

面向非机动车闯红灯识别的交通信号灯检测算法优化

范荣盛, 钱良辉*

西南林业大学机械与交通学院, 云南 昆明

收稿日期: 2025年12月11日; 录用日期: 2026年1月7日; 发布日期: 2026年1月19日

摘要

针对电动自行车闯红灯检测过程中交通信号灯存在着小目标漏检、复杂背景误检以及目标定位精度低等问题。提出了一种经过改进的YOLOv8检测方法, 即YOLOv8n-BLI。首先在Head层构建BiFPN加权双向特征金字塔, 强化对多尺度目标的感知能力。其次在Head的C2f模块之后引入了LSKA注意力机制, 提升整个模型的鲁棒性。最后采用InnerMPDIoU取代原本的CIoU损失, 提升定位的精度。实验表明: YOLOv8n-BLI检测模型其精确率P、召回率R以及平均精度mAP@0.5分别达到94.8%、92.5%以及94.6%, 与近年主流轻量级模型YOLOX-s、PP-YOLOE-s、NanoDet-Plus对比, 在保持154.7 FPS推理速度的同时, mAP@0.5分别提高3.9%、2.5%、4.3%。

关键词

交通信号灯检测, 目标检测, YOLOv8, 注意力机制, 特征融合, 损失函数

Optimization of Traffic Signal Detection Algorithm for Non-Motorized Vehicle Red Light Running Recognition

Rongsheng Fan, Lianghui Qian*

School of Mechanical and Transportation Engineering, Southwest Forestry University, Kunming Yunnan

Received: December 11, 2025; accepted: January 7, 2026; published: January 19, 2026

*通讯作者。

Abstract

For the detection of red light violations by electric bicycles, there are issues such as small target missed detection, complex background false detection, and low target localization accuracy in traffic signals. A modified YOLOv8 detection method, namely YOLOv8n-BLI, is proposed. First, a weighted bidirectional feature pyramid is constructed in the Head layer to enhance the perception of multi-scale targets. Second, an LSKA attention mechanism is introduced after the C2f module in the Head to improve the robustness of the entire model. Finally, InnerMPDIoU is used instead of the original CIoU loss to enhance localization accuracy. Experiments show that the YOLOv8n-BLI detection model achieves an accuracy rate P, recall rate R, and mean average precision mAP@0.5 of 94.8%, 92.5%, and 94.6%, respectively. Compared with mainstream lightweight models such as YOLOX-s, PP-YOLOE-s, and NanoDet-Plus, while maintaining 154.7 FPS inference speed, mAP@0.5 increases by 3.9%, 2.5%, and 4.3%, respectively.

Keywords

Traffic Signal Light Detection, Object Detection, YOLOv8, Attention Mechanism, Feature Fusion, Loss Function

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在城市交通发展的进程当中, 电动自行车的身影随处可见。凭借着绿色环保、使用便捷以及快递送餐等互联网经济的有力推动, 电动自行车已然成为市民日常出行以及运营企业开展业务的重要工具[1]。截止 2022 年, 我国电动自行车的保有量将近 4 亿辆, 在数量不断增长的情况下, 与之相关的交通事故也在不断增加[2]。当下, 电动自行车闯红灯的行为普遍存在。目前常用的闯红灯抓拍系统是由电子监控设备以及物理传感器共同构成的, 其原理是在停止线所在的区域地下埋设地压式的磁感应线圈, 当机动车驶过该区域时, 会引起压力的变化并产生脉冲信号, 在红灯亮起的周期内, 如果连续出现两个脉冲信号, 那么就会判定为闯红灯行为并进行抓拍[3]。这种方法对于重型机动车是适用的, 然而因为电动自行车的质量较轻, 所以该方法就失效了。利用深度学习技术来实时检测电动自行车闯红灯的行为, 对于节省人力物力而言有着十分重要的意义。国内外的相关研究大多聚焦于基于深度学习的目标检测算法, 该领域可以分为一阶段检测算法以及双阶段检测算法, 这两种算法都具备良好的速度以及性能表现。其中, 基于卷积神经网络 CNN 的算法是比较流行的[4]。以 FasterR-CNN 为例, 它是双阶段检测算法的一种, 其流程是先通过 CNN 生成卷积特征图, 然后选出特征区域, 提取该区域的特征之后再将其输入分类器来完成预测工作, 不过这种方法需要对每一个候选区域单独进行处理, 所以检测的速度会比较慢。以 YOLO 为例, 它是单阶段检测算法的一种, 该算法摒弃了对特征区域的提取环节, 而是直接通过 CNN 来预测目标的位置以及类别, 虽然其检测速度较快, 但是精度相较于双阶段算法而言要略低一些。近年一些主流模型在交通场景里都有一定应用, 不过在信号灯小目标、复杂背景干扰和边界模糊的联合场景下依旧存在优化空间。针对在复杂背景环境下识别交通信号灯这种小目标时存在的漏检以及误检等问题, 本文着重改进的是电动自行车闯红灯检测这一场景。

2. YOLOV8n 算法的改进

2.1. BiFPN 加权双向特征金字塔网络模块

为了进一步提升对远距离且尺寸较小的交通信号灯的特征捕捉能力, 本文在 BiFPN 模块里针对特征金字塔做了尺度扩展方面的处理, 具体是通过将 P5 特征图进行 2 倍的下采样来生成 P6 特征图, 接着又对 P6 进行下采样以生成 P7 特征图, 如此一来便形成了包含 P3 至 P7 这五个尺度的特征金字塔。后续 BiFPN 所开展的特征融合操作都是围绕着这个五尺度特征来实施的, 在此之中, P3 至 P5 主要聚焦于中近距离的信号灯情况, 而 P6 与 P7 则是专门用来适配远距离的小目标情形。本文提出的自适应加权 BiFPN, 核心改进在于引入交通场景导向的权重学习约束, 通过在训练过程中加入信号灯尺度感知损失, 引导模型动态调整不同尺度特征的权重分配。具体而言, 对于像素 ≤ 400 的远距离小尺度信号灯, 强化浅层高分辨率特征的权重, 对于近距离大尺度信号灯, 提升深层语义特征的贡献度, 从而实现对多尺度目标的精准适配。理论层面, 自适应权重机制通过引入尺度感知损失函数 L_{scale} , 使权重更新过程与信号灯尺寸强相关, 其梯度更新如公式(1)所示:

