

改进 YOLOv8 的非机动车违规行为检测方法^{*}

李强¹ 南新元² 蔡鑫² 杨仕伟¹

(1.新疆大学电气工程学院 乌鲁木齐 830017; 2.新疆大学智能科学与技术学院 乌鲁木齐 830017)

摘要: 针对当前检测算法在非机动车不规范驾驶行为检测容易出现漏检误检的问题,提出一种改进 YOLOv8 的非机动车违规行为检测方法 YOLO-CSSM。首先在 Backbone 和 Neck 中构建了 SPD-Conv 网络模块,以提高对微小目标的学习能力,加强模型在复杂背景条件下的特征提取能力;其次分别在 Backbone 和 Neck 网络中引入 DCNv2 和 SegNext Attention 注意力机制模块,重新设计了 C2f-DCNv2,突出非机动车和驾驶人重要特征信息,提高模型特征融合能力;最后使用 WIoU 损失函数的思想改进 MPDIoU,将原 CIoU 替换为 Wise-MPDIoU,用来解决正负样本不平衡带来的问题。该算法在自建非机动车不规范驾驶行为数据集上进行验证,实验结果显示,改进后的 YOLOv8 算法在自建非机动车不规范行为驾驶数据集上的精确率 P、召回率 R 和平均精度均值 mAP@0.5 为 89.4%、90.0% 和 93.6%,比传统的 YOLOv8 算法分别提升了 3.3%、5.4% 和 4.5%,取得了更好的检测精度和效果。并以非机动车违规行为检测算法为基础,使用 PyQt5 设计开发了非机动车违规行为识别检测系统。

关键词: YOLOv8;DCNv2;Wise-MPDIoU;非机动车不规范行为;注意力机制;特征融合

中图分类号: TN407 **文献标识码:** A **国家标准学科分类代码:** 510.40

Improved YOLOv8-based method for non-motor vehicle violation detection

Li Qiang¹ Nan Xinyuan² Cai Xin² Yang Shiwei¹

(1. School of Electrical Engineering, Xinjiang University, Urumqi 830017, China;

2. School of Intelligent Science and Technology, Xinjiang University, Urumqi 830017, China)

Abstract: Addressing the issue of false negatives and positives in non-motor vehicle irregular driving behavior detection with the current detection algorithm, an improved target detection algorithm, YOLO-CSSM, was proposed based on YOLOv8. The Backbone and Neck were enhanced with an SPD-Conv network module, which improved the model's ability to learn from small targets and extract features under complex backgrounds. Subsequently, DCNv2 and SegNext Attention modules were integrated into the Backbone and Neck networks, respectively, to emphasize important feature information of non-motor vehicles and drivers, enhancing the model's feature fusion capability. The MPDIoU was improved using the concept of the WIoU loss function, replacing the original CIoU loss function with Wise-MPDIoU to mitigate the imbalance between positive and negative samples. Validated on a self-built dataset of non-motor vehicle irregular driving behaviors, the improved YOLOv8 algorithm demonstrated precision, recall and mean average precision (mAP@0.5) of 89.4%, 90.0% and 93.6%, respectively, showing improvements of 3.3%, 5.4% and 4.5% over the traditional YOLOv8 algorithm, achieving better detection accuracy and effectiveness. And Based on the non-motorized vehicle violation detection algorithm, a non-motorized vehicle violation recognition and detection system was designed and developed using PyQt5.

Keywords: YOLOv8;DCNv2;Wise-MPDIoU;non-motor vehicle irregular behavior;attention mechanism;feature fusion

0 引言

非机动车作为绿色交通的重要载体,具有灵活便捷、低

碳高效等优势^[1]。但骑行中普遍存在的未佩戴头盔、违规载人、分心驾驶等行为,显著提升了交通事故风险。研究表明,规范佩戴头盔可降低 63% 的头部受伤率和 88% 的颅脑

收稿日期:2025-03-28

* 基金项目:中央引导地方科技发展资金(ZYYD2025CG06)、国家自然科学基金(62303394)、新疆维吾尔自治区自然科学基金(2022D01C693)、新疆维吾尔自治区高校基本科研业务费科研项目(XJEDU2023P025)资助

损伤率,其缓冲机制能有效吸收 83% 的冲击动能^[2]。这些违规行为不仅违背交通法规,更直接威胁公共安全。因此,深化非机动车安全研究兼具政策优化和工程实践双重价值:既为完善交通管理体系提供数据支撑,又能通过安全技术推广与行为干预,系统提升城市交通韧性。

国内外研究非机动车不规范驾驶行为主要集中在深度学习领域,其中基于卷积神经网络(convolutional neural networks, CNN)^[3]的检测算法占据主流。该类算法主要分为双阶段检测算法和单阶段检测算法。双阶段检测算法以基于区域的卷积神经网络(region-based convolutional neural network, R-CNN)系列为代表,通过选择性搜索生成候选区域进行目标检测,能够有效处理目标尺度变化和遮挡问题,但因需对每个候选区域单独提取特征和分类,计算效率低、检测速度慢。单阶段检测算法以单次多框检测器(single shot multibox detector, SSD)^[4]和 YOLO(you only look once)系列为代表,SSD 在多种尺度特征图上检测,对小目标检测能力较强,速度优于 R-CNN,但深层特征图检测小目标时性能有限;YOLO 系列采用端到端训练,实时检测效率高,但在检测精度上稍逊于双阶段检测方法。

近年来有很多学者都对非机动车不规范驾驶行为展开了研究,晏舒童^[5]以 YOLOv4 为基础,通过轻量化主干网络,使用深度可分离卷积,引入注意力机制,并通过遗传算法优化锚框,提高了模型的检测速度,但在面对小目标和复杂场景时会牺牲检测精度。马斌^[6]以 YOLOv5s 算法为基准网络,引入 CBAM 注意力机制和空间到深度模块来减少细粒度信息丢失,但会导致模型的复杂度和检测时间增加。张国鹏等^[7]通过在 YOLOv8 使用 BiFPN 优化特征融合模块,增强了算法的特征提取与特征表达能力,使模型能够更好地处理多尺度特征,提高了对小目标的检测性能,但在实际应用中仍可能存在一定程度的误检和漏检情况。杨璐钱等^[8]提出了 Yolov8-scG 算法,通过引入 C2f-CGblock 模块和替换普通卷积与下采样算子,有效改善了模型对低分辨率图像和小目标的检测精度,提升了对图像细节的捕捉能力。但模型泛化性欠佳,对于道路强光环境等特定场景会出现一定的检测误差。周顺勇等^[9]提出了 Helmet-YOLO 算法,通过引入 PCAHead 检测头、渐进式特征金字塔网络结构和 HelmetIoU 边界框损失函数,有效地提升了模型在复杂场景下的检测能力。但增加了模型的复杂性。这可能导致在资源受限的设备上部署时面临挑战。

