

BSD-YOLO: 基于动态稀疏注意力与自适应检测头的小目标车辆检测方法

杨彪, 王学, 官铮, 龙萍

(云南大学信息学院, 云南 昆明 650504)

摘要: 在智能交通监控系统中, 复杂场景下的小目标车辆检测面临特征分辨率低、遮挡干扰严重、模型计算冗余及边界框回归精度不足等挑战。为兼顾检测精度与边缘设备部署效率, 提出一种基于动态稀疏注意力与轻量化双分支结构的改进 YOLOv8 检测框架。首先设计双向路由稀疏注意力机制(ReBiAttention), 通过双层动态路由筛选关键特征, 增强对小目标浅层特征的保留能力; 随后结合 GSConv 与 VoV-GSCSP 模块, 在减小计算量的同时动态调整多尺度特征权重; 并在检测头部分引入改进型 DynamicHead 结构, 实现多任务自适应优化; 最后改进 ShapeIoU 损失函数, 引入形状与尺度感知机制, 提升定位精度。在 UA-DETRAC 数据集上的实验表明, 改进模型较基线 YOLOv8n 的 Precision, Recall 与 mAP@0.5 分别提升 8.739%, 1.685% 和 7.225%, 参数量减少 4.3%。该方法为复杂交通场景下的小目标车辆精准检测提供了高效解决方案。

关键词: YOLOv8; 注意力机制; 轻量化; 深度学习; 小目标检测

中图分类号: U 495; TP 391.41; TP 183 DOI: 10.11996/JGj.2095-302X.2026010099

文献标识码: A 文章编号: 2095-302X(2026)01-0099-12

BSD-YOLO: a small target vehicle detection method based on dynamic sparse attention and adaptive detection head

YANG Biao, WANG Xue, GUAN Zheng, LONG Ping

(School of Information Science and Engineering, Yunnan University, Kunming Yunnan 650504, China)

Abstract: In intelligent traffic monitoring systems, small target vehicle detection in complex scenes faces challenges such as low feature resolution, severe occlusion interference, computational redundancy, and insufficient bounding-box regression accuracy. To balance detection accuracy with deployment efficiency on edge devices, an improved YOLOv8 framework based on dynamic sparse attention and a lightweight dual-branch structure was proposed. The method first introduced a bidirectional routing sparse attention mechanism (ReBiAttention) that enhanced the retention of shallow features for small targets by dynamically filtering key features through a two-level routing strategy. Subsequently, GSConv and VoV-GSCSP modules were integrated to reduce computational cost while dynamically adjusting multi-scale feature weights. An improved DynamicHead was applied for multi-task adaptive optimization, and a modified ShapeIoU loss function with shape- and scale-aware weighting was employed to improve localization accuracy. Experiments on the UA-DETRAC dataset showed that, relative to baseline YOLOv8n, Precision, Recall, and mAP@0.5 increased by 8.739%, 1.685%, and 7.225%, respectively, while the parameter count decreased by 4.3%. This method provided an efficient solution for accurate detection of small-target vehicles in complex traffic scenarios.

Keywords: YOLOv8; sparse attention; lightweight; deep learning; small target detection

收稿日期: 2025-06-16; 定稿日期: 2025-08-18; 通信作者: 官铮, E-mail: guanzheng@ynu.edu.cn

Received: 16 June, 2025; Finalized: 18 August, 2025; Corresponding author: GUAN Zheng, E-mail: guanzheng@ynu.edu.cn

基金项目: 国家自然科学基金(61761045); 云南专家工作站项目(202305AF150045)

Foundation items: National Natural Science Foundation of China (61761045); Yunnan Expert Workstation (202305AF150045)

车辆作为道路交通系统的核心参与者^[1],其精准检测对智能交通管理、自动驾驶安全及道路监控效能提升具有关键意义。早期基于传统机器学习的方法,如支持向量机(Support Vector Machine, SVM)^[2]和 Haar 特征分类器^[3]依赖人工设计特征与浅层分类器构建检测模型,虽在受限场景下具备实时处理优势,但受限于特征表达能力不足、环境鲁棒性弱等技术瓶颈制约,故对于小目标检测,手工设计的特征难以捕捉细微纹理差异,且分类器易受光照变化、遮挡干扰影响,导致复杂场景下的泛化性能显著下降。在车流密集区域与复杂道路场景中,小目标车辆的漏检与误检问题已成为阻碍相关技术实际落地应用的核心瓶颈^[4],亟待具有兼顾检测精度与推理效率的创新性解决方案。深度学习技术的非线性拟合能力为该技术的发展提供了更具代表性的解决方案^[5]。现有基于深度学习的小目标检测方法主要聚焦于模型的精度与速度^[6]:

1) 在特征提取不充分和信息丢失问题方面,已有研究^[7]将视觉 Transformer (Vision Transformer, ViT),即多头转置注意力+通道注意力与 CNN (Convolutional Neural Network) 并联,用于局部-全局特征提取;并通过分组渐进式残差结构进行端到端无监督特征融合,大幅提升了红外-可见光融合图像的对比度、清晰度及定量性能指标。

2) 在检测精度效率提升方面,双层路由策略^[8]通过内容感知的计算资源动态分配,实现了关键区域聚焦与无效计算过滤,但其采用的 BiFormer (Bi-Level Transformer) 架构仍存在结构性局限^[9]:如预设尺寸的区域划分($S \times S$)难以适配多尺度目标的形态,导致小目标丢失关键信息; Top- k 在进行候选区域筛选时,需要预先计算完整亲和度矩阵,故内存与计算开销随特征图分辨率呈二次增长,难以满足实时处理的需求; Value 端局部特征增强受限于深度可分离卷积的固定范围,跨区域细节挖掘能力不足;且固定的 Top- k 筛选策略仅依据相关性分数选取前 k 个区域,忽略了那些虽然得分略低、但与目标在语义上联系紧密的区域,容易造成关键信息遗漏。

3) 在传统轻量化方面^[10],深度可分离卷积^[11]等技术虽能降低计算量,但通道独立性导致特征表达能力退化,尤其对小目标检测精度影响显著。与此同时,特征金字塔网络 (Feature Pyramid Network, FPN)^[12]等固定层级连接策略受困于通道冗余与空

间信息衰减,高层语义特征补偿能力不足,难以平衡多尺度细节与全局语义的融合效率。为此,LI 等^[13]提出基于动态特征路由的轻量化金字塔重构模块 Slim-Neck,其核心突破在于构建可微分稀疏连接机制:通过门控分离卷积 (Gated Spatial Convolution, GSConv) 将标准卷积解耦为空间感知与通道调制分支;金字塔特征动态路由协议通过可学习门控函数对跨层级特征流进行非线性加权,提升关键特征传递效率并抑制冗余干扰;多粒度特征蒸馏机制在融合阶段实施通道级重要性排序时,仅保留前 30% 高响应特征进行跨尺度聚合,使计算复杂度 (Floating Point Operations, FLOPs) 降至传统 FPN 的 40%,同时维持高效特征表达能力。