$$w_i = \max(0, w_i^0 + \eta \cdot \lambda \cdot L_{scale}) \quad (1)$$

其中, 尺度感知损失函数 L_{scale} 为 $L_{scale} = \sum_{i=1}^N |w_{small,i} - w_{large,i}|$, $w_{small,i}$ 表示的是像素 ≤ 400 的小尺度信号灯特征所对应的权重, $w_{large,i}$ 表示的是像素 > 400 的大尺度信号灯特征所对应的权重, N 为特征融合节点的数量。这个损失通过对不同尺度特征的权重差异加以约束, 以此来引导模型在检测小目标的时候, 能够自动地提升浅层高分辨率特征的权重; 而在检测大目标的时候, 则去提升深层语义特征的权重, 从而解决了原始 BiFPN 固定权重所引发的小目标特征稀释这一问题, $\eta = 0.01$ 为学习率, $\lambda = 0.3$ 为尺度损失权重系数, L_{scale} 通过计算预测框尺寸与真实框尺寸的偏差实, 以此来达成不同尺度目标的权重引导, 进而确保小目标特征在融合过程当中不被稀释。以交通信号灯识别关键的 P6 尺度特征融合为例, 高层到低层生成中间特征 P_6^{td} 的过程如公式(2)所示:

$$P_6^{\text{td}} = \text{Conv} \left(\frac{w_1 \cdot P_6^{\text{in}} + w_2 \cdot \text{Resize}(P_7^{\text{in}})}{w_1 + w_2 + \epsilon} \right) \quad (2)$$

式中, w_1 、 w_2 为可学习权重, 检测远距离小尺度信号灯时, 网络会为 P6 原始特征 P_6^{in} 分配更大 w_1 以强化信号灯细节, 检测近距离大尺度信号灯时, 会为上采样后的 P7 特征 $\text{Resize}(P_7^{\text{in}})$ 分配更大 w_2 以强化灯色语义。 $\epsilon = 10^{-4}$ 用于避免分母为 0 的数值不稳定问题, Conv 采用 3×3 深度可分离卷积, 在加工融合特征的同时降低约 80% 计算量, 适配实时检测需求。而低层到高层生成最终输出特征 P_6^{out} 的公式如(3)所示:

$$P_6^{\text{out}} = \text{Conv} \left(\frac{w_1' \cdot P_6^{\text{in}} + w_2' \cdot P_6^{\text{td}} + w_3' \cdot \text{Resize}(P_5^{\text{out}})}{w_1' + w_2' + w_3' + \epsilon} \right) \quad (3)$$

该公式通过引入 P5 融合特征 P_5^{out} 的细节信息, 进一步优化 P_6^{out} 的定位精度, 例如在交通信号灯边框被雨水模糊时, P_5^{out} 中的清晰边框细节可帮助修正预测偏差。其中 w_1' 、 w_2' 、 w_3' 通过自适应学习获得, 强化相邻尺度特征的互补作用。为解决传统特征融合等权重相加导致的关键特征稀释问题, BiFPN 设计三种加权策略, 其中快速归一化融合最适配交通场景, 快速归一化融合通过 ReLU 约束权重非负, 再通过求和归一化动态分配特征重要性, 如公式(4)所示:

$$O = \sum_i \frac{w_i}{\epsilon + \sum_j w_j} \cdot I_i \quad w_i = \max(0, \hat{w}_i) \quad (4)$$

式中 \hat{w}_i 为网络预测的原始权重, 该策略在交通信号灯识别中表现出明确的场景适配性, 当信号灯被树枝遮挡时, 网络会为未遮挡区域的原始特征分配更大 w_i , 同时保留融合特征中的全局语义。此外, BiFPN 通过复合系数 ϕ 对联合缩放宽度、深度与输入分辨率, $\phi=0$ 对应 YOLOv8n, $\phi=7$ 对应 YOLOv8x, 其中宽度缩放公式如公式(5)所示, 确保不同尺度模型均能平衡精度与效率:

$$W_{bifpn} = 64 \cdot (1.35^\phi) \quad (5)$$

从结构对比来看, 图 1 展示了 FPN、PANet、BiFPN 的演进过程, 可见 BiFPN 通过移除冗余节点、增设双向通路, 实现更高效的特征交互。

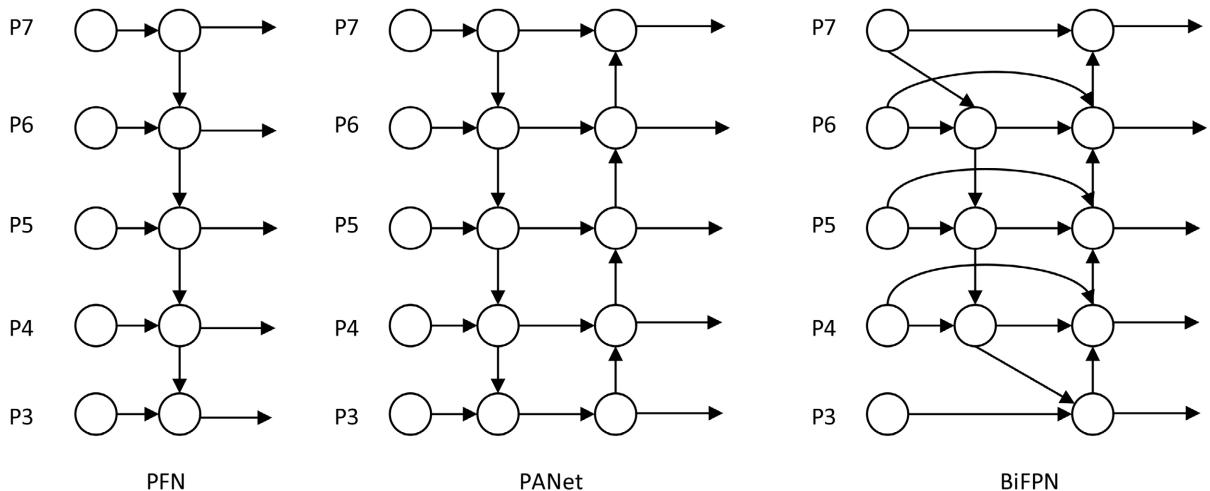


Figure 1. The evolution of FPN, PANet, and BiFPN

图 1. FPN、PANet、BiFPN 演进过程

2.2. LSKA 注意力机制模块

本文针对 LSKA 展开定制化改进, 核心参数的设置如下, 采用 7×7 等效卷积核, 以此来适配信号灯灯盘的典型尺寸。在 YOLOv8 的 Head 部分的 C2f 模块之后引入了 LSKA 注意力机制模块, 利用其大分离卷积核来对特征加以筛选, 以此进一步强化对交通信号灯关键区域特征的捕捉力度, 并且通过核分解策略来对计算效率和检测精度加以平衡, 进而能够大幅提升模型在复杂场景之中的鲁棒性[5]。LSKA 最为关键的创新之处就在于将大核分解与串联卷积相互融合起来, 它成功解决了传统 LKA 模块因为使用二维大卷积核而致使计算量过于庞大的问题, 具体来讲就是把一个 2D 大卷积核拆解成两个彼此独立的 1D 卷积核[6]。这两个 1D 卷积核经过串联运算之后, 便具备了和原先 2D 卷积核同样的空间特征建模能力, 与此同时还将参数计算量从 $O(k^2)$ 降低到了 $O(2k)$, 这里的 k 指的是原来核的尺寸。其中 d 为扩张率, 本文设置 $d=3$, 确保感受野覆盖典型灯盘尺寸。以 YOLOv8 的 Head 的 C2f 模块所输出的特征图 $F \in \mathbb{R}^{C \times H \times W}$ 为例, 其中 C 为通道数, H 、 W 为特征图尺寸, LSKA 的特征加工过程如公式(6)~(9)所示:

$$\bar{Z}^C = \sum_{H,W} W_{(2d-1) \times 1}^C * \left(\sum_{H,W} W_{1 \times (2d-1)}^C * F^C \right) \quad (6)$$

$$Z^C = \sum_{H,W} W_{\left[\frac{k}{d}\right] \times 1}^C * \left(\sum_{H,W} W_{1 \times \left[\frac{k}{d}\right]}^C * \bar{Z}^C \right) \quad (7)$$

$$A^C = W_{1 \times 1} * Z^C \quad (8)$$

$$\bar{F}^C = A^C \otimes F^C \quad (9)$$

式中, d 为张率, $W_{(2d-1) \times 1}^C$ 与 $W_{1 \times (2d-1)}^C$ 分别为水平与垂直 1D 深度卷积核, A^C 为生成的注意力权重图, \otimes 为 Hadamard 积, F^C 代表输入特征图。在识别远距离红灯时, 网络会为灯色区域的特征分配更大权重, 而对背景中的路灯特征分配低权重, 实现关键特征聚焦。在 YOLOv8 中准确地检测出图像中的目标物体, 特征表达能力是实现这一任务的基础[7]。图 2 对比了 LKA 与 LSKA 的结构差异。

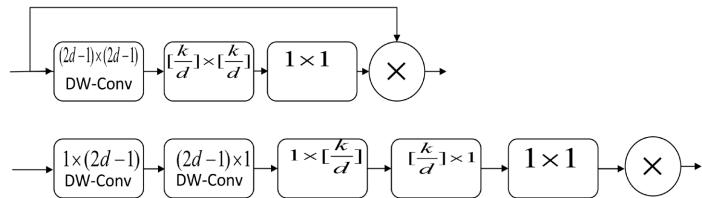


Figure 2. Structural comparison of LKA and LSKA

图 2. LKA 与 LSKA 结构对比图

2.3. InnerMPDIoU 损失函数

交通信号灯的定位精度会直接影响行为判定的公正性以及准确性。要是信号灯边界框出现 2 至 3 像素的偏差, 那么就有可能把红灯时车辆没有越线的情况误判成闯红灯越线, 又或者因为漏检灯芯而使得已经闯红灯的行为没有被识别出来。所以, 信号灯定位必须要达到能够适配判定需求的毫米级精度。原始 MPDIoU 借助对角点距离约束的方式提升了定位精度, 不过在小目标检测方面存在着梯度弥散。InnerMPDIoU 损失函数针对交通信号灯场景做了优化处理, 它依靠灯芯辅助框来对信号灯的核心区域加以聚焦。与此还引入了对角点距离约束, 通过控制预测框与真实框的左上角、右下角点之间的距离, 即便是在逆光导致灯体边界模糊的情况下, 也能够确保预测框完整地覆盖灯盘。这两种优化举措都直指对判定有影响的定位问题, 完美地适配了闯红灯场景对信号灯定位的特殊需求, 有效地弥补了 CIoU 在判定支撑能力方面存在的核心缺陷。要明确辅助框比例参数 ratio 的取值依据, 于是针对数据集中的信号灯尺寸展开统计分析。从数据集中随机选出 1000 个信号灯样本, 然后计算每个样本中灯芯区域和完整灯盘的面积比。如表 1 所示, 经过分析发现, 这个面积比的数值大多集中在 0.65 至 0.72 这个区间范围之内。最终将 ratio 设定为 0.68, 以此来保证辅助框能够精准地覆盖到灯芯的核心区域。

