

基于改进 YOLOv5s 算法的交通信号灯检测^①



王 军^{1,2}, 葛宝康¹, 程 勇¹

¹(南京信息工程大学 软件学院, 南京 210044)

²(南京信息工程大学 科技产业处, 南京 210044)

通信作者: 葛宝康, E-mail: 202212210051@nuist.edu.cn

摘 要: 针对交通信号灯检测中目标尺度小、检测精度低的问题, 提出一种改进 YOLOv5s 的交通信号灯检测算法. 首先, 构建一种特征金字塔模块 RSN-BiFPN, 充分融合不同尺度的交通信号灯特征, 以减少目标漏检和误检. 其次, 引入新的特征融合层和预测头, 提高网络对小目标的感知性能, 增强检测准确性; 最后, 采用 EIoU 函数优化损失, 加快网络收敛速度. 通过在 S2TLD 公开数据集上进行的大量的实验结果表明, 本文所提方法相较于基础网络, 精确率提升 4.1%, 达 96.1%; 召回率提升 3%, 达 95.9%; 平均精确度提升 1.9%, 达 96.5%. 同时, 改进后的算法实现了更快的检测速度, 达每秒 22.7 帧, 本文方法有效实现交通信号灯快速、准确地检测, 可广泛应用于交通道路中信号灯分析相关研究.

关键词: 交通信号灯检测; YOLOv5s; 小目标; 特征金字塔; EIoU 损失函数

引用格式: 王军, 葛宝康, 程勇. 基于改进 YOLOv5s 算法的交通信号灯检测. 计算机系统应用, 2023, 32(12): 243-252. <http://www.c-s-a.org.cn/1003-3254/9319.html>

Traffic Light Detection Based on Improved YOLOv5s Algorithm

WANG Jun^{1,2}, GE Bao-Kang¹, CHENG Yong¹

¹(School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China)

²(Science and Technology Industry Division, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: Aiming at the small target scale and low detection accuracy in traffic signal detection, this study proposes a traffic signal detection algorithm based on improved YOLOv5s. Firstly, a feature pyramid module RSN-BiFPN is constructed to fully integrate traffic signal features of different scales to reduce target missed detection and false detection. Secondly, a new feature fusion layer and prediction head are introduced to improve the perception performance of the network for small objects and enhance detection accuracy. Finally, the EIoU function is adopted to optimize the loss and accelerate network convergence. Experiments conducted on the public dataset S2TLD show that compared with the basic network, the precision rate of the proposed method is increased by 4.1% at 96.1%, the recall rate is 95.9% with an increase of 3%, and the average precision is increased by 1.9%, reaching 96.5%. Meanwhile, the improved algorithm achieves a faster detection speed of 22.7 frames per second. The proposed method can realize rapid and accurate detection of traffic lights and can be widely employed in the research on analyzing traffic lights.

Key words: traffic light detection; YOLOv5s; small target; feature pyramid; EIoU loss function

近年来, 随着自动驾驶技术的快速发展, 对于自动驾驶场景中行人、交通标志等目标的识别研究日益增

加. 其中, 交通信号灯检测是一项关键任务, 其利用车载摄像头拍摄的照片和视频来对信号灯进行分析. 然

① 基金项目: 国家自然科学基金 (41875184, 41975183)

收稿时间: 2023-05-21; 修改时间: 2023-06-26; 采用时间: 2023-07-07; csa 在线出版时间: 2023-09-19

CNKI 网络首发时间: 2023-09-21

而,在获取的图像中存在交通信号灯目标尺度过小和有用特征稀缺的问题,这增加了目标检测误检和漏检的可能性.因此,迫切需要一种准确度高且检测速度快的算法来解决这一问题.

深度学习技术在计算机视觉领域取得了巨大的成功,特别是在目标检测领域.与早期手工设计的方法相比,现阶段基于深度学习的目标检测算法已经成为主流,并可以分为以下两类:(1)传统的两阶段目标检测算法,如Faster R-CNN^[1]、VFNet^[2]、CenterNet2^[3]等,这些方法首先需要有一个网络来定位图像中可能的物体区域,随后利用另一个网络对物体进行分类.尽管以Faster R-CNN为代表的两阶段检测器在目标检测任务中具有高精度的优点,但由于检测速度较慢,难以满足交通信号灯检测实时性的要求.(2)单阶段目标检测算法包括YOLO9000^[4]、YOLOv3^[5]、YOLOv4^[6]、ScaledYOLOv4^[7]、YOLOv5、YOLOX^[8]、FCOS^[9]、DETR^[10]等.YOLO系列^[4-8,11]是一种典型的单阶段目标检测算法,检测速度快、准确率高,因此成为交通信号灯检测任务的首选算法.

尽管目标检测算法在计算机视觉领域取得了一定成绩,但在小目标检测方面仍存在一定问题与难点.现阶段,小目标检测多以主流目标检测算法为基础,从数据增强、检测层、检测头、特征提取和注意力机制等方面提升小目标的检测性能.例如,Wang等^[12]通过使用AF-FPN替换原有的特征金字塔,提高了YOLOv5对多尺度目标的检测性能;同时,还提出了一种自动学习的数据增强方法,通过丰富数据集来提高模型的适应性和鲁棒性.Chen等^[13]引入了专门用于检测小目标的检测头和增强注意力,检测精度提高6.68%.Wang等^[14]提出了SG-YOLO算法,结合SE注意力模块和Ghost模块,使用SG-Bottleneck结构,参数量减少32.1%.Chu等^[15]使用Trans模块和LD-Head减少冗余特征干扰,提高小目标检测性能.Bi等^[16]提出了基于YOLOv4的YOLO-RFB算法,通过剪枝主干和引入虫洞卷积层等方法,成功提高了检测交通标志中小目标样本的效果.

