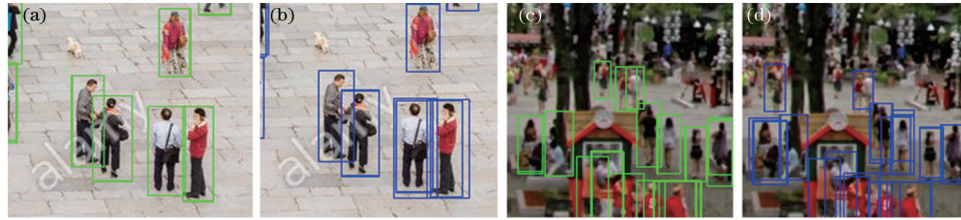


图 8 Tight Loss 对模型性能的影响。(a)(c)经过 Tight Loss 微调之后的预测结果;(b)(d)未经 Tight Loss 调整的预测结果  
Fig. 8 Influence of Tight Loss function on model performance. (a)(c) Prediction results after Tight Loss fine-tuning; (b)(d) prediction results without Tight Loss adjustment

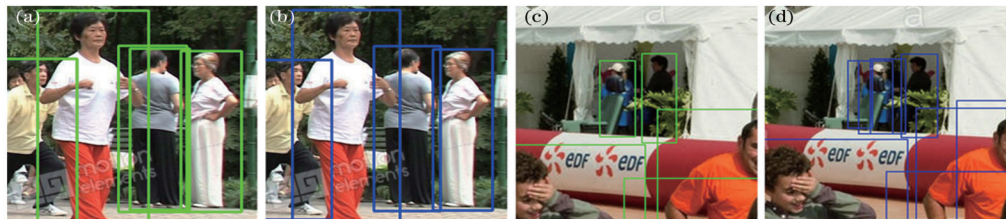


图 9 模型综合性能对比。(a)(c)改进 YOLOv3 的预测结果;(b)(d)原始 YOLOv3 的预测结果  
Fig. 9 Comparison of comprehensive performance of models. (a)(c) Prediction results generated by improved YOLOv3; (b)(d) prediction results generated by original YOLOv3

## 5 结 论

针对当前基于深度学习的检测器在面对密集遮挡目标场景时表现出的不足之处,提出了三点创新:1)设计了一种 Tight Loss 函数以提升同一目标的多个预测框的紧凑性,从而降低假阳率;2)设计了一种高分辨率金字塔,以提升高重叠目标深度特征的可区分性;3)提出了一种空间注意力检测头,利用注意力机制提升网络对于目标的专注度,减少因为相邻子特征的高相似性而产生的冗余预测框数量,进而提升模型推理速度。为验证所提算法的有效性,在密集行人数据集 CrowdHuman 上进行了实验。实验结果表明,高分辨率金字塔可以为严重重叠目标提供可区分的深度特征,这体现在召回率的显著提升上。此外,Tight Loss 在高分辨率金字塔的基础上进一步提升了预测框的紧凑性,从而降低了检测结果中假阳目标的数量,这体现在对丢失率的有效抑制上。SAPH 模块可以有效减少冗余预测框的数量,这体现在对推理速度的显著提升上。综上所述,所提算法可以有效提升模型对遮挡目标的检测性能。

### 参 考 文 献

[1] 孙迎春,潘树国,赵涛,等.基于优化 YOLOv3 算法的交通灯检测[J].光学学报,2020,40(12):1215001.  
Sun Y C, Pan S G, Zhao T, et al. Traffic light detection

从图 9(a)、(b)的对比中可以看出所提算法对模型中严重遮挡目标的检测性能的提升效果。从图 9(c)、(d)的对比中可以看出 Tight Loss 函数使得预测框更加紧致后带来的假阳率下降的效果。

based on optimized YOLOv3 algorithm[J]. Acta Optica Sinica, 2020, 40(12): 1215001.

- [2] Mou L C, Zhu X X. Spatiotemporal scene interpretation of space videos via deep neural network and tracklet analysis[C]//IEEE International Geoscience and Remote Sensing Symposium, July 10-15, 2016, Beijing, China. New York: IEEE Press, 2016: 1823-1826.
- [3] Kopsiaftis G, Karantzas K. Vehicle detection and traffic density monitoring from very high resolution satellite video data[C]//IEEE International Geoscience and Remote Sensing Symposium, July 26-31, 2015, Milan, Italy. New York: IEEE Press, 2015: 1881-1884.
- [4] 郭烈,马跃,岳明,等.驾驶特性的识别评估及其在智能汽车上的应用综述[J].交通运输工程学报,2021,21(2):7-20.  
Guo L, Ma Y, Yue M, et al. Overview of recognition and evaluation of driving characteristics and their applications in intelligent vehicles[J]. Journal of Traffic and Transportation Engineering, 2021, 21(2): 7-20.
- [5] 王润民,朱宇,赵祥模,等.自动驾驶测试场景研究进展[J].交通运输工程学报,2021,21(2):21-37.  
Wang R M, Zhu Y, Zhao X M, et al. Research progress on test scenario of autonomous driving[J]. Journal of Traffic and Transportation Engineering, 2021, 21(2): 21-37.
- [6] 沈岷,李舜酩,柏方超,等.路面车辆实时检测与跟踪的视觉方法[J].光学学报,2010,30(4):1076-1083.