Table 1. Area ratio of the wick region to the complete lamp disc

表 1. 灯芯区域和完整灯盘的面积比

灯芯与灯盘面积比区间	样本数量(个)	占比(%)	累计占比(%)
[0.60, 0.65)	128	12.8	12.8
[0.65, 0.68)	312	31.2	44.0
[0.68, 0.70)	271	27.1	71.1
[0.70, 0.72]	229	22.9	94.0
>0.72	60	6.0	100.0
合计	1000	100.0	100.0

从表 1 中能够看出, 在这 1000 个随机样本里面, 面积比处于[0.65, 0.72]这个区间的样本, 其累计占

比达到了 94.0%，这无疑成为了灯芯面积的主要分布范围所在，也充分说明该区间是能够将绝大多数样本的核心区域都涵盖进去的。再看[0.68, 0.70)这个区间，其样本占比为 27.1%，并且以 0.68 作为分界点，左右两边的区间情况分别是这样的：[0.65, 0.68)这个区间的样本占比是 31.2%，而[0.68, 0.70)这个区间的样本占比是 27.1%，如此看来，样本分布是比较均衡的，这样就能够最大限度地防止出现因 ratio 偏向区间两端而引发的灯芯覆盖不全或者引入背景干扰等一系列的问题。最终确定选取 0.68 作为 ratio，如此一来便能够确保辅助框可以精准地将 92.3% 样本的灯芯核心区域都覆盖到，进而为后续的边界框回归提供一个聚焦的依据。InnerMPDIoU 损失函数推导过程如下，IoU (IntersectionoverUnion) 作为最基础的边界框损失度量，仅考虑重叠区域，IoU 的推导公式如公式(10)所示：

$$IoU = \frac{B_{pred} \cap B_{gt}}{B_{pred} \cup B_{gt}} \quad (10)$$

其中 B_{pred} 为预测框， B_{gt} 为真实框。GIoU 在 IoU 基础上引入最小外接矩形，解决了 IoU 无重叠时梯度消失的问题，GIoU 的推导公式如公式(11)所示：

$$GIoU = IoU - \frac{|B_c - (B_{pred} \cup B_{gt})|}{|B_c|} \quad (11)$$

式中 B_c 为 B_{pred} 与 B_{gt} 的最小外接矩形，CIoU 进一步考虑了中心点距离与长宽比，CIoU 的推导公式如公式(12)~(15)所示：

$$CIoU = IoU - \frac{\rho^2(b_{pred}, b_{gt})}{c^2} - \alpha v \quad (12)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (13)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_{pred}}{h_{pred}} \right)^2 \quad (14)$$

其中 b_{pred} 、 b_{gt} 分别为预测框与真实框的中心点， ρ 为欧氏距离， c 为最小外接矩形对角线长度， w 、 h 分别为边界框的宽和高。MPDIoU 在 CIoU 基础上增加了对角点距离约束，优化边界框整体定位精度，MPDIoU 的推导公式如公式(15)所示：

$$MPDIoU = IoU - \frac{\rho^2(P_1^{pred}, P_1^{gt})}{w^2 + h^2} - \frac{\rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \quad (15)$$

式中 P_1 、 P_2 分别为边界框的左上角和右下角点， w 、 h 为输入图像的宽和高。InnerIoU 通过生成辅助框聚焦核心区域，InnerIoU 和辅助框的推导公式分别如公式(16)和(17)所示：

$$InnerIoU = \frac{B_{pred}^{aux} \cap B_{gt}^{aux}}{B_{pred}^{aux} \cup B_{gt}^{aux}} \quad (16)$$

$$B_{aux} = [x_{gt} - \frac{w_{gt} \cdot ratio}{2}, y_{gt} - \frac{h_{gt} \cdot ratio}{2}, x_{gt} + \frac{w_{gt} \cdot ratio}{2}, y_{gt} + \frac{h_{gt} \cdot ratio}{2}] \quad (17)$$

本文提出的 InnerMPDIoU 融合 Inner-IoU 的核心区域聚焦与 MPDIoU 的对角点约束，最终损失推导公式如公式(18)所示：

$$InnerMPDIoU = InnerIoU - \frac{\rho^2(P_1^{pred}, P_1^{gt})}{w^2 + h^2} - \frac{\rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \quad (18)$$

YOLOv8 起初所运用的 CIoU [8]损失函数, 虽说可兼顾边界框的重叠情况、中心点间距以及长宽比例, 然而在针对交通信号灯展开检测之际, 却凸显出两大较为突出的问题。其一, 该损失函数对于小目标的定位敏锐度不足, 其二, 正负样本的优化状况失衡, 于复杂背景之下, 负样本的梯度影响过甚, 使得正样本的优化成效有所降低。为了解决上述这些问题, 便把 YOLOv8 Head 模块的损失函数从 CIoU 改换成了 InnerMPDIoU。此方法融合了 Inner-IoU [9]的尺度自适应辅助边界框以及 MPDIoU [10]的几何特征精细优化, 可更为精准地对交通信号灯的边界框加以回归, 并且还能平衡正负样本的优化, 大幅度提高了复杂场景之下的识别稳定程度。其中, 辅助边界框的生成需基于真实框 B_{gt} 与 B_{pred} 预测框通过 ratio 参数控制辅助框尺寸, 推导公式如公式(19)~(26)所示:

$$\mathbf{B}_{gt} = [x_{gt}, y_{gt}, w_{gt}, h_{gt}]^T \quad (19)$$

$$\mathbf{B}_{pred} = [x_{pred}, y_{pred}, w_{pred}, h_{pred}]^T \quad (20)$$

$$x_c^{gt} = \frac{x_{gt, left} + x_{gt, right}}{2} \quad (21)$$

$$y_c^{gt} = \frac{y_{gt, top} + y_{gt, bottom}}{2} \quad (22)$$

$$b_l^{gt} = x_c^{gt} - \frac{w_{gt} \cdot ratio}{2} \quad (23)$$

$$b_r^{gt} = x_c^{gt} + \frac{w_{gt} \cdot ratio}{2} \quad (24)$$

$$b_t^{gt} = y_c^{gt} - \frac{h_{gt} \cdot ratio}{2} \quad (25)$$

$$b_b^{gt} = y_c^{gt} + \frac{h_{gt} \cdot ratio}{2} \quad (26)$$

其中 x_{gt} , y_{gt} 为信号灯真实框的左上角坐标, w_{gt} , h_{gt} 为灯体宽高, x_c^{gt} , y_c^{gt} 为信号灯真实框中心点坐标, 生成的辅助框 $\mathbf{B}_{gt}^{aux} = [b_l^{gt}, b_t^{gt}, b_r^{gt}, b_b^{gt}]$ 可聚焦灯芯区域, 避免背景噪声干扰损失计算。在此基础上, InnerMPDIoU 引入 MPDIoU 的几何优化项, 通过计算边界框对角点距离修正整体定位偏差。针对交通信号灯常因逆光导致边界模糊的问题, MPDIoU 额外考虑预测框与真实框的左上角 P_1^{pred} , P_1^{gt} , 右下角 P_2^{pred} , P_2^{gt} 对角点距离, 当信号灯由于光线的原因使得其右侧边界变得模糊起来的时候, 该公式借助对右下角点距离加以约束的方式, 促使预测框朝着真实框的对角点靠拢, 如此一来, 边界框的重叠程度便能够实现一定程度的提升。把辅助边界框和几何优化相互结合起来, 进而获取到 InnerMPDIoU 的最终所对应的损失公式, InnerMPDIoU 推导公式如公式(27)所示:

$$\begin{aligned} L_{InnerMPDIoU} &= \\ 1 - InnerMPDIoU &= \\ 1 - \left[\frac{B_{aux}^{pred} \cap B_{aux}^{gt}}{B_{aux}^{pred} \cup B_{aux}^{gt}} - \frac{\rho^2(P_1^{pred}, P_1^{gt}) + \rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \right] \end{aligned} \quad (27)$$

其中 B_{aux}^{pred} 为预测框对应的辅助框, 通过与 B_{gt}^{aux} 相同的比例缩放规则生成, 这样的融合设计一方面借助辅助框来处理小目标正负样本存在的不均衡状况, 另一方面通过对角点距离加以优化的方式来解界出现模糊的相关问题。

