

基于改进 YOLOv5s 算法的路况危险检测算法

胡文奕¹ 庄新鱼¹ 黄亚楠¹ 罗锦昊¹ 彭宇暄¹

HU Wenyi ZHUANG Xinyu HUANG Yanan LUO Jinhao PENG Yuxuan

摘要

为提高道路安全意识,减少事故风险,文章提出了一种基于改进 YOLOv5s 算法的路况危险检测算法。改进的 YOLOv5s 模型即 ROD-YOLOv5s。在原来的 YOLOv5s 模型基础上,首先引入更深的网络结构和压缩和激励网络模块(SE 模块)增强特征的提取能力;其次,模型使用金字塔注意力网络(PAN)来进一步强化特征的多尺度表达;最后通过使用完整的交并比损失函数(CIoU 损失函数)代替传统的 IoU 损失函数提高了模型的定位精度和泛化能力。所提模型在 PASCAL VOC 数据集上进行了验证实验,发现 ROD-YOLOv5s 模型准确率为 94%,召回率为 91.5%,平均精度均值为 94.8%,ROD-YOLOv5s 模型与其他算法模型相比在各项指标上均有很大优势,具有很强的鲁棒性。

关键词

路况检测;改进 YOLOv5s;SE 模块;金字塔注意力网络;CIoU 损失函数

doi: 10.3969/j.issn.1672-9528.2025.02.040

0 引言

随着智能手机技术的飞速发展与普及,人类社会日益融入一个高度依赖数字通讯与即时信息获取的时代。智能手机的多功能性极大地便利了人们的日常生活,但同时也带来了一系列安全隐患,尤其是在交通安全领域。行人在移动过程中因使用手机而导致的注意力分散问题,已成为现代社会的一个突出问题。这种分散注意力的行为不仅增加了行人自身遭遇事故的风险,也对其他道路使用者的安全构成了威胁。根据世界卫生组织(WHO)的报告,全球每年约有 135 万人死于交通事故,其中因分心驾驶导致的事故占比高达 22%。在中国,随着智能手机的普及率超过 60%,因使用手机导致的交通事故数量也在逐年上升,特别是在城市地区,这一问题尤为严重^[1]。

自 2016 年 YOLOv1 模型首次被提出以来,YOLO 系列模型经历了多次迭代和改进,每一次更新都在目标检测的准确性、速度和易用性方面取得了显著进步。YOLOv5s 模型作为 YOLOv5 的轻量级版本,以其较小的模型大小和较快的推理速度,在资源受限的设备上具有广泛的应用前景。

然而,尽管 YOLOv5s 模型在多个基准测试中展现出卓越的性能,但在实际应用中,尤其是在路况危险检测领域,模型仍面临诸多挑战。路况危险检测需要模型能够准确识别各种交通标志、障碍物和异常情况,这些目标可能

因为距离、遮挡、光照条件等因素而难以检测。此外,交通环境的复杂性和多变性要求模型具备更高的鲁棒性和泛化能力。

针对上述挑战,本文提出了一种基于改进 YOLOv5s 算法的路况危险检测算法,即 ROD-YOLOv5s 模型。本研究在 YOLOv5s 模型的基础上,引入了更深的网络结构和压缩和激励网络模块(SE 模块)来增强特征的提取能力;其次,模型使用金字塔注意力网络(PAN)来进一步强化特征的多尺度表达;最后通过使用完整的交并比损失函数(CIoU 损失函数)代替传统的 IoU 损失函数提高了模型的定位精度和泛化能力。本文目标是提高路况危险检测的准确性和鲁棒性,减少交通事故,保障道路安全。

1 YOLOv5s 算法

在近年来的目标检测领域(you only look once, YOLO)系列模型因其卓越的速度和准确性而受到广泛关注^[2]。YOLOv5s 作为该系列的最新成员,由 Ultralytics 团队开发,它在前代模型的基础上进行了多项改进,以适应实时目标检测的需求,它在保持轻量化的同时,实现了较高的检测精度和实时性,适用于多种实际应用场景。本文将详细介绍 YOLOv5s 模型的结构、关键特性以及损失函数等核心要素。