上述方法虽有效提高小目标检测精度,然而小目标对象分辨率低、像素占比小,在使用卷积网络进行特征提取时,所提供信息有限,不足以对小目标进行表达.同时,随着卷积神经网络多次下采样,特征图尺寸不断减小,小目标特征信息难以提取,导致检测过程中

出现严重的误检、漏检.

为此,本文提出了一种改进YOLOv5s的交通信号灯检测算法,以提高小目标检测平均精确度并降低漏检率和误检率.首先,引入新的特征融合层和检测头,增强算法的感知能力,尤其是对小目标的检测能力和在光照变化较大的场景下的适应性.其次,采用多尺度特征融合网络RSN-BiFPN,通过融合不同尺度的特征信息,实现了更准确的交通信号灯定位,降低漏检和误检的风险.最后,对定位损失函数进行改进,更好地约束模型的学习过程,提升算法的准确性.综合上述改进措施,改进后的算法在误检率和漏检率方面都显著降低,从而提高了精确度.

1 YOLOv5s 原理

YOLOv5是一种广泛应用于目标检测的深度学习模型,包括4个版本:YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x.为了在速度和精度之间取得平衡,本文选择参数量小、速度快的YOLOv5s网络来进行交通信号灯检测.YOLOv5s由输入层、骨干网络、颈部网络和头部网络组成,其网络结构如图1所示.

1.1 输入端

输入端采用3种数据增强技术,分别为马赛克数据增强、自适应锚点计算和自适应图像尺寸处理.马赛克增强通过随机裁剪来扩充数据集数量,从而提高网络对小目标的检测能力.自适应锚点计算针对不同数据集生成最佳预测框,提高模型检测精度.自适应图像尺寸处理采用添加最小黑边和缩放到标准尺寸的方式,避免特征张量与全连接层不匹配的问题,进而提高模型的效率和稳定性.

1.2 骨干网络

骨干网络采用了Focus下采样、CSP1_X和SPP结构来提取图像特征.Focus模块对输入图像进行下采样和通道增加,生成12维特征图的输出.接下来,通过使用3×3卷积核从特征图中提取32维特征层.通过Focus下采样的方法在提高网络性能的同时,保留了图像信息.CSP1_X采用残差结构,增加了层与层之间反向传播的梯度值,避免梯度消失问题,并获得更细粒度特征信息.SPP模块将任意大小的特征图转换为固定大小的特征向量,增强模型感受野.

1.3 颈部网络

颈部网络采用CSP2_X、FPN和PAN^[17]结构,进

一步提取和整合来自骨干网络的特征信息. CSP2_X 结构通过插入两个残差块来提高网络的特征融合能力, 使网络在保持丰富的特征信息的同时减少计算量. 为了解决多尺度物体检测问题, 采用 FPN 和 PAN 结构

对特征金字塔网络进行扩充和路径聚合. FPN 从上到下传递深层语义特征, 而 PAN 从下到上传递目标的位置信息. 这种自上而下和自下而上的特征信息融合方法可以有效地识别不同尺度的目标.

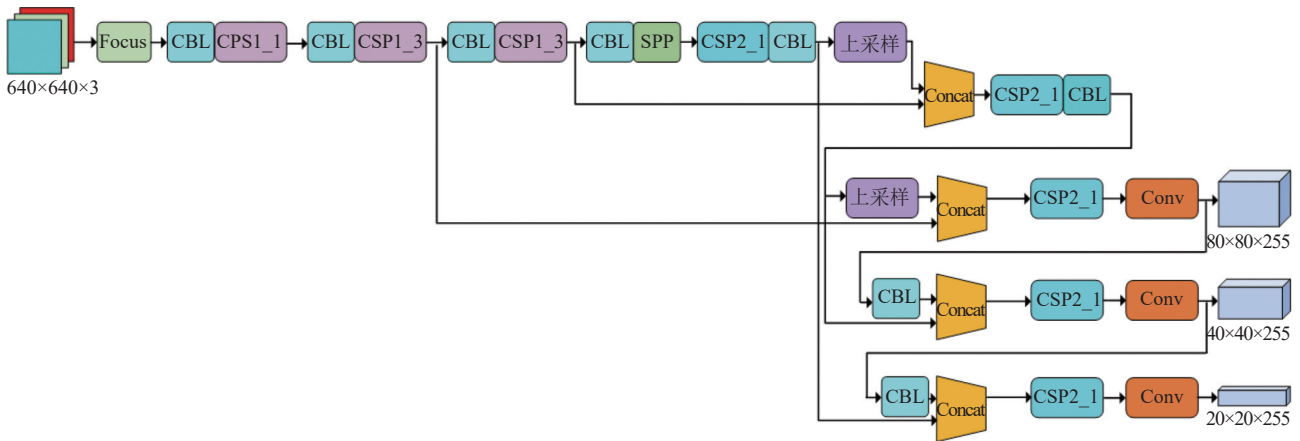


图1 YOLOv5s 网络结构

1.4 头部网络

头部网络采用 bounding box 损失和 NMS 非极大值抑制以提高目标检测的准确性. Bounding box 损失用于训练目标检测模型, 衡量预测框与真实标签框之间的差异. NMS 则用于后处理阶段, 对重叠的检测框进行筛选和抑制, 从而得到最终的目标检测结果. 头部网络由 3 个检测层组成, 每个层生成不同大小的特征图, 用于检测不同大小的目标物体. 每个检测层都输出相应的向量, 这些向量被组合在一起, 生成最终的目标检测结果, 以实现目标对象的精确定位和分类.