除此之外,深度学习技术的突破为小目标检测提供了全新解决方案。以 SSD^[14] (Single Shot MultiBox Detector) 和 RetinaNet^[15] 为代表的单阶段检测器,通过端到端特征学习与高效回归策略,在实时性方面展现出显著优势。但其改进重点集中于通用尺度目标检测,针对交通场景中普遍存在的远距小车辆 (像素面积 $< 32^2$) 及密集遮挡目标,仍存在显著的漏检问题。

两阶段检测器 (如 Faster R-CNN^[16] 和 Cascade R-CNN^[17]) 通过区域建议机制提升了检测精度,但复杂的级联结构导致计算负载过高,难以满足交通监控场景的实时性要求。实测数据表明,在 1 080 p 分辨率视频流处理中, Faster R-CNN 的推理速度仅为 12 FPS,且小目标召回率低于 50%,制约了其工程化应用。

虽然这些方法都取得了显著的成效,但当前小目标车辆检测技术仍面临着以下挑战:

1) 特征表达局限。远距离车辆在深层特征提取过程中面临严重的细节信息损耗,常规卷积操作难以有效捕捉微小目标的纹理特征;

2) 背景混淆干扰。交通标志、护栏等背景物体与车辆局部特征存在相似性,传统注意力机制难以精准聚焦目标区域,易受相似背景特征的干扰而产生误激活;

3) 遮挡敏感性问题。密集车流场景下目标可见区域显著减少,单一检测头结构难以适应部分遮挡目标的几何形态多样性;

4) 效率性能不均衡。为提升小目标检测精度,模型网络层次更大,导致体积和计算量剧增;而轻量化虽可加速识别,却易丧失微小目标的细节,导致精度下降。因此如何在有限算力下兼顾实时与高

精度, 是亟待解决的难题。

针对上述问题, 本文提出了一种基于动态稀疏注意力与轻量化双分支结构的改进 YOLOv8 (You Only Look Once Version 8)检测框架: BSD-YOLO (Bi-directional Sparse Dynamic YOLO), 其中“B”表示双向(Bi-directional), “S”表示稀疏(Sparse), “D”表示动态(Dynamic)。本文通过多层次架构优化设计, 实现检测精度与推理效率的协同提升。大量实验证实, 相比传统全局注意力在 UA-DETRAC 数据集上使 mAPs(Mean Average Precision Scores)提升 7.225%。具体来说, 本文的主要贡献如下:

1) 提出了一种双层路由注意力机制, 将特征图划分为若干区域, 基于区域间相似度进行筛选, 执行特征点注意力并加入局部卷积增强, 实现稀疏计算与全局聚合, 因此大幅削减计算与内存开销, 增强对小目标浅层特征的保留能力;

2) 有效的模块, 例如: GSConv, VoVGSCSP 和 Dynamic head; 其中 GSConv 和 VoVGSCSP 利用多级轻量化瓶颈强化跨层特征融合, 在 YOLOv8 特征金字塔颈部实现效率与精度的更优平衡; Dynamic head 显著降低推理延迟与内存占用, 有效促进小目标检测;

3) 提出一种优化的 ShapeIoU 方法, 通过引入形状引导的自适应权重分配和多尺度边缘感知损失, 增强对小目标特征的感知灵敏度, 从而提升定位精度。

1 相关工作

YOLO 系列^[18-25]将目标检测视作一个端到端的回归问题, 通过单次前向传播即可同时预测边界框及类别概率, 实现了真正意义上的“一次性”检测。且将输入图像划分为 $S \times S$ 的网格, 每个网格负责预测位于其中心的物体的 B 个边界框(坐标与置信度)和 C 个类条件概率, 大幅提升了检测速度与简便性。杨锦辉等^[26]提出了基于瓶颈层优化的改进方案, 通过重构路径聚合网络的瓶颈模块, 集成感受野增强模块(RFB-bottleneck)提升特征表达能力, 并采用 Ghost 卷积替代传统卷积层以压缩模型参数量。该方案虽有效缩减了模型体积, 但引入的结构冗余导致帧率性能出现明显衰减, 制约了实时检测场景的应用潜力。YU 等^[27]提出面向自动驾驶场景的 YOLO-MPAM 检测框架, 其创新包括: 通过多头自注意力机制构建位置感知网络, 动态筛选关键区域特征, 实现目标敏感区域的高效聚焦; 设

计层次化特征交互路径, 融合浅层高分辨率细节与深层语义信息, 显著提升微小目标的定位精度。在针对小目标定位与分类任务的耦合性问题, DAI 等^[28]提出基于注意力机制的统一化动态检测头架构, 通过协同整合多尺度特征关联、局部区域聚焦及任务权重分配 3 层次自注意力机制, 在不增加额外计算开销的前提下显著提升了检测头的特征表征能力。基于 COCO 基准数据集的实验结果表明, 采用 ResNeXt-101-DCN 主干网络时, 该方法以 54.0 AP 的性能指标刷新了当时的最优检测精度。

YOLOv8 作为 YOLO 系列的重要改进版本, 在目标检测、姿态估计、图像分类及语义分割等任务中展现出优异性能。相较于 YOLOv5, 其主要改进包括: ①网络结构优化; ②解耦头设计改进; ③加入 Anchor-free 策略。上述改进使 YOLOv8 在保持高效推理速度的同时, 实现了复杂场景下检测性能的全面提升。

尽管 YOLO 系列最新迭代至 YOLOv12^[29], 但本研究仍选择 YOLOv8n 作为算法改进的基线模型, 主要基于以下几点考虑:

1) YOLOv12 聚焦于将注意力(Area-attention)、大核卷积和 R-ELAN 等集成来提升整体特征表达能力, 而动态稀疏双层路由注意力机制(Refined Bi-level Dynamic Sparse Routing Attention Mechanism, ReBiAttention)聚焦于邻接 Top- k 区域的稀疏建图处理, 有效过滤低质量干扰区域。

2) Dynamic head 则是为 YOLOv8n 设计的可动态调节任务分支的检测头结构, 具有更强的小目标识别能力, 而不是引入复杂的注意力计算。

3) Slim-Neck 模块通过共享卷积降低额外计算, 实现轻量化性能提升, 无需依赖大型算子或硬件加速支持。

综上所述, 本文在轻量模型 YOLOv8n 上注入符合实际应用需求的小目标特征筛选能力, 保证实时性同时提高检测精度。

2 BSD-YOLO 模型网络介绍

本研究以 YOLOv8n 作为基线模型, 重点对骨干网络、Head 层以及损失函数的计算方法进行优化, 以提升模型在车辆检测方面的性能。图 1 为改进后的 YOLOv8 网络模型, 图 1 中右上角以及 Head 部分为本文针对 YOLOv8 所做的网络改进(BSD-YOLO 模型网络)。