截止到目前,非机动车不规范驾驶行测还存在着诸多问题,比如对小目标检测效果差,不同违规行为形态有差异,易造成漏检、误检。并且多数研究只针对未戴头盔这一项不规范行为进行检测,未对其他不规范行为进行检测。本文对上述问题综合考虑,提出一种改进 YOLOv8 的非机动车违规行为检测方法,主要工作如下:1)在 C2f 网络中设计 C2f-DCNv2 模块改善对小目标误检,错检能力;2)在主

干构建 SPD-Conv 加强模型在复杂背景条件下的特征提取能力;3)针对当前研究未对其他不规范行为进行检测,本文收集了常见的 3 种不规范行为图像,包括未佩戴头盔、违规载人、骑行中使用电子设备。对多种违规行为检测,使得对非机动车不规范行为的检测更为完善;4)以非机动车违规行为检测算法为基础,使用 PyQt5 设计开发了非机动车违规行为识别检测系统。

1 YOLO-CSSM

YOLOv8 虽然在目标检测中有良好的表现,然而在处理非机动车不规范行为检测场景中尺度多变且背景复杂的目标时,其性能会受到一定影响。非机动车驾驶中不规范行为较多,包括未佩戴头盔、违规载人、骑行中使用手机等,它们特点各异,同时受到复杂环境的影响,易引发漏检和误检的问题。为了提高检测精度,本文提出一种非机动车不规范行为检测算法:YOLO-CSSM。改进如下:

1)通过在 C2f 网络中融合 DCNv2,该设计使算法对微小物体的特征提取能力得到明显增强,有效提高了复杂场景下的细粒度目标检测精度;

2)将 SegNext Attention 注意力机制整合至 Neck,其特有的特征优化机制能够有效聚焦关键区域,显著提升模型对复杂特征的辨识与表征效能;

3)在 Backbone 和 Neck 中构建 SPD-Conv 网络模块,提高对微小目标的学习能力,加强模型在复杂背景条件下的特征提取能力;

4)使用 WIoU 损失函数的思想改进 MPDIoU,将原 CIoU 替换为 Wise-MPDIoU,用来解决正负样本不均衡所带来的问题。其结构图如图 1 所示。

1.1 C2f-DCNv2 模块

驾驶非机动车的不规范行为可能有多种表现形式,如未佩戴头盔、使用手机、违规载人等,且每种行为的表现方式和环境背景各异。通常卷积操作在处理输入特征时,会将数据划分成若干固定尺寸的网格单元,所有空间位置均共享参数一致的滤波核进行特征映射,并在每个位置上执行卷积运算。每个局部区域在特征图中的位置是固定的,从而实现对整个输入特征图的局部感知与特征提取。然而,这种操作方式对非机动车不规范行为的检测结果并不好。因此,本文选择可变形卷积 DCNv2^[10]来替换原算法 C2f 模块中的 Bottleneck 结构,以更好地适应复杂场景下的行为特征提取。通常在二维卷积操作中,输入特征图中的特定坐标点 P_0 处的特征响应值 $y(P_0)$ 的计算方法为:对输入特征图在位置 P_0 处的邻域区域进行加权求和,其中权重由卷积核决定。具体而言,将输入特征图中对应位置的像素值和卷积核当中的各个元素进行逐元素相乘,并对所得的乘积进行求和运算,最终得到输出特征图中的特征值。

$$y(P_0) = \sum_{P_n \in R} \omega(P_n) x(P_0 + P_n) \quad (1)$$

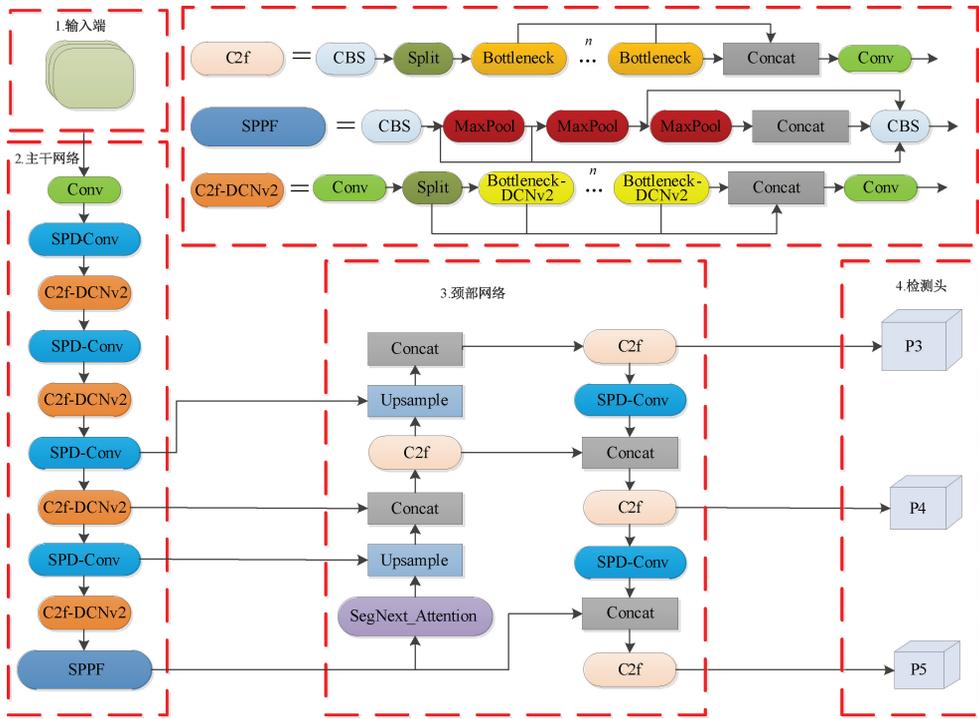


图 1 YOLO-CSSM 算法网络

Fig. 1 YOLO-CSMM network

在可变形卷积中,式(1)中的 $\omega(P_n)$ 表示在空间坐标点 P 对应的滤波核参数与特征张量值 $x(P_0 + P_n)$ 存在映射关系,其中 P_n 遍历感受野覆盖的全部坐标集合, R 则表征滤波核的几何作用域。

$$y(P_0) = \sum_{P_n \in R} \omega(P_n) x(P_0 + P_n + \Delta P_n) \quad (2)$$

式中: $x(P_0 + P_n + \Delta P_n)$ 代表动态调整后的坐标点特征响应,其中 ΔP_n 为自适应的空间位移参数。

改进前后对比如图 2 所示。对比分析表明,相较于常规卷积,可变形卷积创新性地建立了偏移量优化机制,该机制通过参数化偏移量优化采样网格的空间分布,显著提

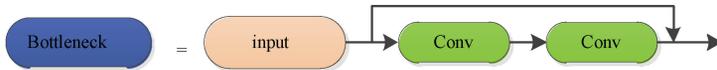
升了卷积核对目标几何特征的适配性。

基于跨阶段连接拓扑的 C2f 结构通过通道维度压缩策略优化计算流程,在浅层特征提取阶段,由于通道间信息差异性显著,需保留足够通道容量来维持细粒度特征多样性。此时网络层具有高分辨率特征丰富但深层语义表征薄弱的特点。值得注意的是,该模块的瓶颈结构创新性地集成可变形卷积核,其动态适应目标几何形变及环境干扰的特性显著优于传统固定卷积。通过残差拓扑实现输入输出特征融合,这种设计既能扩展特征感知范围,又能增强特征表达能力,最终优化模型抗干扰能力与跨场景适应。