2 算法优化

本文基于 YOLOv5s 模型进行改进, 首先, 在保留原模型的 3 个检测层的基础上, 添加了一层具有更大尺度特征的小目标检测层, 以获取更多特征信息, 增强小目标的检测能力. 其次, 将 PANet 特征金字塔替换为本文提出的 RSN-BiFPN 模块, 并应用在特征融合层, 以增强特征的多尺度表达, 从而降低漏检率和误检率. 最后, 使用 EIou 损失函数作为边界框损失函数, 以提升检测精度和速度. 改进后的 YOLOv5s 网络结构如图 2 所示.

2.1 目标检测层优化

针对交通信号灯尺度较小的问题, 通过增加小目

标检测层可以有效减少小目标的漏检率. 在 YOLOv5s 网络中, 特征融合层仅包含第 3、4、5 层的特征信息. 该部分进行两次上采样和两次下采样, 以融合深层特征与浅层特征. 特征融合完成后, 在检测层输出大小为 80×80 、 40×40 、 20×20 的 3 种不同尺度的特征图. 特征图尺度越大, 感受野越小, 包含目标位置信息和局部特征细节越详细, 有利于检测小目标; 特征图尺度越小, 感受野越大, 语义信息越丰富, 但局部细节不明显, 适合检测大目标. 为了提高小目标检测的平均精度, 同时避免由于下采样导致目标的局部细节特征损失过大、小目标信息丢失, 本文引入了一个原本未进行融合的 160×160 的特征图到检测层. 该层经过深层特征的传递以及与浅层特征的融合, 包含更多的小目标轮廓信息以及位置信息, 有利于小目标的定位与识别, 降低小目标的漏检和误检. 同时, 在特征融合网络中增加了一次上采样和一次下采样操作, 从而将最后输出的检测层增加至 4 层. 增加检测层后, 输出的预测框数量也从 9 个相应地增加到 12 个, 其中新增的 3 个预测框长宽比不同且专为小目标检测而设计. 通过这些改进, 本文得到了一个具备四头预测器结构的模型. 该模型在检测微小物体和多尺度物体方面表现出色, 有效提高了对小目标的检测性能.

2.2 RSN-BiFPN 特征金字塔

YOLOv5s 模型采用了 PANet 特征金字塔结构来

解决目标检测中尺度变化的问题. PANet 特征金字塔通过自顶向下和自底向上的路径以及横向连接的方式,

将不同层级的特征进行融合. 其特征融合过程如图 3 所示.