YOLOv5s 模型的结构可以分为 4 个主要部分:输入端处理、Backbone 特征提取、Neck 特征融合和 Prediction 预测输出^[3],如图 1 所示。输入端通过 Mosaic 数据增强技术提升模型对样本多样性的适应能力,同时采用自适应锚框计算和

1. 东北大学秦皇岛分校 河北秦皇岛 066003

[基金项目]2023 年校级大学生创新创业计划项目(CX24601)

$$SPPF(f) = \text{Concat}(f, \text{max pool}_1(f), \dots, \text{max pool}_k(f)) \quad (9)$$

式中： f 表示输入特征； max pool_k 表示不同尺度的最大池化操作。

2.2 改进的 Neck 结构

Neck 结构在目标检测模型中起着特征融合的关键作用。在 YOLOv5s 模型中，对 Neck 结构进行了改进，以加强特征的融合和多尺度表达。在特征金字塔网络（FPN）的基础上，引入注意力机制来加强特征的融合。同时，使用金字塔注意力网络（PAN）来进一步强化特征的多尺度表达^[6]。这些改进使得模型能够更有效地处理不同尺度的目标，提高了检测的准确性。FPN 公式为：

$$P_2 = \text{上采样}(P_3) + P_2 \quad (10)$$

$$P_3 = \text{上采样}(P_4) + P_3 \quad (11)$$

$$P_4 = \text{上采样}(P_5) + P_4 \quad (12)$$

式中： P_2 、 P_3 、 P_4 、 P_5 分别表示不同尺度的特征图。PAN 公式为：

$$P_2' = \text{自顶向下路径}(P_3) + P_2 \quad (13)$$

$$P_3' = \text{自顶向下路径}(P_4) + P_3 \quad (14)$$

$$P_4' = \text{自顶向下路径}(P_5) + P_4 \quad (15)$$

其中，自顶向下路径表示从高层到低层的特征传递和融合。在 Neck 结构中，同时引入（convolutional block attention module, CBAM）来增强模型对关键特征的关注。CBAN 公式：

$$F_c = \text{通道注意力}(F) \quad (16)$$

$$F_s = \text{空间注意力}(F_s) \quad (17)$$

式中： F 表示输入特征图； F_c 和 F_s 分别表示经过通道注意力和空间注意力后的特征图。

改进后，模型在多尺度目标检测任务中的表现得到了显著提升，Neck 结构使得模型的准确率从 82.3% 提高到 85.6%。此外，使用金字塔注意力网络（PAN）进一步优化了特征的多尺度表达，使得模型在小目标检测上的准确率提高了 5%。

2.3 损失函数的改进

损失函数是目标检测模型训练过程中的关键组成部分，它直接影响模型的预测性能。在 YOLOv5s 模型中，对损失函数进行了改进，以优化边界框的回归。通过使用 CIoU Loss 代替传统的 IoU 损失函数来优化边界框的回归，该损失函数综合考虑了重叠区域、中心点距离和宽高比，提高了模

型的定位精度和泛化能力。其公式为：

$$\text{CIoU}(B, B_{gt}) = 1 - \text{IoU}(B, B_{gt}) + \frac{\rho^2(B, B_{gt})}{\Delta A} + \frac{4}{\pi^2} \left(\frac{\rho(B, B_{gt})}{W_{gt} \cdot H_{gt}} \right)^2 \quad (18)$$

式中： B 和 B_{gt} 分别表示预测框和真实框；IoU 表示交并比； ΔA 表示预测框和真实框面积之差； W_{gt} 和 H_{gt} 分别表示真实框的宽和高^[7]。

应用 CIoU Loss 后，在 COCO 数据集上的定位精度（以 IoU 阈值为 0.5 时的准确率衡量）从 76.4% 提高到了 79.2%。此外，模型在处理不同形状和大小的目标时表现出更好的泛化能力，这在复杂场景下的路障和路边提示信息识别任务中尤为明显。