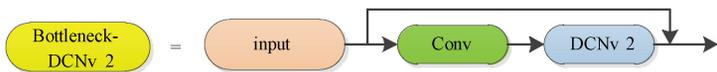
(a) C2f-DCNv2 模块结构

(a) C2f-DCNv2 modul structer



(b) 改进前Bottleneck

(b) Bottleneck



(c) 改进后Bottleneck

(c) Bottleneck-DCNv2

图 2 改进前后对比

Fig. 2 Comparison before and after improvement

1.2 SPD-Conv 模块

在检测微小目标时,如远处的头盔和手机,在非机动车数据集中微小目标所占的像素比例过小,且有一些图片的分辨率也存在此问题。此外,神经网络初始阶段的下采样处理会引发高频细节特征的衰减,这种特征完整性的缺失直接影响模型对微小物体的辨识灵敏度。为了解决这一问题,本文提出引入 SPD-Conv^[11] 模块以替代原始卷积网络中的跨步卷积和池化层,SPD-Conv 架构整合了空间深度变换单元与连续卷积单元,前者通过创新性空间压缩策略实现特征图降维,同时在通道维度构建信息保全机制,有效抑制高频特征损失。这一设计有助于在降低空间分辨率的同时,尽可能保持输入特征的通道层次特征,优化了信息传递和特征表达。该模块在每个空间深度变换单元后集成连续卷积单元,这种设计通过参数自适应机制动态调整下游卷积层的特征维度配置,从而在有效控制计算量的同时,提升模型的表现和计算效率,从而有效控制额外的计算负担。通过对大小为 $S \times S \times C_1$ 的子特征映射

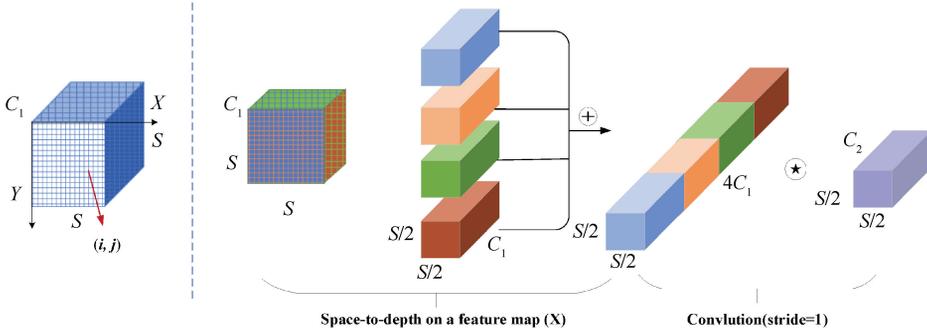


图 3 当比例因子为 2 时 SPD-Conv 的示意图

Fig. 3 Schematic diagram of SPD-Conv at scale=2

SPD 在处理流程结束后,级联配置 C_2 ($stride = 1$) 预设核参数的连续卷积单元,旨在维持关键判别特征的完整性。随后, $C_2 < scale^2 C_1$ 进一步对这些特征进行转换处理。

$$\mathbf{X}' \left(\frac{S}{scale}, \frac{S}{scale}, scale^2 C_1 \right) \rightarrow \mathbf{X}'' \left(\frac{S}{scale}, \frac{S}{scale}, C_2 \right) \quad (4)$$

1.3 SegNeXt-Attention

SegNeXt^[12] 是一种针对语义分割任务优化的卷积神经网络架构,通过创新的卷积注意力机制,显著提升了分割性能,该架构由编码器和解码器组成,其中编码器采用金字塔结构,每个构建块借鉴了 ViT 的结构特点,但未采用自注意力,而是引入了多尺度卷积注意力(MSCA)模块。MSCA 模块的运作机制分为 3 个核心部分:深度卷积操作旨在聚合局部信息,从而有效地提取细节特征;而多分支深度条带卷积则通过并行处理多尺度信息,能够更好地捕捉多层次的上下文信息,每个分支使用不同尺寸的卷积核(如 7×7 、 11×11 和 21×21),通过深度可分离卷积近似实现大卷积核的效果,降低计算复杂度同时提取条状特征,

序列进行切片,实现了特征的高效融合与下采样。

$$\begin{aligned} f_{0,0} &= \mathbf{X}[0:S:scale, 0:S:scale], f_{1,0} = \mathbf{X}[1:S:scale, 0:S:scale], \dots, \\ f_{scale-1,0} &= \mathbf{X}[scale-1:S:scale, 0:S:scale]; \\ f_{0,1} &= \mathbf{X}[0:S:scale, 1:S:scale], f_{1,1}, \dots, \\ f_{scale-1,1} &= \mathbf{X}[scale-1:S:scale, 1:S:scale]; \\ f_{0,scale-1} &= \mathbf{X}[0:S:scale, scale-1:S:scale], f_{1,scale-1}, \dots, \\ f_{scale-1,scale-1} &= \mathbf{X}[scale-1:S:scale, scale-1:S:scale] \end{aligned} \quad (3)$$

对于任意设定的原始张量 \mathbf{X} 可通过多尺度分割策略实现特征重构,具体操作时将输入张量按几何分布模式进行区域划分,形成具有不同空间粒度的子空间表征。以比例因子 2 为例,实施四元空间分割生成 $(S/2, S/2, C_1)$ 维度的分块特征,随后沿通道轴执行多尺度特征金字塔融合,构建出空间维度压缩 2 倍的新特征矩阵 \mathbf{X}' , 其有效实现了原始数据的多分辨率表征。在通道轴上进行了倍增,扩展了 $scale^2$ 倍的规模,图 3 为当 $scale = 2$ 时的图示。