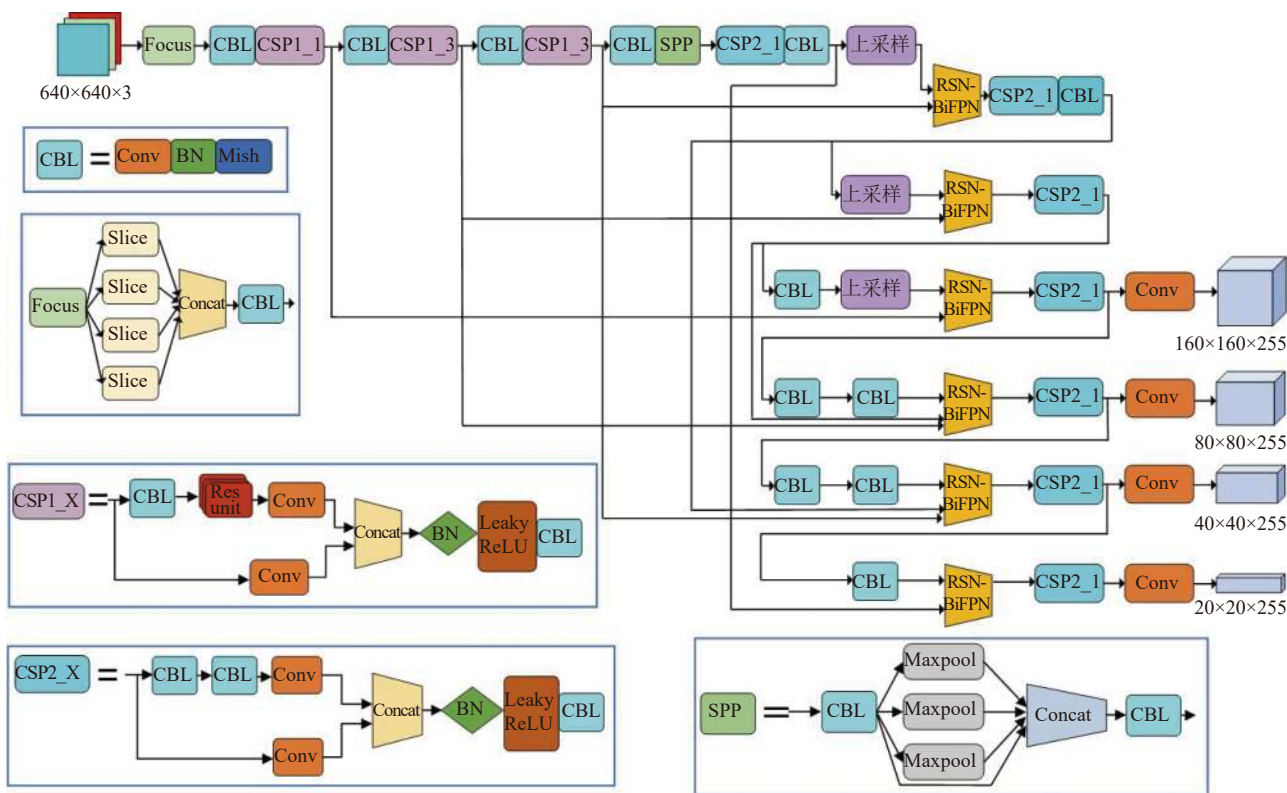


图 2 改进 YOLOv5s 网络结构

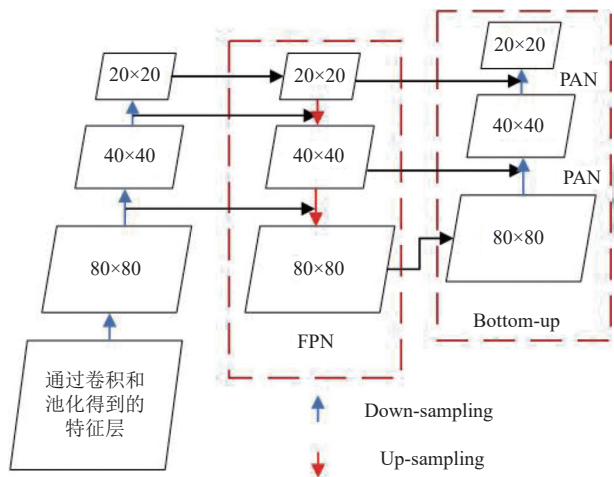


图 3 特征融合过程

首先, YOLOv5s 使用一系列卷积层来提取不同尺度的特征图, 这些特征图具有不同的分辨率, 代表了图像中不同尺度的特征信息. 这些特征图被送入 PANet 进行特征融合.

其次, 自顶向下路径通过对低层级特征图进行下

采样, 得到一组具有相同通道数的特征图, 这组特征图被称为 bottom-up 特征. 同时采用融合操作将高层级特征图的语义信息传递到低层级特征图中, 如图 3 中红线所示. 这样可以将高层级特征图的语义信息与低层级特征图的细节信息相结合, 得到更全局和粗糙的特征表示.

然后, 自底向上路径通过对高层级特征图进行上采样和融合操作, 将低层级特征图的细节信息传递到高层级特征图中, 如图 3 中蓝线所示. 这样可以将不同尺度的特征信息进行融合, 提高检测的准确性. PANet 在自顶向下和自底向上模块增加了一个横向连接操作, 通过将自底向上和自顶向下路径的特征图进行逐元素相加或拼接, 实现特征融合, 如图 3 中黑线所示.

最后, 通过自底向上和自顶向下路径的融合, 得到了多层级的特征金字塔. 为了进一步融合这些特征, PANet 采用了一个特征融合模块. 该模块通过对特征金字塔中的特征图进行上采样和下采样操作, 将不同尺度的特征图调整到相同的大小. 然后将调整后的特

征图进行逐元素相加或拼接,实现特征的融合。

如图4(a),YOLOv5s原始模型仅将第3、4、5层特征纳入到特征融合网络。然而,由于小目标和大目标在尺度上存在显著差异,仅使用这几个特征层可能无法有效处理不同尺度目标之间的差异,一些小目标仍然难以检测,或者可能被错误地归为其他类别。

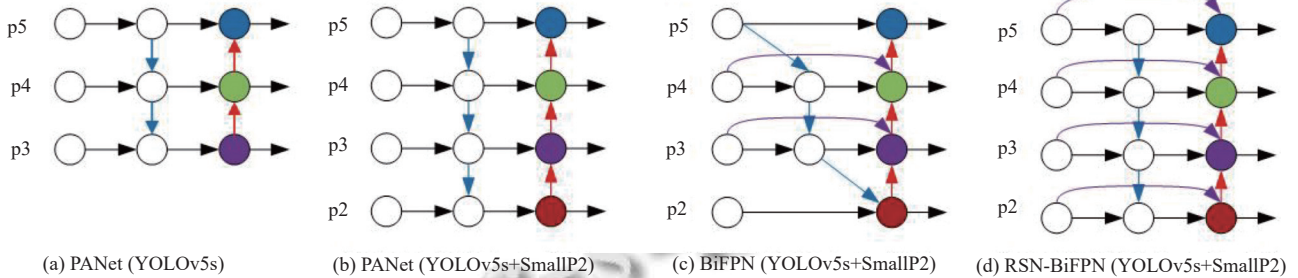


图4 特征金字塔对比

如图4(c)所示,BiFPN在PANet的基础上进行改进。首先,为了简化网络结构,BiFPN删除了只有一条输入边的节点,其原因是这类节点无法进行特征融合。其次,在第3层和第4层的输入节点和输出节点之间增加了一条额外的边,以很小的代价融合更多的特征。最后,BiFPN将每个双向(自上而下和自下而上)路径视为单个特征网络层,并重复应用同一层以实现更高级别的特征融合。BiFPN通过多尺度特征融合、上下文信息引入和优化信息流动等方式,增强对小目标的表达能力,提高小目标检测、定位的准确性。