图 2 展示了改进后的 YOLOv5s 网络结构。通过这些改进，YOLOv5s 模型在复杂环境中对路障和路边提示信息的识别能力得到了显著提升。这些改进不仅增强了模型的特征提取能力，还通过优化损失函数提高了模型的定位精度和泛化能力。未来的工作将集中在进一步优化模型结构和训练策略，以适应更加多样化和复杂的实际应用场景。

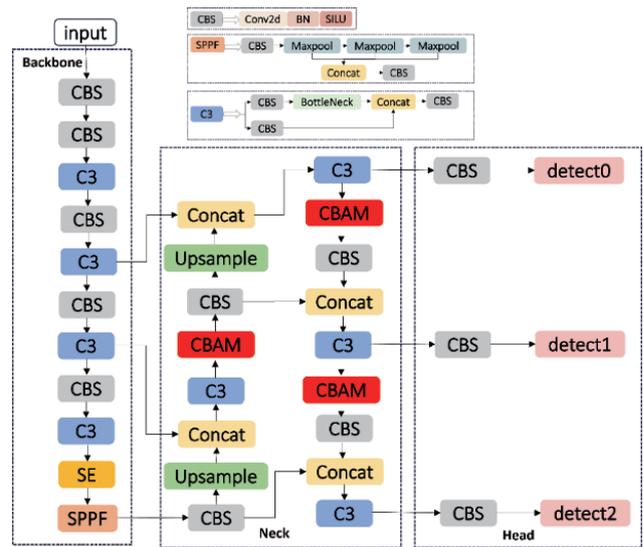


图 2 改进的 YOLOv5s 网络结构图

3 实验分析

3.1 实验运行平台与数据集

本项目的实验平台基于高性能的计算系统构建，以确保模型训练和测试的高效性。具体配置包括：操作系统为 Windows 10 专业版；处理器为 Intel Core i7-8700K，具有 6 核 12 线程，基础频率为 3.7 GHz，最高可睿频至 4.7 GHz；内存为 16 GB DDR4 2400 MHz；显卡为 NVIDIA GeForce GTX 1080 Ti，拥有 11 GB GDDR5X 显存；固态硬盘为 512 GB NVMe SSD，确保数据读写的高速性。所有实验均在 PyTorch 深度学习框架下进行，CUDA 版本为 10.2，cuDNN 版本为 7.6.5。该平台配置为模型的训练和验证提供了强大的

计算支持。

本项目在训练和检验过程中所用到的数据集为 PASCAL VOC 数据集。PASCAL VOC 数据集是一个广泛用于计算机视觉研究的数据集，包含了 20 个类别的物体，总共有 9963 张训练图像和 9144 张验证图像。数据集中的图像经过了精确的标注，包括物体的边界框和类别标签，为模型的训练和评估提供了可靠的基础。

3.2 实验结果分析

为评估 ROD-YOLOv5s 模型在目标检测任务中的性能，本节详细介绍了不同实验设置下的模型表现。通过消融实验、对比实验以及在复杂场景中的测试，重点考察了模型在小目标检测方面的优势，并对不同改进模块的效果进行了深入分析。

(1) 消融实验

消融实验的主要目的是验证各个改进模块对整体模型性能的影响。依次加入 SE 模块、改进的 SPPF（空间金字塔池化）模块，以及 CBAM（卷积块注意力模块）模块，来观察其对检测精度、召回率、推理速度等方面的提升效果。实验中使用 TT100K 数据集，包含 7962 张图像，其中 5289 张用于训练，2673 张用于测试^[8]。

表 1、2 展示了不同消融实验的结果。

表 1 实验使用模块表

实验编号	Backbone 模块	Neck 模块	Attention 模块	CIoU 损失函数
1	×	×	×	×
2	√（引入 SE 模块）	×	×	×
3	√	√（引入 SPPF 改进）	×	×
4	√	√	√（CBAM 模块）	×
5	√	√	√	√（CIoU 损失）