适应实际场景中物体形状; 1×1 卷积用于建模通道间关系,输出作为注意力权重,重新加权 MSCA 的输入特征。数学上, MSCA 的输出可以表示为:

$$Att = Conv_{1 \times 1} \left(\sum_{i=0}^3 Scale_i (DWConv(F)) \right) \quad (5)$$

其中, F 是输入特征, $DWConv$ 表示深度卷积, $Scale_i$ 表示不同尺度的卷积分支,这种设计使得 SegNeXt 能够有效地从局部到全局提取多尺度上下文信息,实现空间和通道维度的自适应性,从而在目标检测任务中表现出色。MSCA^[11] 结构图如图 4 所示。

解码器则从不同阶段收集多级特征,并使用 Hamburger 模块来进一步提取全局上下文。

在非机动车不规范行为检测领域, SegNeXt 架构的应用可以提高对非机动车细节特征的捕捉能力,尤其是在复杂交通场景中,通过多尺度卷积注意力机制,模型能够更好地识别非机动车的不规范行为,如违规载人、行驶中使用手机等这种对细节和多尺度上下文信息的敏感性,使得 SegNeXt 在提高检测精度和鲁棒性方面具有潜在优势,有

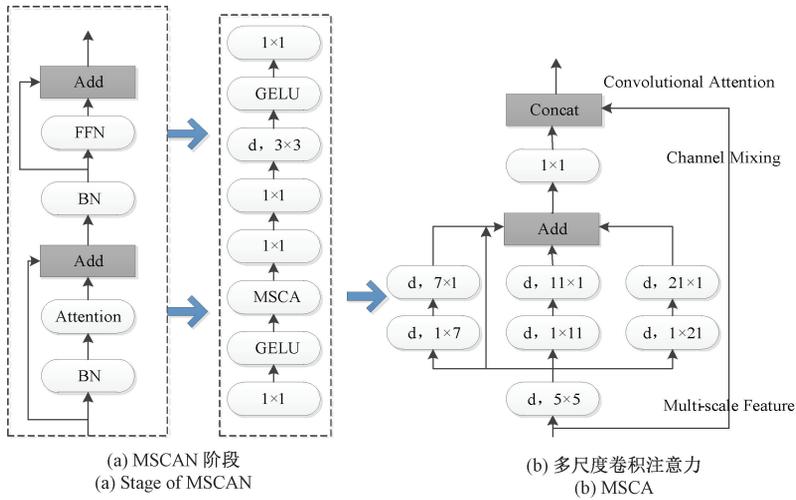


图 4 MSCA 和 MSCAN 图示

Fig. 4 Illustration of the proposed MSCA and MSCAN

助于减少漏检和误检,提升非机动车不规范行为检测的整体性能。

1.4 Wise-MPDIoU

原模型 YOLOv8 使用的是 CIoU^[13] 损失函数,CIoU 不仅考虑了预测框与真实框的重叠面积,同时引入了中心点距离以及长宽比,其计算公式如下:

$$CIoU = IoU - \left(\frac{\rho^2(B^{pred}, B^{gt})}{c^2} + \alpha v \right) \quad (6)$$

$$v = \frac{4}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^{pred}}{h^{pred}} \right) \quad (7)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (8)$$

在检测任务中, v 被用来衡量预测框与真实框的尺寸比, α 是一个稳定参数, $\rho^2(B^{pred}, B^{gt})$ 式(6)计算的是预测框 B^{pred} 与真实框 B^{gt} 几何质心的空间距离度量,其中参数 c 作为归一化补偿系数,其计算依据来源于真实框的横向跨度 w^{gt} 与纵向维度 h^{gt} ,同时模型预测框的宽高参数 w^{pred} 和 h^{pred} 参与空间对齐计算。即使原 CIoU 在衡量框的长宽比方面提供了有效的度量,现有方法在 anchor 长宽比例的参数化定义上存在模糊性,同时高精度回归实例的优化过程存在阻滞,这种双重困境将引发正负类别样本分布失衡现象,进而影响了模型的训练效果。为此,本文将原 CIoU 替换为 Wise-MPDIoU,用来解决正负样本不均衡所带来的问题。WIoU^[14] 通过对预测框和真实框之间的区域进行加权,解决了 IoU 可能存在的偏差问题,其计算公式如下:

$$WIoU = \frac{\sum_{i=1}^n \omega_i IoU(b_i, g_i)}{\sum_{i=1}^n \omega_i} \quad (9)$$

其中, b_i 代表第 i 个预测物体框的位置坐标, g_i 代表第 i 个物体标注框的真实坐标, ω_i 是与真实标注框相关联

的权重值。MPDIoU^[15] 旨在通过直接最小化预测框与真实框之间对角两点间的点距离,从而优化物体检测中的框定位精度。该方法在非机动车不规范行为在小目标检测任务中,模型展现了更优的准确性和鲁棒性,有效地提升了定位精度。其具体的计算公式如式(10)所示。

$$MPDIoU = IoU - \frac{\rho^2(P_1^{pred}, P_1^{gt})}{w^2 + h^2} - \frac{\rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \quad (10)$$

其中, $P_1^{pred}, P_1^{gt}, P_2^{pred}, P_2^{gt}$ 表示框与框之间对角线的点,度量函数 $\rho^2(P^{pred}, P^{gt})$ 用于量化预测点与真实点之间的空间间距,其中检测框的横向跨度和纵向维度分别由参数 w 和 h 进行参数化表征。本文使用 WIoU 损失函数的思想改进 MPDIoU,以 Wise-MPDIoU^[16] 替换 CIoU,有助于模型更好地学习目标位置的准确位置,提高模型的检测准确率,其计算公式如下:

$$Wise-MPDIoU = WIoU - \frac{\rho^2(P_1^{pred}, P_1^{gt})}{w^2 + h^2} - \frac{\rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \quad (11)$$

优化后的损失函数在非机动车不规范检测任务中,通过梯度修正机制有效抑制离群样本引发的异常梯度幅值与有害扰动,实现跨类别样本的梯度均衡,进而提升模型在未知数据域的泛化性能,最终达成高精度目标定位效果。

2 实验设计

2.1 数据集构建

本文所用的数据集拍摄于新疆维吾尔自治区乌鲁木齐市和库尔勒市,该数据集中有从不同角度和场景下拍摄的非机动车及违规行为的图像,总量为 2 664 张,将类别分为五类: non-motor (非机动车)、helmet (佩戴头盔)、withouthelmet (未佩戴头盔)、override (违规载人)、usephone (骑行使用手机),将数据集按照 8:1:1 的比例划分为训练