本文在BiFPN的基础上,提出了一种名为RSN-BiFPN的结构,其结构如图4(d)所示。RSN-BiFPN首先将第2层特征加入到特征融合结构中,以融合更多不同尺度的特征。其次,保留第2层和第5层结点以减少深层语义信息的丢失。最后,在每一层的输入节点和输出节点之间增加了一条额外的边,实现跨尺度的连接。并引入了可学习的权重因子来表征不同输入特征的重要程度,同时在自上而下和自下而上的多尺度特征融合中反复应用这些权重。改进后的网络一定程度上增加少量的计算复杂度,但有效减少复杂场景下的漏检和误检现象。下面是RSN-BiFPN的两个特征融合公式:

$$P_i^{td} = Conv \left(\frac{w_1 \cdot P_i^{in} + w_2 \cdot P_{i+1}^{in}}{w_1 + w_2 + w_3 + \varepsilon} \right) \quad (1)$$

$$P_i^{out} = Conv \left(\frac{w'_1 \cdot P_i^{in} + w'_2 \cdot P_i^{td} + w'_3 \cdot P_{i-1}^{out}}{w'_1 + w'_2 + w'_3 + \varepsilon} \right) \quad (2)$$

为了解决这个问题,本文引入了第2层特征层到特征融合网络中,如图4(b)所示。由于网络添加了小目标检测层,将原本不参与特征融合的第2层特征层加入特征融合网络中,过多保留浅层语义信息导致网络的深层语义信息丢失严重,使得网络特征相对复杂。因此,更多地保留这些相对深层的语义信息显得尤为重要。

其中, P_i^{td} , P_i^{out} 分别表示自上而下路径上第*i*层的中间过渡特征以及自下而上第*i*层的最后输出特征, $Conv$ 代表对总体计算结果进行卷积操作, ε 是防止梯度消失的超参数。在式(1)中, w_1 , w_2 代表当前输入层的权重参数以及下一层输入所需要计算的中间过渡特征。在式(2)中, w'_1 , w'_2 , w'_3 分别代表当前层的输入权重,当前层中过渡单元输出的权重以及前一层的输出的权重。

2.3 损失函数优化

边界框损失函数在目标检测任务中起着关键作用,对于准确的目标定位和识别具有至关重要的影响。在目标检测任务中, IoU (intersection over union) 被广泛用于衡量预测边界框与真实边界框之间的重叠程度。然而,传统的 IoU 存在一些问题。首先, IoU 仅关注边界框之间的重叠程度,而未考虑位置和形状的差异。这使得在目标形状变化较大或存在重叠目标的情况下, IoU 无法提供准确的损失度量。

$CIoU$ (complete intersection over union) 是一种改进的 IoU 度量方法,综合考虑了边界框的位置、形状和大小之间的关系。然而, $CIoU$ 在一些特定情况下仍然存在一些局限性。例如,当目标发生尺度变化或具有较大的长宽比时, $CIoU$ 可能无法准确地衡量边界框的重叠程度,影响检测的准确性。YOLOv5s模型采用 $CIoU$ 损失函数作为边界框损失函数, L_{CIoU} 计算公式如式(3)所示。其中 IoU 表示重叠面积, α 是用做权衡的一个参数, ν 是衡量长宽比的一致性参数。

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{(c_w)^2 + (c_h)^2} + \alpha v \quad (3)$$

其中, α 和 v 计算公式如式(4)和式(5)所示:

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (4)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (5)$$

为了解决 CIoU 的局限性, 本文采用了改进的边界框损失函数 EIoU (enhanced intersection over union). 图 5 展示了 CIoU 和 EIoU 损失函数在预测框回归过程中的迭代对比. 红色框和绿色框代表预测框的迭代过程, 蓝色框表示真实框, 黑色框为预设的锚框. 通过观察图中的变化, 可以发现 CIoU 无法同时调整宽度和高度, 而 EIoU 能够有效解决这个问题.

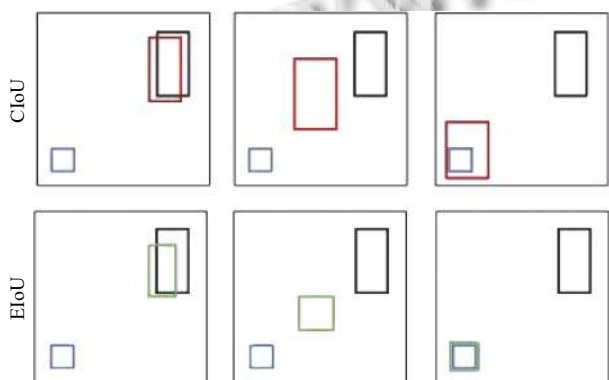


图 5 CIoU 与 EIoU 的迭代对比

EIoU 在 CIoU 的基础上进一步考虑了目标的尺度和长宽比信息, 通过引入额外的尺度项, 能够更准确地度量边界框之间的重叠程度. 与 CIoU 相比, EIoU 在处理尺度变化较大的情况下表现出更好的鲁棒性和准确性. EIoU 损失函数由重叠损失 L_{IoU} 、中心距离损失 L_{dis} 和宽高损失 L_{asp} 组成, 其定义如式(6)所示. 通过优化交通信号灯检测的收敛速度和定位精度, EIoU 损失函数有助于降低回归结果不准确的可能性.

$$\begin{aligned} L_{EIoU} &= L_{IoU} - L_{dis} - L_{asp} \\ &= 1 - IoU + \frac{\rho^2(b, b^2)}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \end{aligned} \quad (6)$$

其中, c_w 和 c_h 分别表示覆盖预测框和真实框的最小边界框的宽和高; w 和 w^{gt} 分别为预测框和真实框的宽度; h 和 h^{gt} 分别是预测框和真实框的长度; c 是可以同时包含预测边界框和真实边界框的最小闭包区

域的对角线距. IoU 为预测边界框与真值边界框的交并比.