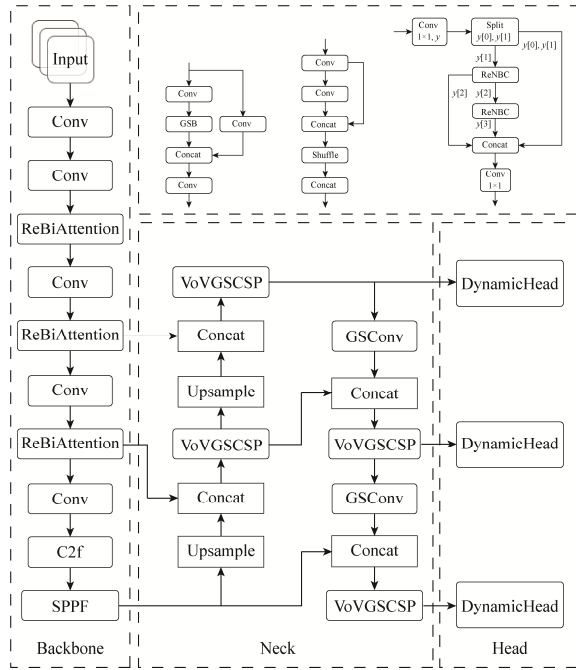


图 1 BSD-YOLO 模型结构

Fig. 1 BSD-YOLO model structure

2.1 动态稀疏双层路由注意力机制

在视觉 Transformer 架构中,注意力机制作为核心组件,能够有效捕捉长距离依赖关系并融合全局信息,从而显著提升模型表征能力。然而,若要计算所有空间位置间的成对交互,将会导致较高的计算复杂度与内存占用。为应对这个挑战,诸多研究者尝试引入预设稀疏结构,如将注意力计算限定于局部窗口,或采用轴向、扩张等方式划定计算范围,以此降低资源消耗。

本文基于对双层路由策略的深入剖析及现有方案的问题洞察提出 ReBiAttention。该机制在维持全局信息聚合能力的同时,通过自适应筛选关键区域,大幅削减计算与内存开销。与固定窗口或轴向稀疏等内容无关的预设模式不同,其能够依据输入特征的实际分布,动态构建区域级亲和图,并只会保留与各查询最相关的 Top- k 区域,从而在保障模型性能的前提下,实现长距离依赖的高效捕捉。

如图 2 所示, ReBiAttention 模块主要包含以下流程:

1) 运用 1×1 卷积对输入特征进行通道变换与压缩,等价于 BiFormer 中的输入投影操作,将投影后的特征图划分为 $S \times S$ 个不重叠区域,每个区域包含 HW/S^2 个特征向量,以实施粗粒度筛选。

2) 针对每个候选区域,计算其与其他区域之间的特征相似度关系,并以此构建区域间的有向

关联图,用于后续的路由选择。对每个查询区域,仅保留邻接度最高的 k 条边(即 Top- k 相关区域),滤除大部分无关区域,有效地将计算复杂度降至 $O\left(\frac{4}{3}(HW)^3\right)$ 。继而对留存的候选区域集合,执行常规点乘注意力,实现跨区域精细信息融合。

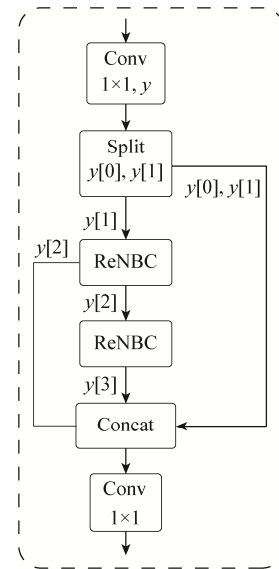


图 2 ReBiAttention 结构图

Fig. 2 ReBiAttention structure diagram

为了研究 k 对小目标检测性能的影响,本文在 Top- k routing 中分别令 $k \in \{1, 2, 3, 4, 5, 6, 7, 8\}$,并在 UA-DETRAC 验证集上对比 mAP@0.50 和计算量 GFLOPs。并为了避免在图像尺寸较小或分块数量过少时造成 Top- k 越界,本实验对 k 进行限制,确保其不超过 p^2 (p^2 表示图像被划分为若干 $p \times p$ 个 Non-overlapping 局部窗口后,超过总窗口数量会使权重稀释和无效连接,反而使注意力泛化)。结果表明,当 $k \geq 4$ 时模型取得最佳精度-速度平衡,同时计算开销也处于可接受范围,且当 $k=4$ 时计算量最小,因此本文最终采用 $k=4$ 作为默认设置,见表 1。

表 1 k 值对 ReBiAttention 的敏感性分析Table 1 Sensitivity analysis of k value to ReBiAttention

Top- k	Precision	Recall	mAP@0.50	mAP@0.50:0.95	GFLOPs
1	0.563 24	0.619 81	0.599 37	0.416 07	11.1
2	0.623 38	0.593 59	0.602 03	0.445 74	<u>11.2</u>
3	<u>0.669 23</u>	0.571 61	0.623 59	0.448 92	11.3
4	0.673 39	0.613 42	<u>0.642 58</u>	0.444 12	11.3
5	0.654 76	0.624 12	0.631 84	0.470 34	11.4
6	0.632 35	0.625 91	0.636 18	0.457 68	11.5
7	0.615 29	<u>0.628 39</u>	0.633 24	0.457 65	11.5
8	0.615 95	0.656 82	0.648 06	<u>0.468 26</u>	11.6

注:加粗数据表示最优值;下划线数据表示次优值。

3) 在 Value 张量上引入深度卷积增强项(Local context enhancement), 以捕捉邻域空间细节。最终将注意力输出重组为原始空间形状, 经后续层归一化与多层感知机(Multilayer Perceptron, MLP)完成特征更新, 形成下一层 Transformer 的输入。

ReBiAttention 的双向路由机制借助 Top- k token 筛选策略, 仅保留与目标相关的特征区域参与计算。这种稀疏设计尤其契合交通场景中小目标分散分布的特性。通过层次化路由路径, 该机制实现浅层细节特征(如车辆边缘、车轮纹理)与深层语义信息(如车辆类别、方向)的协同优化。在 ReNBC 中采用深度可分离卷积(Depthwise Separable Convolution, DSC)替代标准卷积, 结合通道维度动态压缩策略, 进一步提升计算效率。此外, 模块化设计使其可无缝集成至 YOLOv8 框架, 无需调整损失函数或训练策略, 显著提升 YOLOv8 在小目标车辆检测中的精度与效率, 同时缓解复杂交通场景下的背景干扰与遮挡敏感问题。

2.2 轻量化特征金字塔重构

Slim-Neck(图 3)通过门控函数动态调整权重, 但无法主动过滤与小目标无关的空间区域, 导致冗余计算开销; 且多粒度通道蒸馏需逐层排序裁剪, 增加了运行负担。此外, 若只采用 DWConv 会因通道独立运算破坏跨通道耦合, 削弱细粒度特征表达能力。

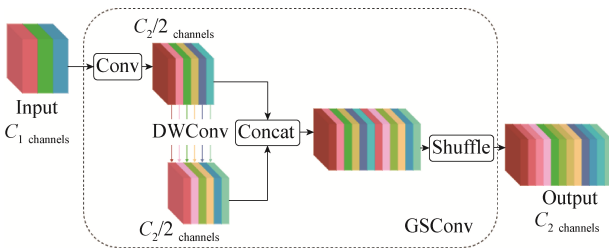


图 3 GSConv 结构示意图

Fig. 3 Schematic diagram of the GSConv structure

针对上述问题, 对原始动态特征路由的轻量化金字塔重构模块进行改进, 将图 4 中的 3 个模块分别引入 YOLOv8 的 Neck 层(图 1), 有效解决部分问题:

- 1) GSConv 负责通道压缩与轻量化卷积, 通过“半精度”标准卷积与深度可分离卷积拼接及通道混洗操作, 缓解 DWConv 的表达缺陷;
- 2) VoVGSCSP 在跨层融合中引入多级 GSBottleneck 串联与恒通道捷径, 借助 1×1 卷积升降维及残差连接, 实现高效特征融合。

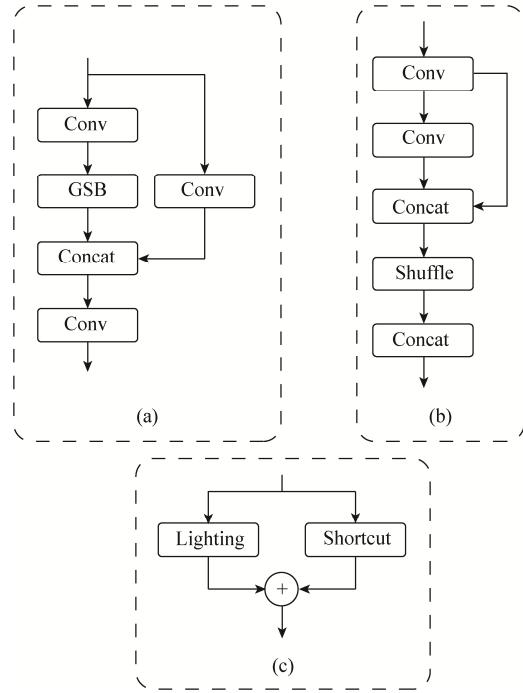


图 4 模块结构图

Fig. 4 Module structure diagram ((a) VoVGSCSP; (b) GSConv; (c) ReBiAttention)

综上, 新构建的 GSConv 与 VoVGSCSP 模块, 有机融合标准卷积、深度可分离卷积与残差聚合, 通过通道拼接与 Shuffle 消除 DWConv 的表达瓶颈, 利用多级轻量化瓶颈强化跨层特征融合, 在 YOLOv8 特征金字塔颈部实现效率与精度的更优平衡。

2.3 多任务自适应检测头(Dynamic head)

Dynamic head 借助尺度感知(Scale-aware)、空间感知(Spatial-aware)和任务感知(Task-aware) 3 层自注意力机制, 在无额外计算开销的情况下统一处理分类与定位任务。但该设计对特征层的三维张量(层级 \times 空间 \times 通道)进行全量注意力计算, 在边缘设备部署时会面临推理延迟与内存占用过高的问题。本研究基于 YOLOv8 原生检测头, 针对 Dynamic head 全张量注意力导致的计算与显存开销问题, 提出分层通道压缩、多路并行分支及可变堆叠深度的改进方案, 在确保表征能力不变的前提下, 显著降低推理延迟与内存占用。

- 1) 针对不同特征层级分辨率差异, 本方案首先在各检测分支 $cv1(x)$ 前引入基于层级自适应的通道压缩策略: 对高分辨率特征采用较小的中间通道数, 对低分辨率特征保留更大通道数。此策略在最大程度保留语义信息的同时, 有效减少高分辨率层的计算量。随后, 在分支末端仅执行一次统一重投影, 生成回归分支 $cv2$ 与分类分支 $cv3$ 所需的输出

通道，避免重复升降维操作，降低冗余计算。

2) **Dynamic head** 对各层特征均采用相同数量的 **DyHeadBlock** 堆叠，致使高分辨率层产生不必要的深度注意力开销，而低分辨率层堆叠深度不足，难以捕捉全局语义。为此，本文通过配置向量 $\{n_i\}$ 动态指定每个层级的堆叠次数：对 P3 和 P4 等细粒度层采用较浅的堆叠数 n_3 和 n_4 ，对 P5 等粗粒度层采用更深的堆叠数。该改进在保证多尺度信息充分融合的同时，降低了整个注意力子网络的 FLOPs。

3) 为进一步缓解分类与回归任务的特征冲突，降低单一注意力层内的通道开销，本模型在 **Dynamic head** 输出后，立即拆分为 2 个并行子网：回归分支 (**Regression Subnetwork, REGS**) 包含一层 3×3 卷积，用于回归特征预处理；分类分支 (**Classification Subnetwork, CLSS**) 包含一层 3×3 卷积，用于分类特征预处理。这避免了原始设计中在同一张量上对 C 通道同时进行 **Task-Aware** 注意力计算的高开销，通过轻量级瓶颈快速提取任务所需信息。

4) 回归与分类子网的输出在通道维度拼接，形成与 **YOLOv8 Detect** 兼容的四维预测张量，既保持端到端可导性，又实现高效并行计算。

2.4 损失函数

虽然传统的 **ShapeIoU** 方法可以通过关注边界框本身的形状和尺度来计算损失，从而使边界框回归更加准确；但是依旧存在 3 个问题：① 权重 w_w ， h_h 仅基于真实值 (**Ground Truth, GT**) 计算，忽略了预测框的形状偏差；② 中心距离与形状成本在同一尺度上加权，无法对极端长条或超扁平目标做更细腻的距离与形状平衡；③ 对小目标时，尺度本身微弱，导致惩罚项趋近于常数，难以显著影响梯度方向。因此，本文提出了改进 **ShapeIoU**：

1) 将权重计算扩展为同时依赖预测框和真实框的长宽，即

$$\omega_w = \frac{2(\omega_p^\alpha + \omega_g^\alpha)}{(\omega_p^\alpha + \omega_g^\alpha) + (h_p^\alpha + h_g^\alpha)} \quad (1)$$

$$h_h = \frac{2(h_p^\alpha + h_g^\alpha)}{(\omega_p^\alpha + \omega_g^\alpha) + (h_p^\alpha + h_g^\alpha)} \quad (2)$$

式中： ω_g^α 和 h_g^α 体现了对真实框长宽的直接依赖，且直接参与了分子与分母的计算，使得最终权重不仅反映预测框的形状，也同时反映了真实框的长宽比例； α 表示指数，控制对形状差异的敏感度。该

改进使得损失函数能够实时反映预测框的形状特征，从而对模型在极扁或超长目标上的定位误差进行更强惩罚。

2) 在中心距离项上，将原全局距离统一归一化改为按水平和垂直维度分别处理，即

$$D = \frac{h_h(x_p - x_g)^2 + \omega_w(y_p - y_g)^2}{c^2} \quad (3)$$

$$c^2 = \left[\max\left(x_p + \frac{\omega_p}{2}, x_g + \frac{\omega_g}{2}\right) - \min\left(x_p - \frac{\omega_p}{2}, x_g - \frac{\omega_g}{2}\right) \right]^2 + \left[\max\left(y_p + \frac{h_p}{2}, y_g + \frac{h_g}{2}\right) - \min\left(y_p - \frac{h_p}{2}, y_g - \frac{h_g}{2}\right) \right]^2 \quad (4)$$

式中：在水平 x 和垂直 y 方向分别加权归一化，计算中心偏差，利用式(1)中已混合真实框信息的 ω_w 和 h_h 对不同方向偏移赋予差异化关注；对长条形目标和扁平目标分别增强水平与垂直误差的惩罚。