集、测试集与验证集。违规行为数量分布如表 1 所示。

表 1 违规行为数量分布表

Table 1 Distribution table of the number of violations

类别	数量
未佩戴头盔	1 596
佩戴头盔	1 733
非机动车	2 247
骑行使用手机	853
违规载人	786

2.2 实验环境和评估指标

实验操作系统为 Windows11、采用 GPU 型号为 RTX3090 显卡, 显存 24 GB, CPU 型号为 14vCPU Intel (R) Xeon(R) Gold 6330 CPU@2.00 GHz, 深度学习框架为 PyTorch2.0.0, 编程语言为 Python-3.8(ubuntu20.04)。

本文评价指标体系涵盖目标检测领域。对于目标检测模型的性能评估, 则依据精确率 (precision, P)、召回率 (recall, R)、平均精度均值 (mean average precision, mAP)、模型参数量 (parameters) 以及每秒十亿次浮点运算量 (giga floating-point operations per second, GFLOPs)^[17] 等指标进行。 $mAP@0.5$ 指的是在置信度阈值为 0.5 时, 对 N 个类别的平均精度均值进行计算, 而 $mAP@0.5-0.95$ 则是

在置信度阈值从 0.5~0.95 范围内, 以 0.05 为步长, 对 N 个类别的平均精度均值进行计算。这些指标的数值越高, 表明模型的性能越优越。非机动车不规范行为检测优先保障精度, FPS 仅在特定场景下成为关键约束。所以本研究不考虑 FPS, 所以本文各个评价指标的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$AP = \int_0^1 P(R) dR \quad (14)$$

$$mAP = \frac{1}{n} \sum AP \quad (15)$$

其中, TP 为被样本预测积极的正样本; FP 为被样本预测消极的负样本; FN 为被样本预测消极的正样本。

3 实验与结果分析

3.1 不同注意力机制分析对比

为探究注意力机制对模型的影响, 本研究在基模型 YOLOv8 架构中集成多种注意力, 通过 $mAP@0.5$ 指标开展对比测试。实验数据显示, 采用 SegNext 注意力机制的改进方案在多目标检测任务中展现出最优综合效能提升, 具体定量分析结果如表 2 所示。

表 2 不同注意力机制的对比

Table 2 Comparison of different attention mechanisms

模型	P/%	R/%	$mAP@0.5$ /%	$mAP@0.5-0.95$ /%
YOLOv8	86.1	84.6	89.1	60.4
YOLOv8-EMA ^[18]	85.4	86.3	89.4	58.7
YOLOv8-CBAM ^[19]	88.8	84.3	90.2	60.1
YOLOv8-SimAM ^[20]	87.7	83.4	90.4	60.1
YOLOv8-CPCA ^[21]	88.1	84.9	90.3	60.0
YOLOv8-SegNext	88.6	85.1	91.1	59.1

根据表 2 中的实验结果, 在精确率 P 方面, CBAM 注意力机制位居第一, 而本文所提出的 SegNext 在此指标上仅低于 CBAM0.02%。与原始模型 YOLOv8 相比, SegNext 的精确率提高了 2.5%。在召回率 R 方面本文所使用注意力机制略低于 EMA 注意力机制, 与基准模型相比, YOLOv8 升高了 0.7%; SegNext 注意力机制在平均精度方面性能最强, 与原模型相比提高了 2.0%, 与其他注意力机制相比平均提高了 1.22%。由上述可知, 对于精确率 P 或召回率 R 方面, 本文所使用注意力机制略低于 CBAM 与 EMA, 该指标未能全面反映模型在不同置信度阈值下的整体表现。就 $mAP@0.5$ 而言, SegNext 相较于 EMA 和 CBAM 表现更优。 $mAP@0.5$ 通过综合准确率与召回率双指标, 能够在动态调整概率判定阈值的多场景下, 全面评估模型的综合效能表现, 以更全面的方式评估模型性

能。基于全面评估的视角, 本文提出的改进型注意力机制 SegNext 在各项指标上优于其他注意力机制, 显著提升了非机动车不规范行为检测的准确性。

3.2 损失函数对比实验

为更好的说明 YOLO-CSSM 模型中使用 Wise-MPDIoU 损失函数的优越性, 设计损失函数对比实验, 在 YOLO-CSSM 模型中分别采用 WIoU 各个版本、CIoU、SIoU^[22]、MPDIoU 和 Wise-MPDIoU 损失函数。

由表 3 可知采用 Wise-MPDIoU 损失函数 $mAP@0.5$ 最高。Wise-MPDIoU 综合考虑了重叠与非重叠区域、中心点间距以及宽高比差异等多种因素, 同时优化了计算过程的复杂性。与其他损失函数相比, Wise-MPDIoU 通过最小化预测边界框和真实标注边界框对角线两点之间的距离, 实现了更高效的边界框回归。在非机动车不规范行

为检测中,Wise-MPDIoU 通过参数优化机制引导网络精准定位目标的几何中心,该损失函数设计显著提升了物体识别任务的定位精度。

表 3 不同损失函数的对比

Table 3 Comparison of different loss functions

损失函数	P/%	R/%	mAP@0.5/%
WIoUv1	87.6	85.4	90.5
WIoUv2	88.0	84.3	90.2
WIoUv3	90.1	90.2	92.7
CIoU	86.1	84.6	89.1
SIoU	87.2	85.3	91.3
MPDIoU	87.7	85.4	91.6
Wise-MPDIoU	89.4	91.5	93.6

表 4 改进结构消融实验对比结果

Table 4 Compare the results of improved structural ablation experiments

序号	C2f-DCNv2	SPD-Conv	SegNext Attention	Wise-MPDIoU	mAP@0.5/%	P/%	R/%	Params/ 10 ⁶
1					89.1	86.1	84.6	3.0
2	✓				91.7(+2.6%)	88.1	83.1	3.0
3	✓	✓			92.1(+3.0%)	89.8	81.4	3.5
4	✓		✓		92.7(+3.6%)	86.6	85.4	3.2
5	✓	✓	✓		93.1(+4.0%)	90.1	80.8	3.6
6	✓	✓	✓	✓	93.6(+4.5%)	84.4	87.6	3.6

从图 5 可以看出改进后的模型在收敛速度、稳定性、最终精度上都优于 YOLOv8,未出现过拟合、欠拟合的现象,说明改进过后的模型优于原模型。

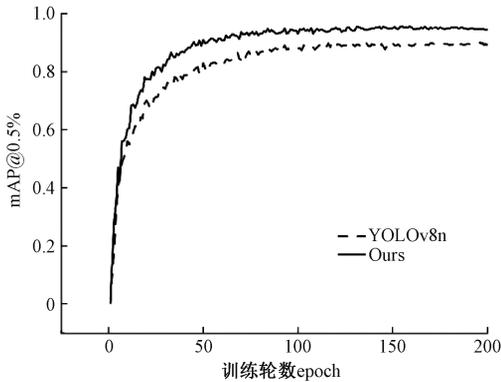


图 5 mAP@0.5/%对比曲线

Fig. 5 mAP@0.5/% comparison curve

3.4 不同算法实验对比

为验证改进算法 YOLO-CSSM 在非机动车不规范行为检测算法中的优越性,本文选取了 YOLOv3s、YOLOv6s、YOLOv9、RT-DETR 和 YOLOv10 等主流公开模型作为基准算法,同时鉴于当前非机动车违规行为检测领域的研究相对有限,特别引入基于头盔检测的优秀改进