3 实验设计与结果分析

3.1 实验环境与数据集

本文的实验环境配置见表 1, 在本文的消融实验、对比分析实验以及与其他算法的对比实验中, 使用了相同的实验配置.

表 1 实验环境配置

配置名称	版本参数
处理器	Intel® Xeon(R) Gold 5218R CPU@ 2.10 GHz
运行内存	64 GB
显卡	NVIDIA A10
深度学习框架	PyTorch 3.9
CUDA	11.7

为了评估本文所提出的改进 YOLOv5s 模型, 本文实验数据集采用由 Kaggle 数据众包平台提供的 S2TLD 小尺度交通信号灯数据集, 经过扩充后共包含 6 095 张图像, 包括多种天气条件下的交通信号灯目标. 每幅图像的大小为 1080×1920 像素, 包含 4 种不同类别的标签: 红灯 (red)、黄灯 (yellow)、绿灯 (green) 和关闭 (off), 数量分别为 4 080 个、6 175 个、375 个和 1 530 个. 数据集按 9:1 的比例被划分为训练集和测试集, 图 6 展示了数据集中的部分图像.



图 6 数据集图像示例

3.2 评价指标

为了评估改进后的算法在交通信号灯检测方面的有效性, 本实验采用以下指标进行评估: 精确度 (Precision)、召回率 (Recall)、多类别平均精度 (mAP) 和每秒帧数 (FPS).

(1) 精确度 (*Precision*)、召回率 (*Recall*) 的计算公式如下:

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

其中, *TP*指正确检测到的目标数, *FP*指错误检测到的目标数, *FN*指未检测到的目标数; *Precision*表示正确预测结果占有所有预测正样本的百分比; *Recall*表示正确预测结果在所有正样本中所占的百分比.

(2) 多类别平均精度 (*mAP*) 的计算公式如下:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (9)$$

其中, *n*指需要分类的类别数, *AP_i* (*average precision*) 表示目标类的平均精度, *mAP*表示在不同召回率下得到的精度的平均值.

(3) 目标检测网络的检测速度通常用FPS (每秒帧数) 来衡量. FPS表示网络模型每秒能够处理的图像数量, 数值越大代表网络模型处理图像的速度越快.

3.3 损失函数对比分析

在本实验的模型训练过程中, 采用了Mosaic-9数据增强技术, 并将图像输入尺寸设置为640×640. 初始学习率设置为0.01, 最终学习率设置为0.1. 模型的优化采用随机梯度下降法 (SGD), 进行150个训练周期 (epochs), 每个批处理的大小为8.

为验证优化损失函数是否能够提升模型的性能、加快网络的收敛速度, 在本实验中, 对比GIoU^[18]、DIoU^[19]、SIoU^[20]、CIoU^[19]和EIoU^[21]等损失函数, 其效果如图7所示, 其损失值及*mAP*如表2所示.

在本次实验中, 以YOLOv5s-GIoU为例, 该模型采用GIoU作为边界框损失函数, 其他模型采用类似的命名方式. 通过对图7的分析, 观察到模型在大约20个epoch后开始收敛. 与其他损失函数相比, 使用EIoU损失函数, 模型的收敛速度更快, 并且在训练结束时具有更小的损失值. 模型的损失值在第120个epoch时基本稳定. 为了确保对比实验的有效性, 本文选择第120个epoch时的损失值和*mAP*进行分析.

综合考虑图7中的损失曲线和表2中的损失值与*mAP*值, 使用EIoU损失函数相比CIoU损失函数, 边界框损失下降0.00038, *mAP*提升0.6%. 与其他损

失函数模型相比, EIoU损失函数的模型达到了最低的损失值0.02056和最高的*mAP*值95.2%, 表现出最佳性能.

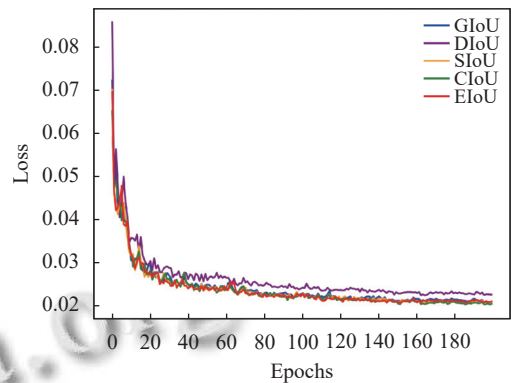


图7 损失函数曲线对比图

表2 不同模型的损失值及*mAP*

模型	损失值	<i>mAP</i> (%)
YOLOv5s-GIoU	0.02102	95.1
YOLOv5s-DIoU	0.02279	94.9
YOLOv5s-SIoU	0.02122	95.0
YOLOv5s-CIoU	0.02094	94.6
YOLOv5s-EIoU	0.02056	95.2

3.4 消融实验结果分析

为验证改进YOLOv5s交通信号灯检测算法的有效性, 设计了消融研究来验证添加检测层、使用RSN-BiFPN特征金字塔和修改损失函数对模型在检测交通信号灯方面的性能影响. 本文提出的改进后的YOLOv5s模型在S2TLD数据集上的消融实验结果见表3, 其中“√”代表使用此改进措施.