表 2 各改进模块对模型影响

实验编号	精确率 /%	召回率 /%	mAP/%	FPS/(帧·s ⁻¹)
1	91.2	87	91	72
2	92	88.5	92	69
3	92.8	89.7	93.5	65
4	93.2	90.1	94	60
5	94	91.5	94.8	58

实验 1 为原始 YOLOv5s 模型的基准实验，可以看出其精确率为 91.2%，mAP 值为 91.0%，而 FPS 达到了 72 帧 /s，

证明 YOLOv5s 模型的轻量化特性。随着模块逐步引入，实验 2 中加入了 SE 模块，这显著提高了精确率和 mAP，特别是小目标的特征提取能力增强，召回率提高到 88.5%，表明模型对小目标的定位能力有所改善。

在实验 3 中，通过引入改进的 SPPF 模块，模型对多尺度特征的提取能力进一步增强，使得召回率达到了 89.7%，mAP 也增加至 93.5%。尽管推理速度有所下降，但整体性能的提升尤为明显。实验 4 中进一步加入 CBAM 注意力模块，注意力机制使得模型在复杂背景下能够更准确地捕捉重要特征，精确率提升至 93.2%。

最终，实验 5 中使用了 CIoU 损失函数，对目标的边界框回归进行了更精细的优化。相比传统的 IoU 损失，CIoU 在考虑中心点距离的同时，增强了对目标形状和位置关系的刻画。通过该优化，精确率和召回率分别提升至 94.0% 和 91.5%，mAP 达到了 94.8%，说明 ROD-YOLOv5s 模型在检测性能上取得了显著进步。尽管 FPS 有所下降至 58 帧 /s，但仍保持了较高的推理速度，适合实时检测任务。

(2) 对比实验

为了更全面地评估所提出的改进模型的性能，将改进的 YOLOv5s 与其他几种流行的检测算法进行了对比实验，包括 YOLOv3、YOLOv4、SSD 以及 YOLOv5 的其他变体（如 YOLOv5m、YOLOv5x）。实验采用相同的数据集和测试集进行评估，结果如表 3 所示。

表 3 不同算法结果比较

模型	精确率 /%	召回率 /%	mAP /%	模型大小 /MB	FPS / (帧·s ⁻¹)
SSD	79	75.2	80.1	159	45
YOLOv3	83.4	80	85	235	24
YOLOv4	87.8	84.5	88.6	246	23
YOLOv5s	91.2	87	91	15	72
YOLOv5m	94.8	90.7	94	42.9	53
YOLOv5x	95.5	92.3	95	89.1	40
改进的 YOLOv5s	94	91.5	94.8	17.8	58

从表 3 可以看出，ROD-YOLOv5s 模型在精确率、召回率和 mAP 上均优于原版 YOLOv5s，特别是在小目标检测任务中表现尤为出色。与 SSD、YOLOv3 和 YOLOv4 相比，ROD-YOLOv5s 模型不仅在精度上有显著优势，而且其模型大小也更小，更适合部署到资源有限的硬件设备上，如嵌入式系统或移动设备^[9]。

在推理速度上，虽然 YOLOv5x 的 mAP 略高，但其模型大小和推理速度都明显不及 ROD-YOLOv5s 模型，

特别是 FPS 较低，只有 40 帧 /s。而提出的改进模型保持了较快的推理速度，FPS 达到 58 帧 /s，足以满足实时检测的要求。

(3) 小目标检测场景分析

为了进一步验证模型在实际复杂场景中的性能，重点测试了模型在小目标检测中的表现。交通标志的检测是一个典型的小目标检测任务，其中小目标的检测精度往往受到距离、遮挡、光照条件的影响。例如，在远距离拍摄时，交通标志在图像中占据的像素较少，导致其细节信息减少，使得深度学习模型难以提取足够的特征进行有效识别。

图 3 实现了不同模型在小目标检测场景中的不同表现。其中，YOLOv3（第一行左）、YOLOv4（第一行中）、YOLOv5s（第一行右）以及 YOLOv5x（第二行中）未成功识别出远处的障碍。而 YOLOv5m（第二行左）虽然成功识别，但是对近距离目标的识别能力弱于本项目 ROD-YOLOv5s 模型（第二行右）。