3.3 结构消融实验对比

为验证本研究改进点的可行性,对改进后的模块做消融实验,结果如表 4 所示,从表 4 中可以看出在加入设计的 C2f-DCNv2 模块后 mAP@0.5 增加了 2.6%,而无参数数量的增加,说明 C2f-DCNv2 能在不影响模型大小的情况下能够提高其检测精度。由于 SPD-Conv 在网络中多加入了一层结构,所以增加了少许参数数量,但 mAP@0.5 提高了 2.6%。SegNextAttention 对 mAP@0.5 的提升效果和 SPD-Conv 几乎相同,参数量少量增加。最后,在使用 Wise-MPDIoU 优化损失函数后,能够解决遮挡,模糊等复杂性问题,提高了精度。改进过后的模型较于基础模型,mAP@0.5 显著提升,参数量少量增加,对非机动车不规范行为的检测有较好的检测效果。

算法与本文提出的方法进行对比分析。为全面评估算法性能,所有对比实验均在相同实验环境下进行,各对比算法的详细实验结果如表 5 所示。

表 5 不同算法对比实验结果

Table 5 Different algorithm comparison experimental results

模型	P/ %	R/ %	mAP@ 0.5/%	mAP@ 0.5-0.95/%	Params/ M
YOLOv3	83.5	87.2	87.8	58.0	103.7
YOLOv6	86.5	84.6	88.8	58.0	4.2
YOLOv8-ScG ^[8]	81.4	83.5	83.4	65.5	2.7
YOLOv9	84.2	83.2	88.8	59.8	2.0
RT-DETR	84.7	82.4	87.3	57.5	29.2
YOLOv10	83.4	82.9	88.8	58.1	2.7
Faster-RCNN	37.5	40.2	39.5	10.6	28.7
YOLOv11s	89.2	86.0	90.2	61.5	9.4
YOLOv11x	89.4	87.6	90.8	60.2	56.8
YOLOv12n	88.7	86.7	89.2	59.5	2.5
Helmet-YOLO ^[9]	85.8	83.9	88.5	60.7	12.4
YOLO-CSSM	89.4	90.0	93.6	60.2	3.6

YOLO-CSSM 算法在非机动车不规范行为自建数据集上的性能显著优于其他模型。以 89.4%的 mAP@0.5、

90.0%的 $mAP@0.5:0.95$ 和 93.6%的 Recall 全面领先所有对比模型,同时以仅 3.6 M 参数量实现轻量化,较 YOLOv11x(56.8 M)减少 93.7%参数却精度更高,且在小目标检测上优于多数主流模型。相较于 YOLOv8-ghost 与 Helmet-YOLO,其 $mAP@0.5$ 提升 8%,参数量仅增加 0.9 M;相比两阶段模型 Faster-RCNN,参数量仅为 12.5%但精度提升 51.9%。该模型在精度、效率与适应性上实现突破,尤其适合资源受限场景(如边缘设备、实时视频分析),相较于原始模型及其他对比模型,均展现出明显的优势。

3.5 泛化性对比实验

为验证本算法泛化能力,将 YOLOv8 与改进算法 YOLO-CSMM 分别在公开头盔数据集 TWH D 和自建数据集训练对比,由表 6 可以得出,YOLO-CSMM 在两个数据集上具有良好的泛化能力,在仅增加少量参数量的情况下满足实时检测的要求并且精度大幅度提高。

表 6 YOLOv8 和 YOLO-CSMM 在不同数据集上的对比实验结果

Table 6 Comparative experimental results of YOLOv8 and YOLO-CSMM on different datasets

数据集	模型	$mAP@0.5/\%$	Parameters/M
自建	YOLOv8	89.1	3.0
	YOLO-CSMM	93.6	3.6
TWH D	YOLOv8	80.6	3.0
	YOLO-CSMM	84.5	3.5

3.6 检测效果对比

为量化验证算法优化成效,本研究对基模型与改进模型在验证集上的性能开展对比实验。

图 6 与图 7 为两种模型检测结果对比图,选自不同设备、不同角度、不同时间与场景拍摄下的非机动车图像。当面对如图 6(a)、(c)与图 7(a)、(c)和目标较大的非机动车时,YOLOv8 和 YOLO-CSMM 都能准确识别,但面对如手机和远处的小目标非机动车时,YOLOv8 在提取非机动车特征时,由于卷积操作导致信息丢失,对于小目标来说,这会大幅度降低小目标的检测能力,而 YOLO-CSMM 模型采用 C2f-DCNv2 模块避免小目标丢失问题,能对小目标准确识别。如图 6(e)、(b)与图 7(e)、(b),YOLOv8 模型对冗余背景与非机动车较为关注,导致将路过行人错误为非机动车,YOLO-CSMM 模型通过 SegNext 注意力机制提高对非机动车的关注度,能对非机动车精准识别,如图 6(f)与图 7(f),YOLOv8 模型在小目标分类与定位损失不易计算,且手机等小目标易受图像背景及噪音干扰,导致模型在检测时产生重复框,而 YOLO-CSMM 模型采用 SPD-Conv 模块通过保留更多的细粒度信息和空间信息,SPD-Conv 有助于模型学习更有效的特征表示,减少了检测过程中的重复框问题。YOLO-CSMM 模型网络相较于 YOLOv8,在整体精度方面表现出了显著提升,并且在漏检问题上得到了有效改进。该网络能够更精确地检测到各种不规范行为以及小型目标,表现出较强的鲁棒性和检测性能,该算法在复杂环境中的目标识别准确率显著提升,展现出优异的场景适应性与跨域泛化能力。