3) 保持原有 **ShapeIoU** 的形状差异度量，并同步应用新权重，即

$$\omega_w = h_h \frac{|\omega_p - \omega_g|}{\max(\omega_p, \omega_g)}, \quad \omega_h = \omega_w \frac{|h_p - h_g|}{\max(h_p, h_g)} \quad (5)$$

$$S = (1 - e^{-\omega_w})^4 + (1 - e^{-\omega_h})^4 \quad (6)$$

$$L_{\text{ShapeIoU}} = 1 - \text{IoU}(b_p, b_g) + D + \frac{1}{2}S \quad (7)$$

通过以上改进，**YOLO-ShapeIoU** 在保持 **IoU** 基础不变的前提下，显著增强了对预测框自身形状和尺度差异的敏感性，从而在综合定位精度与训练稳定性方面较原始 **ShapeIoU** 实现了明显提升。

3 实验结论与分析

3.1 数据集和实验配置

Albany 大学与中国科学技术大学共同推出了 **UA-DETRAC (University at Albany Detection and Tracking Dataset)** 数据集^[30]，旨在为车辆检测与跟踪提供标准测试平台。该数据集由 Canon EOS 550D 相机摄取，录有 10 h 视频资料，帧率为每秒 25 帧，总计超过 14 万帧图像和 121 万个标注对象框。其中，训练集由 60 个视频序列组成，共包含 83 791 张图像，而测试集由 40 个视频序列构成，共 56 340 张

图像。该数据集面临多种挑战,如图 5 所展示的典型场景,涉及轿车、巴士、卡车等 4 类车型,并覆盖晴朗、夜晚、多云、下雨 4 种气候条件,以及高车辆密度和遮蔽情况。在车辆检测实验中,首先排除没有标签的图像,得到 82 085 张图像的训练集和 56 167 张图像的测试集。为了避免过拟合,每 10 帧采样一次,最终获得 8 209 张训练图像和 5 617 张评估图像。



图 5 UA-DETRAC 数据集中的代表性场景

Fig. 5 Representative scenarios in the UA-DETRAC dataset

在本实验中,选用的 GPU 设备为 NVIDIA GeForce RTX 4090(NVIDIA 公司),而 CPU 设备则是 Core i9-12900K @ 3.20 GHz(英特尔公司,中国成都制造)。实验在 Windows 11 操作系统下完成,深度学习框架采用 PyTorch 1.7.0 版本,搭配 CUDA 12.1 和 Python 3.8。为保障算法性能对比的公正性,所有算法均以一致的超参数配置进行训练与验证。实验环境配置见表 2。

表 2 训练参数设置

Table 2 Training parameter settings

编号	参数	设置
1	epochs	300
2	Batch	8
3	imgsz	640
4	workers	4
5	optimizer	SGD
6	close_mosaic	0
7	patience	50
8	warmup_epochs	3.0
9	warmup_momentum	0.8
10	lr0	0.01
11	lrf	0.01
12	mosaic	1.0
13	weight_decay	0.000 5

3.2 评估指标

在车辆检测领域,通过精确率(Precision, P)、召回率(Recall, R)、mAP@0.50 及 mAP@0.50:0.95 来评估实验成效。

1) 准确度。所有判定为阳性的样本中,真实阳性样本所占的比例为

$$P = \frac{TP}{TP + FP} \quad (8)$$

2) 召回率。所有实际阳性样本中,正确识别的阳性样本所占的比例为

$$R = \frac{TP}{TP + FN} \quad (9)$$

式中: TP 表示真阳性样本数; FP 表示假阳性样本数; FN 表示被错误判断为阴性的实际阳性样本数。

3) 平均准确度(Average Precision, AP)。表示某类样本 P-R 曲线下的面积为

$$AP = \int_0^1 PdR \quad (10)$$

4) mAP@0.50。指 IoU 阈值为 0.50 时,各分类的平均精度。

5) mAP@0.50:0.95。涵盖 IoU 阈值从 0.50 到 0.95(间隔 0.05)的所有分类的平均精度。

6) 模型计算量(Giga Floating-Point Operations Per Second, GFLOPs)。模型前向推理时所需的总浮点运算次数,单位为 Giga FLOPs,即

$$GFLOPs = \frac{\sum_1^L FLOPs_l}{10^9} \quad (11)$$

式中: $FLOPs_l$ 表示第 l 个算子(卷积、全连接等)所需的浮点运算数; L 表示算子总数。本文统计时以输入分辨率 640×640 计算。

7) 模型参数量(Params)。模型中所有可学习参数的总数,单位为百万(M),即

$$Params = \sum_1^L N_l \quad (12)$$

式中: N_l 表示第 l 个层的参数个数。

以上 7 项指标涵盖了检测准确性、定位精度及模型复杂度 3 个方面,能够全面评估所提 BSD-YOLO 在小目标车辆检测任务中的表现。

3.3 实验结果分析

3.3.1 动态稀疏感知主干网络实验比较

在算法效能评估研究阶段,本研究构建了动态稀疏感知主干网络的横向对比验证体系。为系统验证 ReBiAttention 模块的模型增强效果,实验设计采用 UA-DETRAC 基准数据集作为验证平台,通过模块替换策略实施横向对比。在实验参数设置层面,构建了全参数统一化的对照实验框架——包括网络拓扑架构、优化器配置以及超参数设定在内的所有变量均严格参照基准方案执行(表 2),从而有效确保不同 ReBiAttention 模块间横向性能对比的可靠性,为算法改进提供可量化的实验依据。

依据实验数据对比(表 3), 本研究选用的基础模型为 YOLOv8n 结构。在技术优化层面, 以“+”标志在 YOLOv8 原架构上进行的模块优化, 尤其是对 ReBiAttention 的创新性特征整合策略——在 C2f 模块中对 ReBiAttention 进行优化。在模块对比实验中, 本方法控制了除 ReBiAttention 插入位置以外的所有网络结构与训练设置(表 2), 以确保实验的公正性和横向可比性。ReBiAttention 模块在不同插入位置下的模型参数量保持在 2.93~2.95 M 左右, 整体计算复杂度稳定在 8.0 GFLOPs 左右, 说明该模块可灵活集成并控制参数增长。

在实验过程中, 先后尝试了多种 ReBiAttention 模块的集成方式, 包括轻量插入(如 C2f 位置)与全局集成(最终版本)。尽管最终版本在参数量(3.45 M)与计算量(11.6 GFLOPs)上相较基线 YOLOv8n 有所上升, 但在 mAP@0.50 和 Recall 这 2 项核心指标上分别提升了约 9.97%和 9.80%。该结果表明, 通过在允许的资源范围内适度提升模型复杂度, 可显著增强检测精度, 尤其在小目标密集场景中表现更佳。相比于轻量化配置, 最终方案在权衡参数效率与检测精度方面取得了更优结果, 体现了 BSD-YOLO 在实用性能与部署适应性之间的平衡优势。

3.3.2 动态检测头和不同网络结构实验比较

本研究在 UA-DETRAC 官方数据集上按 4:1 的比例划分训练集与验证集, 分别对原始 YOLOv8n 和改进后的 BSD-YOLO (包括 SN 模块与 DyHead-small 检测头)模型进行了 300 Epoch 的统一训练与评估, 在实验参数设置层面与表 2 保持一致。由表 4 的实验对比数据可知, 优化后的算法在保持实时检测性能的同时, 在核心性能指标上表现出明显的优越性。本实验分为 2 个部分:

1) 系列基线对比。YOLOv12 在 mAP@0.50 上略优于 YOLOv8n, 但在 Recall 与 mAP@0.50:0.95 上差距不大; YOLOv3 虽参数量庞大但精度表现亦可圈可点; YOLOv5 和 YOLOv6 在轻量化与精度间均未取得最佳平衡。

2) 增量模块对比。在基线 YOLOv8n 上, ReBiAttention 的 Recall 与 mA 明显提升; SN 模块在“轻量化”方向上表现突出, 参数量下降 6.7%且 mAP@0.50 微增; DyHead-small 检测头在 Precision 和 mAP 2 项核心指标上提升最为显著 (Precision+12.7%, mAP@0.50+5.98%), 同时参数量较基线下降 16.3%, 证明动态检测头在小目标尺度自适应方面的显著优势。

表 3 添加不同 ReBiAttention 模块的结果比较

Table 3 Comparison of the results of adding different ReBiAttention modules

实验方法	Precision	Recall	mAP@0.50	mAP@0.50:0.95	GFLOPs	Params/M
YOLOv8n	0.60881	0.598 17	0.589 04	0.431 27	8.1	3.00
+C2f-ReBiAttention	0.636 09	0.576 85	0.592 03	0.399 98	8.1	2.95
+C3-ReBiAttention	0.639 66	0.486 80	0.571 48	0.417 49	8.0	2.93
+CPN-ReBiAttention	0.646 37	0.563 50	0.598 16	0.424 37	8.0	2.93
+CSC-ReBiAttention	0.642 02	0.566 64	0.584 40	0.409 42	8.0	2.94
+ReBiAttention	0.615 95	0.656 82	0.648 06	0.468 26	11.6	3.45
Ours	0.696 20	0.615 02	0.661 59	0.468 31	7.9	2.87

注: 加粗数据表示最优值。

表 4 网络结构对比

Table 4 Comparison of network structures

实验方法	Precision	Recall	mAP@0.50	mAP@0.50:0.95	Params/M
YOLOv8n	0.608 81	0.598 17	0.589 04	0.431 27	3.00
YOLOv12	0.713 72	0.514 93	0.603 74	0.433 87	2.51
YOLOv6	0.570 75	0.543 44	0.556 63	0.413 99	4.23
YOLOv5	0.648 75	0.561 17	0.588 45	0.406 21	2.50
YOLOv3	0.679 22	0.578 32	0.607 95	0.453 78	103.66
YOLOv8n+ReBiAttention	0.615 95	0.656 82	0.648 06	0.468 26	3.45
YOLOv8n+SN	0.613 18	0.599 30	0.601 96	0.426 42	2.80
YOLOv8+Dyhead-small	0.735 32	0.590 65	0.648 83	0.475 72	2.51
Ours	0.696 20	0.615 02	0.661 59	0.468 31	2.87

注: 加粗数据表示最优值。

从表 5 数据可以看出, 在 YOLOv8 原始模型中引入一个 AFPN 检测头, 虽参数量大幅减少 29.67%, 但各类检测精度均有下滑, 故并不符合本文精度上的要求。而引入 Dyhead 检测头后使 Precision 提高了 13.15%, mAP@0.5 提高了 5.979%, mAP@0.5:0.95 提高了 4.445%, 体现了 Dyhead-small 检测头对小目标车辆数据集 UA-DETRAC 的适配性为后续阶段更精准的车辆检测奠定了基础。

本研究通过对照实验验证了改进型 YOLOv8 在车辆检测任务中的技术优势。实验数据显示, 相较于基准模型, 优化后的网络架构在保持计算效率的前提下实现多项性能突破, 模型综合性能达到当前最优水平。

3.3.3 消融实验比较

本研究采用模块化验证方法评估系统优化方案的有效性。通过设计控制变量实验, 重点考察 ReBiAttention 模块、Neck 网络拓扑结构调整、检测头增强以及损失函数调整策略四者的协同优化效果。实验平台选用 YOLOv8 系列框架, 该架构包含 n/s/m/l/x 5 种规模梯度, 本研究基于计算效能与模型精度的平衡考量, 选定轻量级 YOLOv8n 作为基础参照体系。经过完整训练周期后, 在标准验

证集上获得的量化评估数据见表 6。

本研究通过一系列消融实验, 全面验证了网络架构优化方案的有效性。从实验数据来看, 各种改进策略在提升模型检测性能和收敛速度方面均表现出积极效果。首先, 与基线模型 YOLOv8n 相比, 模型 A 在主干网络中引入了 ReBiAttention 模块, 尽管参数量仅增加 0.45 M, 但其检测精度(Precision)和平均精度(mAP@0.50)分别提升了 0.714%和 5.902%, 同时计算负荷保持稳定。这表明, 适度增加的参数开销能够有效改善特征提取能力, 从而提高检测性能。此外, 经过深度网络重构的模型 B 显示出更为明显的轻量化优势: 其参数量减少了 0.20 M, 同时检测性能也获得了突破——精度度上升 0.337%, mAP@0.50 提高 1.292 个百分点。这一结果验证了模型 B 在协同优化与轻量化设计上的综合效果, 有效解决了高层特征图的感受野过大导致小目标信息丢失的问题。在 Head 部分, 模型 C 通过引入 DynamicHead 小目标检测头, 使得 Precision, mAP@0.50 与 mAP@0.50:0.95 分别提升了 12.651%, 5.979%和 4.445%, 而参数量仅减少了 0.49 M, 这表明在细粒度特征层面的改进同样具有较高的性价比。

表 5 各类检测头对比

Table 5 Comparison of various types of detection heads

实验方法	Precision	Recall	mAP@0.50	mAP@0.50:0.95	Params/M
YOLOv8n	0.608 81	0.598 17	0.589 04	0.431 27	3.00
+AFPN ^[31]	0.580 92	0.520 42	0.541 67	0.391 88	2.11
+AFPN-small	0.630 40	0.576 21	0.588 82	0.431 25	3.67
+AFPN-large	0.637 76	0.492 45	0.548 28	0.384 95	2.11
+ASFFHead	0.636 27	0.567 77	0.585 66	0.416 71	4.38
+Dyhead-base	0.740 31	0.532 93	0.599 11	0.429 29	4.75
+Dyhead-large	0.650 17	0.579 04	0.604 89	0.420 23	13.10
+Dyhead-small	0.735 32	0.590 65	0.648 83	0.475 72	2.51
Ours	0.696 20	0.615 02	0.661 59	0.468 31	2.87