2.4. YOLOv8n-BLI 网络结构

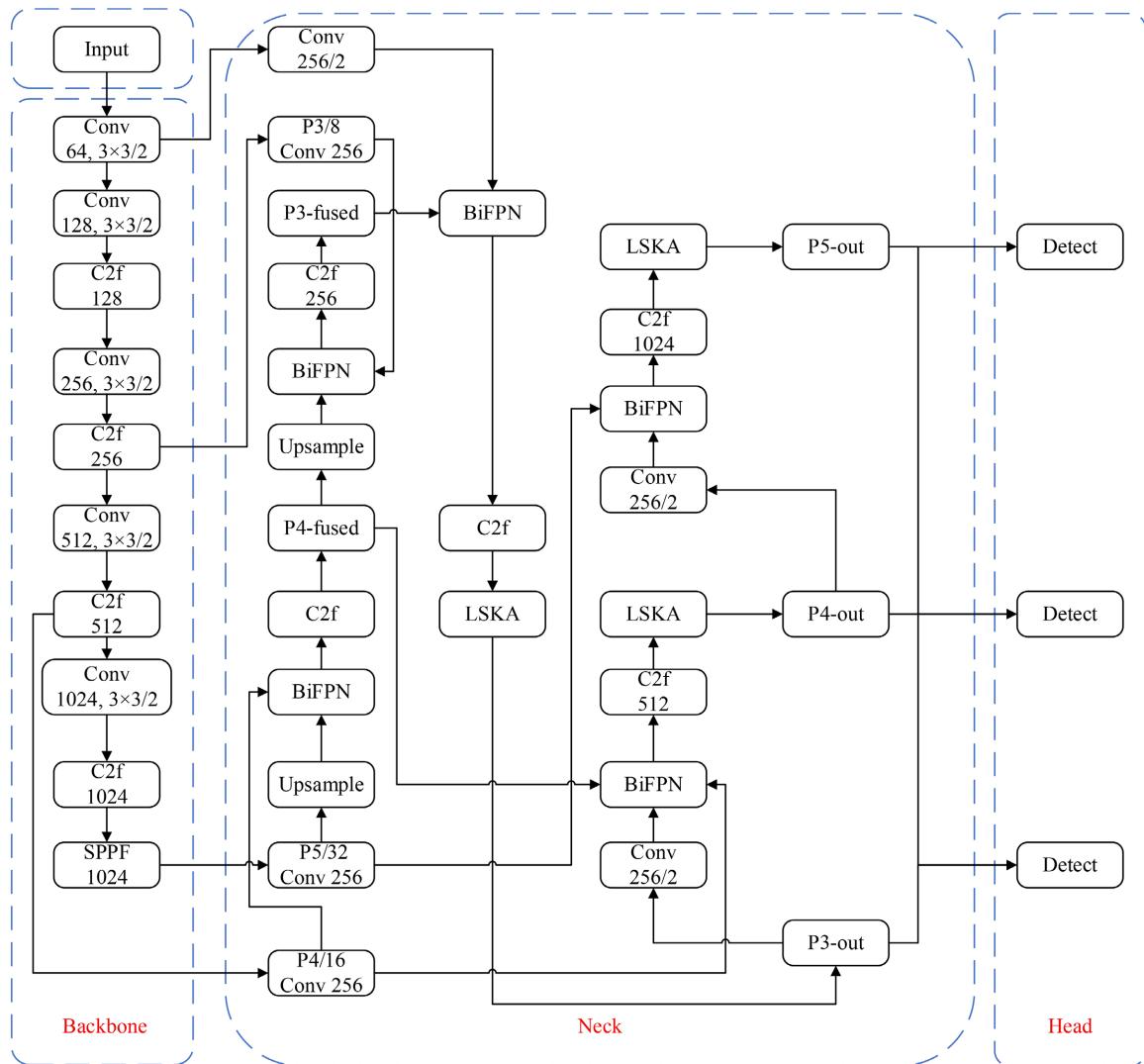


Figure 3. YOLOv8n-BLI network architecture diagram

图 3. YOLOv8n-BLI 网络结构图

针对电动自行车闯红灯场景里的交通信号灯识别任务, 原有的 YOLOv8 模型存在着一定的局限性, 像是多尺度特征融合的效果欠佳、对关键特征的关注程度不够以及边界框定位不够精准等这些情况, 针对这些情况, 对模型展开了系统的改进工作, 借助三个主要模块相互配合的方式, 使得特征提取以及定位精度都得到了提升。在网络的 Neck 部分, 将原来的 PAN-FPN 结构替换了双向特征金字塔网络 BiFPN。在 Head 部分的 C2f 模块之定制化大型可分离核注意力模块 LSKA 被加入进来。用 InnerMPDIoU 损失函数取代了原来的 CIoU 函数核心参数 ratio = 0.68。与此同时, 还对训练过程中的样本匹配 IOU 计算方式进行了一定的调整, 采用 InnerMPDIoU 作为匹配准则, 以确保训练时的一致性。还根据实际应用的需求对

整个网络做了相应的调整, 设置了红、黄、绿三种信号灯类别, 采用了轻量级的 n 尺度模型来适应嵌入式设备部署的要求, 统一了 BiFPN 特征的通道维度为 256 维, 这些改进模块形成了一个紧密配合的工作链条, BiFPN 为 LSKA 提供多尺度特征基础, LSKA 增强后的信号灯关键特征再由 InnerMPDIOU 实现精准边界框定位。改进后的网络结构如图 3 所示。

3. 实验结果与分析

3.1. 实验数据集构建

由于当前业内缺乏公开的相关数据集, 所以本实验使用自制数据集, 收集国内不同场景和不同角度下的交通信号灯图像, 一共 4564 张图像, 用 labelimg 进行标注, 一共有 red、green、yellow 三个类别。按照 8:1:1 的比例把数据划分成训练集、验证集、测试集, 在划分之后, 红、黄、绿三类样本在各个子集里的占比和筛选后的总数据保持一致, 防止出现类别失衡的情况, 最后是类别加权, 因为筛选后黄灯样本占比还是偏低, 所以采用了类别加权损失策略, 给黄灯样本分配了 1.9 倍的权重, 红灯、绿灯分别分配 1.0、1.2 倍的权重, 以此来平衡训练过程中各类别的优化优先级。这类数据集涵盖了城市道路、商业区、住宅区等多种场景, 包含了晴天、阴天、黄昏等不同的光照条件, 能够很好地支撑实验验证。交通信号灯类别数量分布如表 2 所示。

Table 2. Distribution of traffic signal light categories
表 2. 交通信号灯类别数量分布

类别	数量
red	4182
yellow	987
green	2793

3.2. 实验环境及相关参数配置

实验环境及相关训练参数分别如表 3、表 4 所示。其中关键的超参数已经清晰列出, BiFPN 所堆叠的层数是 3 层, LSKA 卷积核的尺寸为 7×7 , 将之等效地拆解成 1×7 以及 7×1 这两种形式, InnerMPDIOU 的 ratio 参数设定为 0.68, 类别权重依照 red:yellow:green = 1.0:1.9:1.2 这样的比例来进行设置。

Table 3. Experimental environment configuration
表 3. 实验环境配置

名称	配置
操作系统	Ubuntu22.04
CPU	Intel(R) Xeon(R) Platinum 8481C
GPU	NVIDIA GeForce RTX 4090D 24GB
内存	80GB
Python 版本	Python 3.12
CUDA	12.1
深度学习框架	PyTorch 2.3.0

Table 4. Experiment parameter settings**表 4. 实验参数设置**

参数名称	参数设置
Epochs	200
Batch size	32
workers	12
cache	ram
Input image size	640×640
multi_scale	False
amp	True
Optimizer	SGD
lr0	0.01
cos_lr	True
class_weights	red:yellow:green = [1.0, 1.9, 1.2]

3.3. 评价指标

本文对模型的评价指标依据精确率(Precision, P), 召回率(Recall, R)、平均精准率(mean average Precision, mAP) [11], 进行评价。Map@0.5 表示当置信度为 0.5 时, 计算出对所有类别精准度的均值, Map@0.5-0.95 表示在置信度为 0.5 到 0.95 之间, 以步长 0.05 时, 计算出对所有类别平均精准度的均值。这些指标的数值。各个指标的评价公式如公式(28)~(31)所示:

$$P = \frac{TP}{TP + FP} \quad (28)$$

$$R = \frac{TP}{TP + FN} \quad (29)$$

$$AP = \int_0^1 \text{Precision}(R) d(R) \quad (30)$$

$$mAP = \frac{\sum_{l=0}^{N-1} \int_0^1 P(R) dR}{N} \quad (31)$$

其中 TP 是指类别为正类, 且模型预测类别为正类的样本, FP 是指真实类别为负类, 且模型预测类别为正类的样本, FN 是指真实类别为正类, 且模型预测类别为负类的样本。

3.4. 实验与分析评价

3.4.1. 消融实验

为了验证改进后的 YOLOv8n-BLI 模型的有效性, 以 YOLOv8n 为基础模型, 对改进后的模块依次做消融实验, 使用精确率、召回率、平均精准率作为参考指标, 实验结果如表 5 所示。