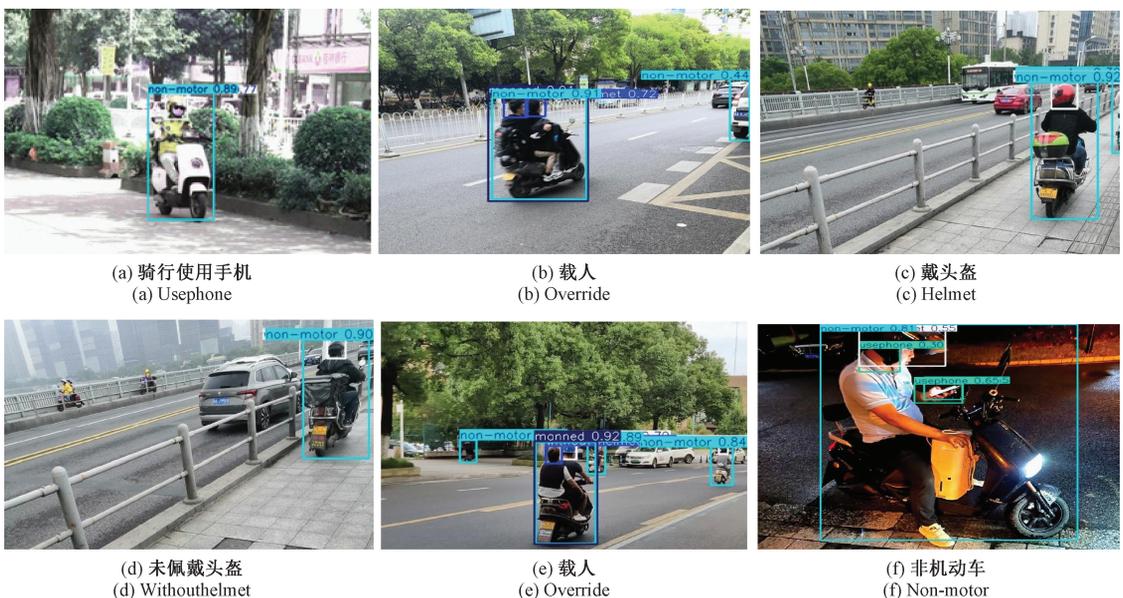


图 6 YOLOv8 模型检测结果

Fig. 6 YOLOv8 model detection results

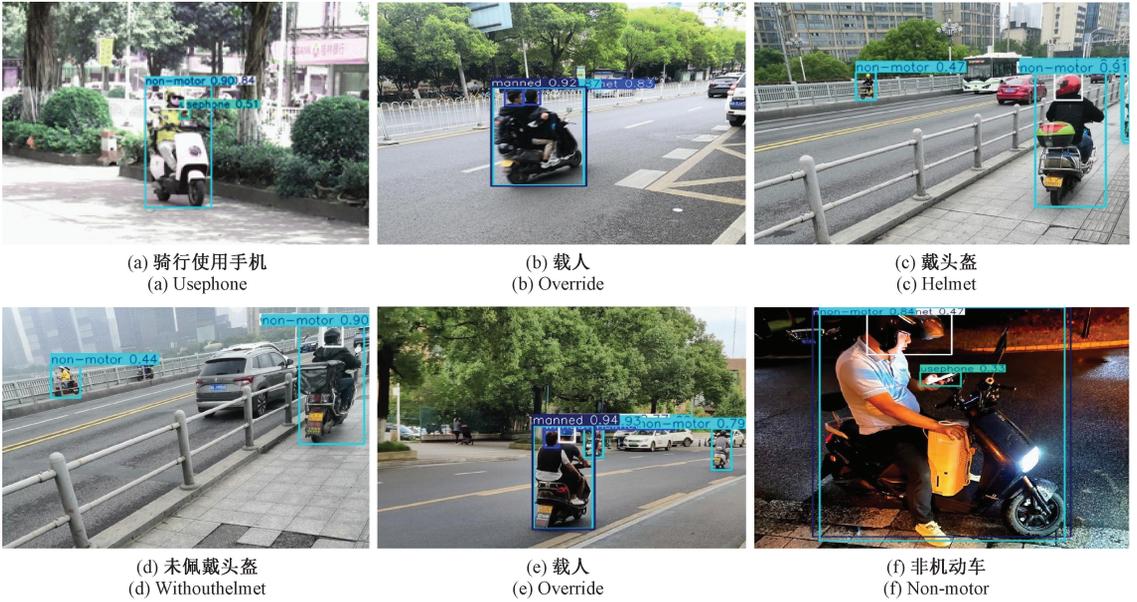


图 7 YOLO-CSSM 模型检测结果
Fig. 7 YOLO-CSSM model detection results

4 系统设计

4.1 系统架构

检测系统的开发基于 Python3.9 版本及 PyQt5, YOLOv8 作为当前主流的实时目标检测算法,具备结构轻量、检测速度快、精度高等优点,适用于本项目中非机动车违规行为的多目标场景的识别需求。在模型训练阶段,系统对包含头盔佩戴情况的图像数据集进行标注与增强,并基于改进的检测头模块进行训练优化,以提升检测精度并减少参数量。在推理阶段,系统支持图像与视频流输入,并对每一帧图像进行预处理操作(如缩放、归一化),随后输入至 YOLOv8 模型进行前向推理,输出包含类别标签、置信度和目标位置坐标的检测结果。

为实现对交通违法行为的自动判别,系统引入逻辑判别模块,通过分析检测框间的空间关系,结合类别标签与置信度信息,自动识别违规行为并进行分类统计。检测结果将实时显示于前端界面中,同时在结果列表中同步展示图像路径、类别、置信度以及对应坐标信息,以便后续存档与查阅。系统的架构设计如图 8 所示。

4.2 系统功能

非机动车危险驾驶识别系统的核心是将从图片、视频和摄像头读取到非机动车驾乘人员驾驶行为信息进行检测识别。因此,在系统内共设计开发三类系统检测模块,分别是:图片检测模块、视频检测模块和摄像头检测模块。检测系统页面如图 9 所示。

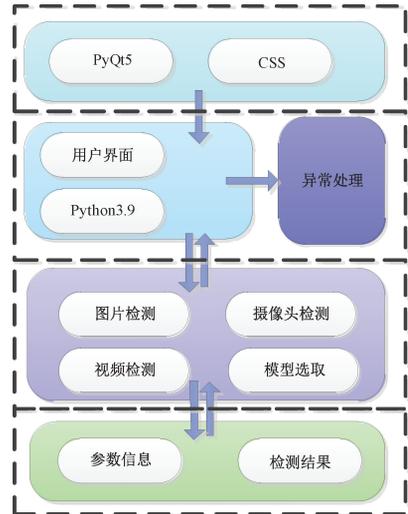


图 8 系统架构
Fig. 8 System architecture



图 9 页面展示
Fig. 9 Page display

5 结 论

为解决非机动车驾驶场景中漏检误检等问题以及对高精度检测的需求,提出了一种改进 YOLOv8 的非机动车不规范行为算法。首先在 Backbone 和 Neck 中引入 SPD-Conv 网络模块,以加强模型对微小目标的学习能力及复杂背景条件下的特征提取能力;其次分别在 Backbone 和 Neck 网络中设计 C2f-DCNv2 和 SegNext 模块,突出非机动车和驾驶人重要特征信息,提高模型特征融合能力;最后使用 WIoU 损失函数的思想改进 MPDIoU,将原 CIoU 替换为 Wise-MPDIoU,用来解决正负样本不均衡所带来的问题。实验表明,改进后的 mAP@0.5 提升至 93.6%,相比基模型提升了 4.5%,漏检率和误检率显著降低。最后以非机动车违规行为检测算法为基础,使用 PyQT5 设计开发了非机动车违规行为识别检测系统。且检测速度满足工程部署需求,未来将在满足实时性约束条件的前提下,后续研究将聚焦于架构轻量化设计,通过双重优化路径提升推理效率与跨场景适应能力。同步实施参数精简策略以降低由参数冗余引发的过拟合风险,从而增强模型对动态多变的交通场景感知需求的适配性。