经过消融实验对比分析, 结果表明: 实验A中, 模型头部添加一个小目标检测层可以提升模型性能. 具体来说, *Precision*值提升1.6%, *Recall*值提升1.4%, *mAP*提升1.2%. 然而, 由于添加了检测层, 计算量增加, 导致FPS下降2.7f/s. 实验B中, 将backbone中的特征金字塔替换为RSN-BiFPN, 这使得模型能够融合更多尺度的特征. 相比于基准模型, 实验B中*Precision*值、*Recall*值、*mAP*值分别提升1.9%、1.4%、0.7%. 额外融合了第2层的特征导致计算量增加, 使得FPS下降3.6f/s. 实验C对损失函数进行改进后, 与未改进的模型相比, *Precision*值提升2.8%, *Recall*值提升2%, *mAP*提升0.6%, FPS下降0.9f/s, 使用不同的损失函数模型有不同的性能. 实验D显示, 同时添加小目标检测

层和修改特征金字塔,在融合更多特征的同时也提升了对小目标的检测效果的稳定性,*Precision* 值、*Recall*

值和 *mAP* 分别提升 2%、0.3% 和 1.3%。然而,由于计算量的增加,FPS 下降 3.1 f/s。

表 3 消融实验结果分析

模型	RSN-BiFPN金字塔	添加小目标检测层	改用ElIoU Loss	<i>Precision</i> (%)	<i>Recall</i> (%)	<i>mAP</i> (%)	FPS (f/s)
YOLOv5s	—	—	—	92.0	92.9	94.6	23.2
A	√	—	—	93.6	94.3	95.8	20.5
B	—	√	—	93.9	94.3	95.3	19.6
C	—	—	√	94.8	94.9	95.2	22.3
D	√	√	—	94.0	93.2	95.9	20.1
本文算法	√	√	√	96.1	95.9	96.5	22.7

同时采用 3 种改进策略,与基准模型相比,本文在不降低检测速度的前提下,成功将模型的准确率提升 4.1%,达 96.1%;召回率提升 3%,达 95.9%;*mAP* 提升 1.9%,达 96.5%。这些改进策略的实验证明了它们对提高模型性能的有效性。图 8 展示了模型改进对各类标签在 *mAP* 上的提升。

和 21.4 f/s。因此,在实时性方面,本文算法也具有一定优势。

表 4 对比实验结果分析

模型	<i>mAP</i> (%)	FPS (f/s)
YOLOv3-tiny	92.4	32.1
YOLOv5s	94.6	23.1
YOLOv7	92.3	14.4
SSD	76.0	27.9
Faster R-CNN	96.3	1.3
本文算法	96.5	22.7

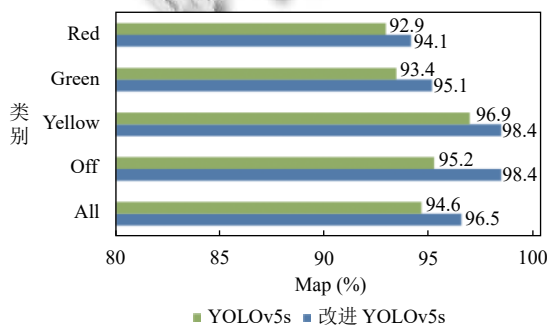


图 8 YOLOv5s 和改进 YOLOv5s 的 *mAP* 对比图

3.5 对比实验结果分析

为了验证本文提出的改进的 YOLOv5s 算法在小尺度交通信号灯检测方面的有效性,本文进行了性能对比实验。将改进的 YOLOv5s 算法与目前主流的目标检测算法,包括 YOLOv3-tiny、YOLOv5s、YOLOv7、SSD 和 Faster R-CNN 进行比较,实验结果如表 4 所示。

通过对比表 4 中的实验结果,可以发现本文提出的改进算法在检测精确度方面优于其他主流模型。与 YOLOv3-tiny、YOLOv5s、YOLOv7、SSD 和 Faster R-CNN 模型相比,改进模型的平均精确度分别提升 4.1%、1.9%、4.2%、20.5% 和 0.2%。

从模型的实时性能来分析,经过测试,本文提出改进算法的 FPS 达 22.7 f/s。虽然与 YOLOv3-tiny 和 YOLOv5s 相比略有降低,但相比于 YOLOv7 和 Faster R-CNN 算法,本文算法的检测速度分别提高 8.3 f/s

通过对对比实验结果可发现,改进后的算法有效实现交通信号灯准确、快速地检测,进一步证明了本文算法的有效性和稳定性。

3.6 检测结果分析

为了验证模型的有效性,本文对检测结果进行可视化分析,同时对检测到的小目标进行局部放大处理。图 9 展示了低亮度小目标的检测情况。在图 9(b) 中,原始 YOLOv5s 模型未能检测到灯光亮度较暗的红色交通信号灯,而在图 9(c) 中,改进后的模型通过添加小目标检测层成功检测到了红灯目标。图 10 展示了多目标的检测情况。在图 10(b) 中,原始 YOLOv5s 模型给出的目标置信度分别为 0.33、0.63、0.63,而在图 10(c) 中,改进后的模型检测出的目标置信度分别提升至 0.74、0.74、0.70,改进后的模型在检测精度上有所提升。图 11 展示了对于远处微小目标的误检情况。通过对比图 11(b) 和图 11(c),可以观察到改进后的模型由于采用了 RSN-BiFPN 特征金字塔,使得模型能够更准确地识别目标特征,避免将远距离的小目标误检为黄灯。综上所述,改进的 YOLOv5s 模型在检测小尺度交通信号灯目标方面表现优于原始 YOLOv5s 模型,同时减少了误检和漏检。改进后的模型具有更好的性能和更高的检测精确性。