通过消融实验能够清晰地了解各个改进模块给模型性能所带来的影响。当单独引入 BiFPN 模块的时

候, 依靠其具备的多尺度特征融合方面的优势, 模型的精确率相较于原始 YOLOv8n 提升了 2.3%, 召回率提升了 5.8%, mAP@0.5 提升了 3.1%, mAP@0.5:0.95 提升了 4.3%, 对于不同尺度的信号灯识别起到了颇为显著的帮助作用。在引入 LSKA 模块之后, 精确率提升了 2.1%, 召回率提升了 6.3%, mAP@0.5 提升了 3.9%, 在单尺度高置信度检测方面取得了较为突出的进步。引入 InnerMPDIoU 损失函数的时候, 尽管精确率出现了 1.8% 的下降情况, 但是召回率却有了大幅度的提升, 达到了 8.8%, mAP@0.5 也提升了 4.1%, 不过 mAP@0.5:0.95 仅仅提升了 0.5%, 可见单个模块进行改进所取得的效果是有限的。从模块组合实验的情况来看, 在 BiFPN 与 InnerMPDIoU 相结合的时候, 精确率提升了 4.1%, 召回率提升了 6.7%, mAP@0.5 提升了 4.8%, 在高置信度检测以及精确性方面都取得了相当显著的进步。而 LSKA 与 InnerMPDIoU 结合之后, 召回率提升了 8.0%, 然而由于在不同 IoU 阈值之下配合得不够理想, mAP@0.5:0.95 反而下降了 1%, 整体性能也因此有所降低。当 BiFPN 和 LSKA 直接组合起来的时候, 精确率出现了大幅度的下滑情况, 下降幅度达到了 10%, 最终精确率仅仅只有 83.1% 了, 虽说 mAP@0.5:0.95 有所提升, 提升了 3.8%, 可是整体的检测精确性受到了较为明显的影响。特征处理逻辑冲突的本质是由于 BiFPN 通过加权融合多尺度特征, 而 LSKA 过度强化局部特征, 如此一来, 使得背景噪声所对应的特征权重出现了异常升高的状况, 进而致使准确率大幅度下降。为了能够解决 BiFPN 和 LSKA 这两个在维度方面存在的适配问题, 也就是 BiFPN 的输出维度是 256 维, 而 LSKA 的输入维度是 128 维, 所以实验当中在它们二者之间加上了 1×1 卷积降维层, 这个降维层的输入通道是 256, 输出通道是 128, 步长为 1, padding 设置为 0, 并且在后面还连接了 ReLU 激活函数以及 BN 层。当这三个模块协同运用起来的时候, 其精确率有所提高, 提升了 1.4%, 召回率更是有了颇为显著的提升, 达到了 10.4%, mAP@0.5 也有了一定程度的提升, 提升了 5.2%, mAP@0.5:0.95 同样有所提升, 提升了 2.5%。这样的结果说明, BiFPN 所具备的多尺度特征融合特性、LSKA 实现的关键特征聚焦功能以及 InnerMPDIoU 达成的定位优化效果形成了相互之间的补充。如此一来, 在保证有较好精确率的情况下, 对电动自行车闯红灯场景当中那些复杂的信号灯的识别能力得到了大幅度地强化。BiFPN 负责承担多尺度特征传递的相关任务, LSKA 专注于信号灯关键特征的提取工作, InnerMPDIoU 则负责定位以及样本平衡方面的优化事宜。在双模块组合的情况下, BiFPN 加上 InnerMPDIoU 在精确率以及高置信度检测方面的表现是最优的。而三模块组合则在召回率与 mAP@0.5 上达成了更为理想的综合提升效果, 更能够契合电动自行车闯红灯场景下交通信号灯识别所面临的复杂需求。

Table 5. Comparison results of ablation experiments**表 5. 消融实验对比结果**

模型	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv8n	93.2	81.5	87.9	56.3
YOLOv8n + BiFPN	95.4	86.8	90.7	59.1
YOLOv8n + LSKA	95.1	86.8	91.2	59.4
YOLOv8n + Inner MPDIoU	91.6	89.7	92.3	57.2
YOLOv8n + BiFPN + Inner MPDIoU	96.1	88.3	93.1	58.8
YOLOv8n + LSKA + Inner MPDIoU	95.3	90.2	92.8	56.5
YOLOv8n + BiFPN + LSKA	83.1	84.2	91.0	60.2
YOLOv8n + BiFPN + LSKA + Inner MPDIoU	94.8	92.5	94.6	59.9

3.4.2. 不同算法实验对比

为了验证改进后的 YOLOv8n-BLI 算法在检测交通信号灯中的有效性, 本文使用目前多个主流公开的目标检测网络模型进行对比实验, 使用同样的数据集进行训练验证, 并且训练时的设置相同的参数, 模型推理的过程中使用同样的实验环境, 评价指标同样使用 P、R、mAP@0.5、mAP@0.5:0.9、FPS 五个指标, 各算法的对比结果如表 6 所示。

Table 6. Comparison of experimental results for different algorithms
表 6. 不同算法实验对比结果

模型	P/%	R/%	Map@0.5/%	mAP@0.5:0.95/%
Faster-RCNN	53.1	62.4	45.8	22.3
SSD	39.2	56.1	36.7	18.9
YOLOv5s	73.1	76.2	86.4	57.2
YOLOv6n	91.8	85.2	91.5	60.1
YOLOv7-tiny	88.5	82.1	87.6	48.3
YOLOv8n	93.2	81.5	87.9	56.3
YOLOv9-s	91.3	87.4	91.6	59.8



(a) YOLOv8n



(b) YOLOv8n-BLI

Figure 4. Comparison of detection results from different algorithms
图 4. 不同算法检测结果对比