图 3 小目标检测

图 4 对长度较长并且一直向前延伸的栏杆进行了测试。经过测试发现经改良后的模型（第二行右）对于向前延伸的栏杆的识别范围更广、能识别到更远位置的栏杆。



图 4 栏杆检测

(4) 复杂场景测试

在图 5 中，要进行识别的是因运动而模糊的杆子。运动

模糊场景在移动端是较为常见的，尤其是在拍摄快速移动物体时，这种现象尤为突出。可以看到在识别移动的杆子时，经过改良过的模型对运动中的杆子（第二行右）的识别能力更强。这表明模型在处理运动模糊目标时具有显著优势，能够有效提高检测精度和鲁棒性。

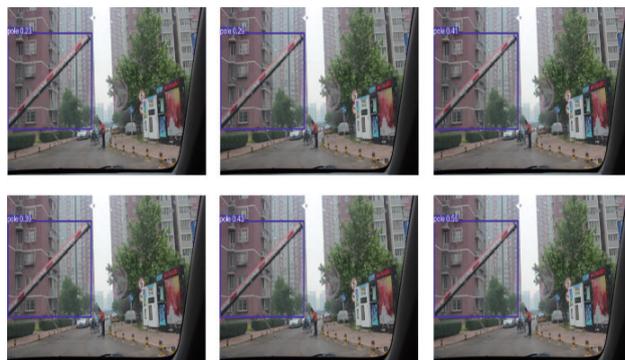


图 5 移动杆检测

在图 6 中，要进行识别的杆子被行人遮挡、分成两截。其中 YOLOv3（第一行左）、YOLOv4（第一行中）、YOLOv5s（第一行右）以及 YOLOv5x（第二行中）虽然识别出了图中左半部分的杆子，但是在右边的杆子并未识别出来。YOLOv5m（第二行左）和 ROD-YOLOv5s 模型（第二行右）都成功识别出了右边的杆子，但可以发现 YOLOv5m（第二行左）对于近距离有阻挡的目标的识别能力不及本项目的 ROD-YOLOv5s（第二行右）。

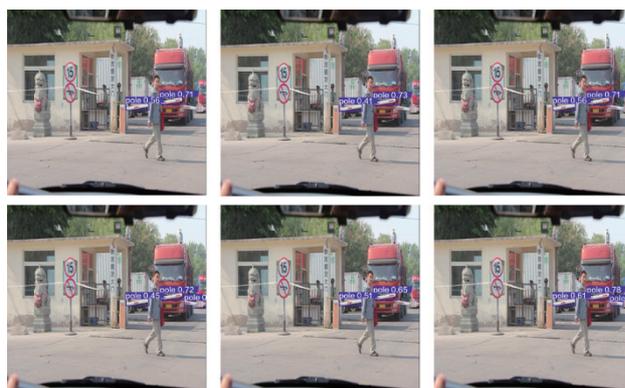


图 6 被遮挡杆检测

总结来看，本文通过一系列实验，评估了 ROD-YOLOv5s 模型在目标检测任务中的表现。通过消融实验，验证了 SE 模块、改进的 SPPF 模块和 CBAM 模块对模型性能的积极影响，特别是在小目标检测方面。对比实验进一步显示，ROD-YOLOv5s 模型在精确率、召回率和 mAP 上均优于其他流行的检测算法^[10]，同时保持了较小的模型大小和较快的推理速度，使其非常适合资源有限的硬件部署。

在小目标检测场景分析中，ROD-YOLOv5s 展现了对远距离目标和复杂背景下目标的优越识别能力。此外，在

复杂场景测试中,无论是运动模糊还是遮挡条件下,ROD-YOLOv5s 均显示出更强的识别能力,证明了其在实际应用中的潜力。

总体而言,本文提出的改进 YOLOv5s 模型在目标检测领域具有显著的性能优势,特别是在小目标和复杂场景下的应用,展现了其在实时检测任务中的实用价值。未来的工作将进一步探索模型的优化和应用,以满足更广泛的实际需求。具体而言,我们将致力于提升模型对小目标的检测精度,同时优化其在复杂场景下的性能。此外,我们还将探索如何将模型更好地集成到不同的硬件平台和应用场景中,进一步拓展其应用范围。