- Shen H, Li S M, Bo F C, et al. On road vehicles real-time detection and tracking using vision based approach [J]. *Acta Optica Sinica*, 2010, 30(4): 1076-1083.
- [7] 赵亮, 胡杰, 刘汉, 等. 基于语义分割的深度学习激光点云三维目标检测 [J]. *中国激光*, 2021, 48(17): 1710004.
- Zhao L, Hu J, Liu H, et al. Deep learning based on semantic segmentation for three-dimensional object detection from point clouds[J]. *Chinese Journal of Lasers*, 2021, 48(17): 1710004.
- [8] 胡杰, 刘汉, 徐文才, 等. 基于三维激光雷达的道路障碍物目标位姿检测算法 [J]. *中国激光*, 2021, 48(24): 2410001.
- Hu J, Liu H, Xu W C, et al. Position detection algorithm of road obstacles based on 3D LiDAR[J]. *Chinese Journal of Lasers*, 2021, 48(24): 2410001.
- [9] 农元君, 王俊杰. 基于嵌入式的遥感目标实时检测方法 [J]. *光学学报*, 2021, 41(10): 1028001.
- Nong Y J, Wang J J. Real-time object detection in remote sensing images based on embedded system[J]. *Acta Optica Sinica*, 2021, 41(10): 1028001.
- [10] 农元君, 王俊杰, 赵雪冰, 等. 遥感目标空间关系检测方法 [J]. *光学学报*, 2021, 41(16): 1628001.
- Nong Y J, Wang J J, Zhao X B, et al. Spatial relationship detection method of remote sensing objects [J]. *Acta Optica Sinica*, 2021, 41(16): 1628001.
- [11] 姬东飞, 丁学明. 基于特定区域去相关的行人检测算法 [J]. *信息与控制*, 2020, 49(4): 404-413.
- Ji D F, Ding X M. Specific region decorrelation feature channel on pedestrian detection[J]. *Information and Control*, 2020, 49(4): 401-413.
- [12] 宣勇, 韩超, 沙文瀚. 改进的 Tiny YOLOv4 算法及在行人检测中的应用 [J]. *激光与光电子学进展*, 2022, 59(12): 1215002.
- Xuan Y, Han C, Sha W H. Improved Tiny Yolov4 algorithm in pedestrian detection[J]. *Laser & Optoelectronics Progress*, 2022, 59(12): 1215002.
- [13] 刘铁根, 刘琨, 戴林, 等. 光电信息事件识别感知关键技术研究进展 [J]. *光学学报*, 2021, 41(1): 0106002.
- Liu T G, Liu K, Dai L, et al. Research progress of key technologies in recognition sensing for opto-electronic information and event[J]. *Acta Optica Sinica*, 2021, 41(1): 0106002.
- [14] Bodla N, Singh B, Chellappa R, et al. Soft-NMS: improving object detection with one line of code[C]//IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 5562-5570.
- [15] Liu S T, Huang D, Wang Y H. Adaptive NMS: refining pedestrian detection in a crowd[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 6452-6461.
- [16] Ma J, Wan H L, Wang J X, et al. An improved one-stage pedestrian detection method based on multi-scale attention feature extraction[J]. *Journal of Real-Time Image Processing*, 2021, 18(6): 1965-1978.
- [17] 邹梓吟, 盖绍彦, 达飞鹏, 等. 基于注意力机制的遮挡行人检测算法 [J]. *光学学报*, 2021, 41(15): 1515001.
- Zou Z Y, Gai S Y, Da F P, et al. Occluded pedestrian detection algorithm based on attention mechanism[J]. *Acta Optica Sinica*, 2021, 41(15): 1515001.
- [18] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2021-03-05]. <https://arxiv.org/abs/1804.02767>.
- [19] Alexey B, Wang C, et al. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2021-05-02]. <https://arxiv.org/abs/2004.10934>.
- [20] Wang C Y, Bochkovskiy A, Liao H Y M. Scaled-YOLOv4: scaling cross stage partial network[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13024-13033.
- [21] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. (2021-07-18) [2021-02-05]. <https://arxiv.org/abs/2107.08430>.
- [22] Shao S, Zhao Z J, Li B X, et al. CrowdHuman: a benchmark for detecting human in a crowd[EB/OL]. (2018-04-30) [2021-02-04]. <https://arxiv.org/abs/1805.00123>.
- [23] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12993-13000.
- [24] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [25] Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts[EB/OL]. (2016-08-13)[2021-02-05]. <https://arxiv.org/abs/1608.03983>.
- [26] Rukhovich D, Sofiiuk K, Galeev D, et al. IterDet: iterative scheme for object detection in crowded environments[M]//Torsello A, Rossi L, Pelillo M, et al. *Structural, syntactic, and statistical pattern recognition*. Lecture notes in computer science. Cham: Springer, 2021, 12644: 344-354.
- [27] Ge Z, Jie Z Q, Huang X, et al. PS-RCNN: detecting secondary human instances in a crowd via primary object suppression[C]//2020 IEEE International Conference on Multimedia and Expo, July 6-10, 2020, London, UK. New York: IEEE Press, 2020: 19870687.