注: 加粗数据表示最优值。

表 6 消融实验

Table 6 Ablation experiments

实验方法	ReBiAttention	SN	Dyhead	ShapelIoU	Precision	Recall	mAP@0.50	mAP@0.50:0.95	Params/M
YOLOv8n					0.608 81	0.598 17	0.589 04	0.431 27	3.00
A	√				0.615 95	0.656 82	0.648 06	0.468 26	3.45
B		√			0.613 18	0.599 30	0.601 96	0.426 42	2.80
C			√		<u>0.735 32</u>	0.590 65	<u>0.648 83</u>	0.475 72	<u>2.51</u>
D	√	√			0.680 19	0.601 61	0.634 58	0.446 55	3.24
E	√		√		0.737 35	0.561 93	0.622 49	0.447 88	3.01
F		√	√		0.663 11	0.572 28	0.626 18	0.445 79	2.40
G	√	√	√		0.710 68	0.570 31	0.636 86	0.456 35	2.85
Ours	√	√	√	√	0.696 20	<u>0.615 02</u>	0.661 59	<u>0.4683 1</u>	2.87

注: 加粗数据表示最优值; 下划线数据表示次优值; √表示使用该模块。

进一步的组合优化实验显示：当在 Backbone 中同时引入 ReBiAttention 模块并对 Neck 层进行修改时，模型 D 的 Precision，mAP@0.50 和 mAP@0.50:0.95 分别提高了 7.138%，4.554% 和 1.528%；而模型 E 则在 Precision 与 mAP@0.50 上分别获得了 12.854% 和 3.345% 的提升，但由于 ReBiAttention 模块部分参数较大，与 DynamicHead 小目标检测头结合后，导致整体模型的计算量和参数量略有增加。相比之下，模型 F 在 SN 后引入 Dynamic head 小目标检测头，通过进一步优化网络框架，不仅实现了参数量下降 0.60 M，同时 Precision 和 mAP@0.50 分别提升了 5.43% 和 3.714%，充分证明了经过优化后的轻量小目标检测头和简化深层网络结构的高效性。

模型 G 综合采用了 3 种关键技术，其数据表现如下：Precision 为 0.710 68，Recall 为 0.570 31，mAP@0.50 为 0.636 86，mAP@0.50:0.95 为 0.456 35，同时模型参数仅为 2.85 M。从数据来看，模型 G 在提高检测精度方面表现突出，其较高的 Precision 说明在检测中误报较少，而适中的 Recall 则表明检测覆盖率尚可。尤其是 mAP 指标的提升，表明该模型在兼顾检测准确性与鲁棒性上取得了较好的平衡。这种优化策略表明，通过将 ReBiAttention 模块用于增强特征提取，SN 用于结构轻量化以及 Dyhead 用于细粒度目标检测，模型在不显著增加参数量的前提下，实现了检测性能的整体提升。

在模型 G 基础上引入 ShapeIoU 损失函数，构建完整改进模型 Ours 后实验数据显示，该方案在参数量仅减少 0.13 M 的情况下，综合检测性能显著提升：mAP@0.50 为 66.16%，mAP@0.50:0.95 为 46.83%，Recall 为 61.50%，ShapeIoU 通过目标形状感知机制优化定位精度，目标长宽比动态调整损失权重，缓解小目标因宽高比异常导致的定位偏差。综合而言，本研究不仅在精度上获得了明显改善，而且由于参数量较低，具有较高的实际应用价值，特别适用于对计算资源要求较高的小目标检测场景。

尽管本研究在 mAP@0.50 和 mAP@0.50:0.95 上均优于其他大部分组合，但在 Precision 和 Recall 这 2 项单独指标上略低于部分组合模型，其主要原因在于：

1) 引入 ShapeIoU 之后，模型在训练时会更加关注不同尺度与长宽比目标的一致度优化，倾向于提升整体 mAP，而非单项 Precision 或 Recall 的极致值。

2) BSD-YOLO 的设计初衷是通过稀疏注意力筛选出高质量特征区域，并结合轻量化检测头提升小目标的检测稳定性。因此，相比极端追求过多保留高阈值预测的 Precision 或尽可能多检测出所有目标(Recall)的模型，BSD-YOLO 在误检(FP)和漏检(FN)之间做了折中，最终在 mAP 指标上获得最佳平衡。

3) 本研究在保持参数量(2.87 M)和计算复杂度接近 SN+Dyhead 组合(2.85 M)的前提下，通过 ShapeIoU 和动态注意力的协同，使得模型对关键小目标的定位更为精确，但这也意味着在某些边缘样本上可能出现略高的漏检，从而导致 Recall 不及仅关注检测头改进的模型 C。

4) 在实际智能监控场景中，综合 mAP 能够更好地反映模型对各种尺度目标的检测效果与稳定性，因此本文重点优化并验证了 mAP，而对 Precision 或 Recall 的轻微下降是可接受的。

3.3.4 检测结果可视化

为了直观比较 BSD-YOLO 与经典模型在同一场景下的检测表现，图 6 将 YOLOv8n，YOLOv5，YOLOv3 以及本方法在 3 个不同时刻、同一视角下的检测结果并列展示。第 1 行，场景为远端车辆小目标分布最为稀疏的时刻，只有本方法成功识别出了最远端的那辆小车，其他 3 种模型均出现了漏检，充分说明了 BSD-YOLO 在极小目标提取上的优势；第 2 行展示的是检测精度和误检漏检的问题，YOLOv5 对远端车辆有漏检情况，YOLOv3 虽有所改进但检测精度仍然与本文改进模型有差距；第 3 行的场景为部分车身被遮挡，其他模型对远端重叠车辆识别精度有所降低，导致漏检与误报；而本方法依托动态稀疏注意力模块与自适应检测头的协同作用，成功捕捉到被遮挡小车，并使检测框准确贴合，其完整性和鲁棒性均显著提升。本模型不仅准确筛选了背景误检，还将紧邻的多辆小车用紧贴目标边缘的检测框完整覆盖，检测精度明显优于其他方法。

进一步地，图 7 给出了各算法在 UA-DETRAC 验证集上的定量对比指标。可以看到，BSD-YOLO 在 mAP 指标上全面领先，较 YOLOv8n 提高了约 7 个百分点，同时将误检率和漏检率降低。由此表明，基于动态稀疏注意力与轻量化双分支结构的改进，不仅在精度上取得了突破，也在实时性和资源占用上保持了良好平衡。