参考文献

- [1] 董苓. 非机动车骑行者不安全行为风险感知研究[D]. 北京:北京交通大学,2015.
DONG L. Research on risk perception of unsafe behaviors of non-motorized vehicle riders[D]. Beijing: Beijing Jiaotong University, 2015.
- [2] 刘思雨. 驾乘电动自行车戴安全头盔的法律规制及完善建议[J]. 公关世界,2022(4):178-180.
LIU S Y. Legal regulation and improvement suggestions for wearing safety helmets when riding electric bicycles[J]. Public Relations World, 2022(4): 178-180.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016; 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.
- [5] 晏舒童. 基于目标检测与跟踪的非机动车违规行为为识别的研究与应用[D]. 南昌:南昌大学,2022.
YAN SH T. Research and application of non-motor vehicle violation recognition based on object detection and tracking[D]. Nanchang: Nanchang University, 2022.
- [6] 马斌. 基于深度学习的非机动车头盔佩戴检测算法研究[D]. 合肥:安徽理工大学,2024.
MA B. Research on non-motorized vehicle helmet wearing detection algorithm based on deep learning [D]. Hefei: Anhui University of Science and Technology, 2024.
- [7] 张国鹏,周金治,马光岑,等. 改进 YOLOv8 的轻量化安全帽佩戴检测算法[J]. 电子测量技术, 2024, 47(17):147-154.
ZHANG G P, ZHOU J ZH, MA G C, et al. Improved lightweight helmet wearing detection algorithm of YOLOv8 [J]. Electronic Measurement Technology, 2024, 47(17):147-154.
- [8] 杨璐钱,胡平,戴家树. 基于 Yolov8-scG 神经网络的电动车头盔佩戴检测算法[J/OL]. 重庆工商大学学报(自然科学版), 1-10 [2025-05-16]. <http://kns.cnki.net/kcms/detail/50.1155.N.20240710.2030.002.html>.
YANG J Q, HU P, DAI J SH. Electric vehicle helmet wearing detection algorithm based on Yolov8 scG neural network [J/OL]. Journal of Chongqing Technology and Business University(Natural Science Edition), 1-10 [2025-05-16]. <http://kns.cnki.net/kcms/detail/50.1155.N.20240710.2030.002.html>.
- [9] 周顺勇,彭梓洋,张航领,等. Helmet-YOLO:一种更高精度的道路安全头盔检测算法[J]. 计算机工程与应用, 2025, 61(2): 135-144.
ZHOU SH Y, PENG Z Y, ZHANG H L, et al. Helmet YOLO: A higher precision algorithm for detecting road safety helmets [J]. Computer Engineering and Applications, 2025, 61(2): 135-144.
- [10] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9308-9316.
- [11] 叶学义,韩卓,蒋甜甜,等. 多尺度反向校正增强和无损下采样的毫米波图像目标检测方法[J]. 电子测量与仪器学报,2025,39(4):50-61.
YE X Y, HAN ZH, JIANG T T, et al. Millimeter-wave image target detection method based on multi-scale reverse correction enhancement and lossless downsampling[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(4):50-61.
- [12] GUO M H, LU C Z, HOU Q, et al. Segnext: Rethinking convolutional attention design for semantic segmentation [J]. Advances in Neural Information Processing Systems, 2022, 35: 1140-1156.
- [13] WANG X, SONG J. CIoU: Improved loss based on

- complete intersection over union for bounding box regression[J]. IEEE Access, 2021, 9: 105686-105695.
- [14] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism [J]. ArXiv preprint arXiv: 2301.10051, 2023.
- [15] MA S, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662, 2023.
- [16] 闫世洋, 罗素云. 基于 SC-YOLOv8 的交通标志检测算法研究[J]. 电子测量技术, 2024, 47(15): 117-124.
YAN SH Y, LUO S Y. Research on traffic sign detection algorithm based on SC-YOLOv8 [J]. Electronic Measurement Technology, 2024, 47(15): 117-124.
- [17] LIU Z, SUN B, BI K. Optimization of YOLOv7 based on PConv, SE Attention and Wise-IoU [J]. International Journal of Computational Intelligence and Applications, 2024, 23(1): 2350033.
- [18] 朱圣博, 魏利胜, 高港, 等. 基于改进 YOLOv8s 的光学遥感小型船舶检测算法[J]. 电子测量与仪器学报, 2024, 38(10): 48-57.
ZHU SH B, WEI L SH, GAO G, et al. Optical remote sensing detection algorithm for small ships based on improved YOLOv8s [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(10): 48-57.
- [19] 王天洋, 刘路, 王太勇, 等. 基于改进 YOLOv8s 的轻量级 PCB 缺陷检测算法[J]. 电子测量与仪器学报, 2025, 39(3): 44-52.
- WANG T Y, LIU L, WANG T Y, et al. Lightweight PCB defect detection algorithm based on improved YOLOv8s[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(3): 44-52.
- [20] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks [C]. International Conference on Machine Learning. PMLR, 2021: 11863-11874.
- [21] HUANG H, CHEN Z, ZOU Y, et al. Channel prior convolutional attention for medical image segmentation[J]. Computers in Biology and Medicine, 2024, 178: 108784.
- [22] GEVORGYAN Z. SIoU loss: More powerful learning for bounding box regression [J]. ArXiv preprint arXiv:2205.12740, 2022.

作者简介

李强, 硕士研究生, 主要研究方向为目标检测、图象识别。

E-mail: 2529467927@qq.com

南新元(通信作者), 教授, 主要研究方向为流程工业系统控制与优化、智能控制系统开发。

E-mail: xynan@xju.edu.cn

蔡鑫, 副教授, 主要研究方向为多智能体系统协同控制、分布式优化与博弈。

E-mail: xincai@xju.edu.cn

杨仕伟, 硕士研究生, 主要研究方向为计算机视觉。

E-mail: ysw12130546@163.com