图9 低亮度小目标检测

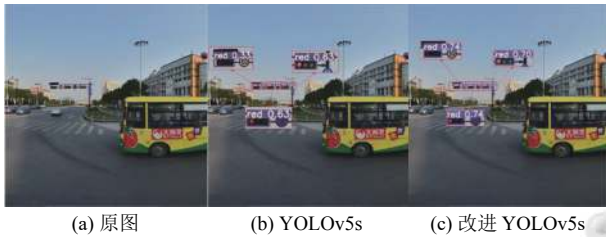


图10 多目标检测



图11 远处微小目标检测

4 总结

针对小尺度交通信号灯检测效果差、误检和漏检等问题,本文对 YOLOv5s 算法进行改进。首先,本文设计一种新的特征金字塔结构 RSN-BiFPN,减少了小尺度交通信号灯目标特征信息的丢失,从而降低误检率和漏检率。同时,通过添加新的特征融合层和预测头,使得模型能够更好地理解和区分不同尺度的交通信号灯。此外,采用 EIoU 损失函数优化输出预测边界框,避免生成过大的检测框,进一步减少误检和漏检问题。改进后 YOLOv5s 模型的 mAP 达到 96.5%,相比原始 YOLOv5s 模型提高 1.9%。此外,该模型的 FPS 达到 22.7 f/s,实验结果表明它不仅具备极好的检测精度,还能满足实时检测的需求。在未来的研究工作中,计划通过获取更多的图像数据集来扩大模型的应用范围。同时,研究采用轻量化网络作为 backbone,构建一个高性能的轻量级检测模型,以部署在小型移动设备上。

参考文献

1 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards

- real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 2 Zhang HY, Wang Y, Dayoub F, *et al.* VarifocalNet: An IoU-aware dense object detector. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 8514–8523.
- 3 Zhou XY, Koltun V, Krähenbühl P. Probabilistic two-stage detection. *arXiv:2103.07461*, 2021.
- 4 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 7263–7271.
- 5 Redmon J, Farhadi A. YOLOv3: An incremental improvement. *arXiv:1804.02767*, 2018.
- 6 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. *arXiv:2004.10934*, 2020.
- 7 Wang CY, Bochkovskiy A, Liao HYM. Scaled-YOLOv4: Scaling cross stage partial network. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 13029–13038.
- 8 Ge Z, Liu ST, Wang F, *et al.* YOLOX: Exceeding YOLO series in 2021. *arXiv:2107.08430*, 2021.
- 9 Tian Z, Shen CH, Chen H, *et al.* FCOS: Fully convolutional one-stage object detection. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 9627–9636.
- 10 Zhu XZ, Su WJ, Lu LW, *et al.* Deformable DETR: Deformable Transformers for end-to-end object detection. *Proceedings of the 9th International Conference on Learning Representations*. OpenReview.net, 2021.
- 11 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 779–788.
- 12 Wang JF, Chen Y, Dong ZK, *et al.* Improved YOLOv5 network for real-time multi-scale traffic sign detection. *Neural Computing and Applications*, 2023, 35(10): 7853–7865. [doi: [10.1007/s00521-022-08077-5](https://doi.org/10.1007/s00521-022-08077-5)]
- 13 Chen Y, Wang JF, Dong ZK, *et al.* An attention based YOLOv5 network for small traffic sign recognition. *Proceedings of the 31st IEEE International Symposium on Industrial Electronics*. Anchorage: IEEE, 2022. 1158–1164.
- 14 Wang QF, Sun X, Yi KX, *et al.* Real time traffic sign

- recognition algorithm based on SG-YOLO. Proceedings of the 21st Asia Simulation Conference. Changsha: Springer, 2022. 86–99.
- 15 Chu JQ, Zhang C, Yan MM, *et al.* TRD-YOLO: A real-time, high-performance small traffic sign detection algorithm. *Sensors*, 2023, 23(8): 3871. [doi: [10.3390/s23083871](https://doi.org/10.3390/s23083871)]
- 16 Bi ZQ, Xu FQ, Shan MJ, *et al.* YOLO-RFB: An improved traffic sign detection model. Proceedings of the 12th International Conference on Mobile Computing, Applications, and Services. Springer, 2021. 3–18.
- 17 Liu S, Qi L, Qin HF, *et al.* Path aggregation network for instance segmentation. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 8759–8768.
- 18 Rezatofghi H, Tsoi N, Gwak J, *et al.* Generalized intersection over union: A metric and a loss for bounding box regression. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 658–666.
- 19 Zheng ZH, Wang P, Liu W, *et al.* Distance-IoU loss: Faster and better learning for bounding box regression. Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2020. 12993–13000.
- 20 Gevorgyan Z. SIoU loss: More powerful learning for bounding box regression. arXiv:2205.12740, 2022.
- 21 Zhang YF, Ren WQ, Zhang Z, *et al.* Focal and efficient IOU loss for accurate bounding box regression. *Neurocomputing*, 2022, 506: 146–157. [doi: [10.1016/j.neucom.2022.07.042](https://doi.org/10.1016/j.neucom.2022.07.042)]
- 18 Rezatofghi H, Tsoi N, Gwak J, *et al.* Generalized intersection over union: A metric and a loss for bounding box regression. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 658–666.

(校对责编: 牛欣悦)