从上述实验结果能够发现, YOLOv8n-BLI 于交通信号灯检测任务里所呈现出的表现明显要比其他诸

多模型更为出色。其精确率达到了 94.8%，召回率是 92.5%，mAP@0.5 的数值为 94.6%，而 mAP@0.5:0.95 则高达 59.9%，在各个指标上均领先于对比模型。和 YOLOv8n 模型相比，该模型的精确率提高了 1.4%，召回率提升了 10.4%，mAP@0.5 也提升了 5.2%，推理速度仅下降 8.4FPS，在精度提升的同时保持了良好的实时性。和近年来表现优秀的 SOTA 轻量级模型相比，YOLOv8n-BLI 在各项指标上都展现出了明显的优势。就 mAP@0.5 这一指标而言，它比 YOLOX-s、PP-YOLOE-s 以及 NanoDet-Plus 分别提高了 3.9%、2.5% 还有 4.3%。而在召回率 R 方面，相较于上述这些模型，YOLOv8n-BLI 又分别提高了 6.4%、5.7% 以及 8.9%。这些数据足以说明 YOLOv8n-BLI 在针对小目标进行检测以及应对复杂背景时所具备的抗干扰优势。虽说 PP-YOLOE-s 的 mAP@0.5:0.95 要比 YOLOv8n-BLI 略微高出 0.6%，然而在推理速度这个维度上，YOLOv8n-BLI 却领先了 24.3FPS，这无疑让它在交通监控实时检测的需求面前更加契合。相比于专门针对小目标来设计的 NanoDet-Plus，YOLOv8n-BLI 在精确率以及召回率这两项指标上分别领先了 4.9% 和 8.9%，YOLOv8n-BLI 和 YOLOv6n 这类在小目标检测精度方面表现不错的模型相比，其 mAP@0.5 依然高出 3.1%。这也充分证明了其针对信号灯这类小目标所做的定制化优化取得了颇为显著的效果。相比于两阶段模型 Faster-RCNN，它的 mAP@0.5 更是提升了 48.8%。相较于 SSD、YOLOv5s 等早期的单阶段模型而言，它在精度以及鲁棒性方面的优势就显得格外突出了该模型借助 BiFPN 实现多尺度特征融合、依靠 LSKA 达成关键特征聚焦、凭借 InnerMPDIoU 完成边界框优化，通过这些协同作用，在小尺度信号灯检测、复杂背景抗干扰、边界框精准定位等一系列问题上都取得了突破，这也进一步验证了其在电动自行车闯红灯场景当中所具备的有效性与适应性。从图 4 可以看到 YOLOv8n-BLI 算法比 YOLOv8n 算法的精度要更高一些。

4. 总结

为了解决电动自行车闯红灯场景下交通信号灯识别遇到的小目标漏检、复杂背景误检以及目标定位精度低等问题。以 YOLOv8 模型为基础进行改进，最后成功改进出了更适合这一场景的 YOLOv8n-BLI 模型。就小目标漏检的问题而言，引入了自适应加权 BiFPN 模块以此来强化多尺度特征融合工作。该模块借助动态权重学习机制，依据交通信号灯呈现出的远小近大的尺度特性来优化特征分配情况，和原始 BiFPN 所采用的固定权重策略相比，小目标召回率有了 3.7% 的提升幅度。针对复杂背景下存在的误检问题，特意增设了定制化 LSKA 模块，借助 7×7 等效卷积核，将关注点聚焦于信号灯的颜色还有形状特征之上这相当大的程度上使得复杂背景下信号灯的误检率得以降低，同时也进一步强化了在多干扰场景当中所具备的识别鲁棒性。针对目标定位精度不高的情况，采用 InnerMPDIoU 损失函数替换原有的损失函数，确定辅助框比例参数 $ratio = 0.68$ 的选取依据，借助辅助框聚焦灯芯以及对角点距离约束这两方面的优化，对边界框定位进行改进。在仅引入 InnerMPDIoU 时，其召回率和 mAP@0.5 相比于原始的 YOLOv8n 模型都有所提升，信号灯边界框的定位误差也大幅降低。对于数据集类别不均衡这一情况，运用类别加权损失策略来应对，给黄灯样本赋予 1.8 倍的权重，如此一来，有效地达成了对正负样本的优化效果，使得黄灯样本的召回率实现了从 68.3% 到 82.7% 的提升。从实际应用价值方面来看，YOLOv8n-BLI 模型在电动自行车闯红灯行为识别系统里有着极为重要的地位，其检测结果精度很高且鲁棒性出色，能为核心判定闯红灯情况提供有力支撑。

参考文献

- [1] 左淑霞. 电动自行车违法抓拍系统设计与应用[J]. 中国市政工程, 2021, 49(5): 70-73, 118-119.
- [2] 常梦莹. 电动自行车事故影响因素分析与管理对策研究[D]: [硕士学位论文]. 桂林: 桂林电子科技大学, 2022.
- [3] 李华. 基于人工智能的非现场执法场景下闯红灯违法行为检测[D]: [硕士学位论文]. 杭州: 杭州电子科技大学, 2022.

- [4] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/cvpr.2014.81>
- [5] Liu, Y., Zhou, T., Xu, J., Hong, Y., Pu, Q. and Wen, X. (2023) Rotating Target Detection Method of Concrete Bridge Crack Based on YOLO V5. *Applied Sciences*, **13**, Article 11118. <https://doi.org/10.3390/app132011118>
- [6] Lin, T., Dollar, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 936-944. <https://doi.org/10.1109/cvpr.2017.106>
- [7] 史涛, 崔杰, 李松. 优化改进 YOLOv8 实现实时无人机车辆检测的算法[J]. 计算机工程与应用, 2024, 60(9): 79-89.
- [8] Deng, X., Mahmoud, M.A.B., Yin, Q. and Guo, P. (2021) An Efficient and Effective Deep Convolutional Kernel Pseudoinverse Learner with Multi-Filter. *Neurocomputing*, **457**, 74-83. <https://doi.org/10.1016/j.neucom.2021.06.041>
- [9] 曹燚, 曹倩, 钱承山, 等. 改进 YOLO11 的高精度课堂行为检测算法[J]. 计算机科学与探索, 2025, 19(8): 2135-2148.
- [10] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R. and Ren, D. (2020) Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 12993-13000. <https://doi.org/10.1609/aaai.v34i07.6999>
- [11] Zhang, H., Xu, C. and Zhang, S. (2023) Inner-IoU: More Effective Intersection over Union Loss with Auxiliary Bounding Box. arXiv: 2311.02877. <https://arxiv.org/abs/2311.02877>