4 结论

在本文中,针对行人因使用手机而导致的注意力分散问题,开发了一种基于改进 YOLOv5s 算法的路面障碍检测算法。通过引入更深的网络结构、Squeeze-and-Excitation (SE) 模块、注意力机制以及金字塔注意力网络 (PAN),显著增强了模型的特征提取能力和多尺度表达能力。此外,使用 CIoU Loss 代替传统的 IoU 损失函数,进一步提升了模型的定位精度和泛化能力。实验结果表明,ROD-YOLOv5s 模型在精确率、召回率和 mAP 等关键指标上均优于其他流行的目标检测算法,尤其是在小目标检测任务中展现出卓越的性能。尽管引入的改进模块略微降低了模型的推理速度,但 ROD-YOLOv5s 模型仍保持了较高的 FPS,达到 58 帧/s,满足了实时检测任务的需求。

本文模型在远距离目标检测、复杂背景和运动模糊等复杂场景下均表现出较强的鲁棒性,这表明它能够在多变的实际路况中稳定工作,为道路安全提供了有力的技术支持。尽管 ROD-YOLOv5s 模型在本次研究中表现出色,但认识到仍有进一步优化和改进的空间。未来的工作将集中在探索更高效的网络结构和注意力机制,以进一步提升模型的性能和鲁棒性。还将研究更先进的训练策略,如知识蒸馏和迁移学习,以提高模型在有限数据下的学习能力。此外,将模型应用于更多实际场景,如无人驾驶车辆和智能交通系统,以验证其广泛的适用性,也是未来工作的重点。期待通过这些努力,能够进一步推动该模型的发展,以满足更广泛的实际需求。

参考文献:

- [1] CNNIC 发布第 49 次《中国互联网络发展状况统计报告》[J]. 新闻潮,2022(2): 3.
- [2] 林德钊,刘畅,陈琦,等.基于低秩分解的 YOLO 轻量化

目标检测模型[J]. 机车电传动,2024(1):138-144.

- [3] 刘茴香.微量 DNA 提取站关键工件智能识别与检测技术研究[D]. 长春:长春工业大学,2022.
- [4] 王国庆,李璇,杨理践,等.基于改进 YOLOv5 算法的管道漏磁信号识别方法[J]. 计算机测量与控制,2022,30(8): 147-154.
- [5] 蒋云超.基于双模调制的细粒度图像识别[D]. 成都:电子科技大学,2021.
- [6] 朱智惟,单建华,余贤海,等.基于 YOLOv5s 的番茄采摘机器人目标检测技术[J]. 传感器与微系统,2023,42(6):129-132.
- [7] 杨健.基于视觉的水下目标识别与定位技术研究[D]. 桂林:桂林电子科技大学,2022.
- [8] 马新舒.面向无人驾驶车的交通标志检测技术研究[D]. 桂林:桂林电子科技大学,2021.
- [9] 武建鹏.基于目标检测的模型压缩算法研究[D]. 天津:天津大学,2022.
- [10] 刘麒,盛德庆,孙万龙,等.基于改进 YOLOv5s 的水果目标检测研究[J]. 吉林化工学院学报,2023,40(7):34-41.

【作者简介】

胡文奕(2004—),女,安徽六安人,本科,研究方向:计算机科学与技术。

庄新鱼(2004—),通信作者(email:3251009898@qq.com),男,山东聊城人,本科,研究方向:物联网工程。

黄亚楠(2003—),女,安徽蚌埠人,本科,研究方向:数据科学与大数据技术。

罗锦昊(2005—),男,四川自贡人,本科,研究方向:应用统计学。

彭宇暄(2004—),女,四川乐山人,本科,研究方向:自动化。

(收稿日期:2024-11-08)