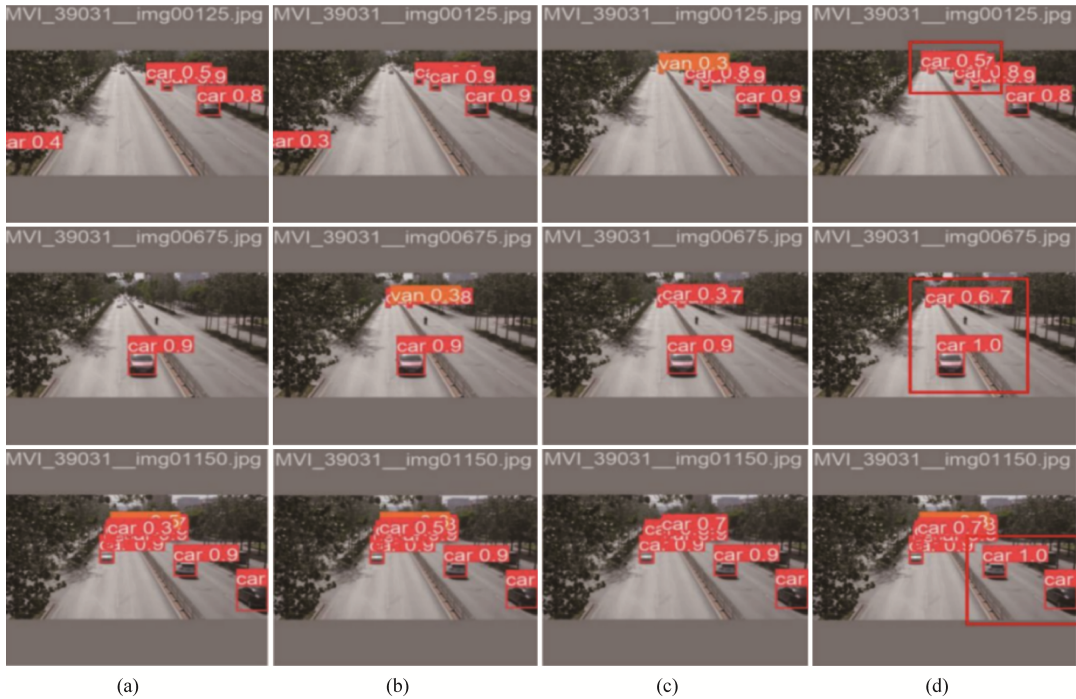


图 6 检测效果可视化对比

Fig. 6 Visual comparison of detection results ((a) YOLOv8n; (b) YOLOv5; (c) YOLOv3; (d) Ours)

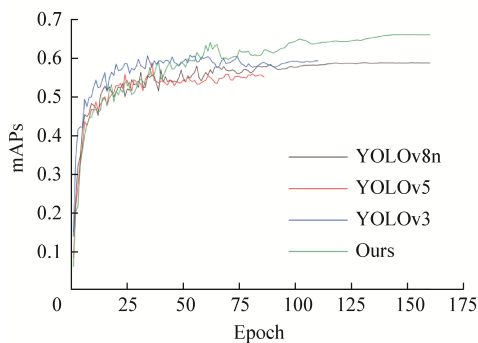


图 7 改进算法指标对比图

Fig. 7 Improved the comparison chart of algorithm indicators

4 结束语

基于 BSD-YOLO 算法, 本文提出一种针对车辆检测任务改进的算法。由实验结果可知, ReBiAttention 模块能够有效提升模型对关键特征区域的关注能力, 从而减少图像中无关信息的干扰; SN(Slim-Neck)的引入使得模型更加轻量化, Dynamic head 检测头增强了模型对小目标的检测能力, 显著提高了密集目标场景中的检测性能。

尽管如此, 本算法在某些极端或复杂环境中的表现仍面临挑战, 如夜间弱光、隧道低照度以及雾霾、雨雪等恶劣气象条件下, 图像质量下降会削弱 ReBiAttention 对细节的捕捉能力; 在车辆高速运动或长距离拍摄时, 小目标边缘信息更为模糊,

Dynamic head 的检测精度有所波动; 此外, 部分场景下的实时性需求(如高速公路监控和无人驾驶决策)对推理延迟提出了更高要求, 需要进一步优化硬件适配与算法并行性。

未来研究将进一步结合多尺度检测技术, 在保持模型实时性和计算效率的同时, 进一步提升小目标检测的准确性, 降低误检率和漏检率, 并针对检测任务进行了改进和验证, 后续工作将基于本方法进一步研究多目标跟踪(Multiple Object Tracking, MOT)性能, 并结合 MOTA 和 IDF1 等指标展开综合评估。推动模型在智能交通、城市规划和自动驾驶等领域的广泛应用, 为智能交通系统的可持续发展提供更加可靠的技术支持。

参考文献 (References)

- [1] 火久元, 苏泓瑞, 武泽宇, 等. 基于改进 YOLOv8 的道路交通小目标车辆检测算法[J]. 计算机工程, 2025, 51(1): 246-257. HUO J Y, SU H R, WU Z Y, et al. Road traffic small target vehicle detection algorithm based on improved YOLOv8[J]. Computer Engineering, 2025, 51(1): 246-257 (in Chinese).
- [2] NAVIA-VAZQUEZ A, GUTIERREZ-GONZALEZ D, PARRADO-HERNÁNDEZ E, et al. Distributed support vector machines[J]. IEEE Transactions on Neural Networks, 2006, 17(4): 1091-1097.
- [3] ARREOLA L, GUDIÑO G, FLORES G. Object recognition and tracking using Haar-like Features Cascade Classifiers: application to a quad-rotor UAV[C]//2022 8th International Conference on Control, Decision and Information Technologies. New York: IEEE Press, 2022: 45-50.
- [4] 杜铨熠. 基于改进 YOLOv8 的无人机航拍交通小目标检测

- 算法研究[D]. 大连: 大连交通大学, 2025.
- DU Q Y. Research on aerial traffic small target detection algorithm in UAV based on improved YOLOv8[D]. Dalian: Dalian Jiaotong University, 2025 (in Chinese).
- [5] 鞠默然, 罗海波, 王仲博, 等. 改进的 YOLO V3 算法及其在小目标检测中的应用[J]. 光学学报, 2019, 39(7): 0715004.
- JU M R, LUO H B, WANG Z B, et al. Improved YOLO V3 algorithm and its application in small target detection[J]. Acta Optica Sinica, 2019, 39(7): 0715004 (in Chinese).
- [6] 濮志远, 罗素云. 复杂交通场景下的目标检测方法[J]. 信息与控制, 2025, 54(4): 632-643.
- PU Z Y, LUO S Y. Object detection method in complex traffic scenarios[J]. Information and Control, 2025, 54(4): 632-643 (in Chinese).
- [7] 孙旭辉, 官铮, 王学. 红外与可见光图像分组融合的视觉 Transformer[J]. 中国图象图形学报, 2023, 28(1): 166-178.
- SUN X H, GUAN Z, WANG X. Vision transformer for fusing infrared and visible images in groups[J]. Journal of Image and Graphics, 2023, 28(1): 166-178 (in Chinese).
- [8] ZHU L, WANG X J, KE Z H, et al. BiFormer: vision transformer with bi-level routing attention[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2023: 10323-10333.
- [9] 黄崇庆, 徐慧英, 张晓雷, 等. BGR-YOLO: 基于 YOLOv8 改进的交通场景下目标检测算法[EB/OL]. (2025-04-08) [2025-05-29]. <https://link.cnki.net/urlid/43.1258.TP.20250408.1455.002>.
- HUANG C Q, XU H Y, ZHANG X L, et al. BGR-YOLO: an improved object detection algorithm under traffic scenarios based on YOLOv8[EB/OL]. (2025-04-08) [2025-05-29]. <https://link.cnki.net/urlid/43.1258.TP.20250408.1455.002>. (in Chinese).
- [10] 刘熠龙, 张自立, 冯冀宁. 基于 UAV-YOLO 的无人机航拍图像轻量化目标检测算法[J]. 现代电子技术, 2025, 48(15): 51-56.
- LIU Y L, ZHANG Z L, FENG J N. UAV-YOLO-based lightweight object detection algorithm for UAV aerial images[J]. Modern Electronics Technique, 2025, 48(15): 51-56 (in Chinese).
- [11] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. [2025-04-16]. <https://arxiv.org/abs/1704.04861>.
- [12] LUO Y H, CAO X, ZHANG J T, et al. CE-FPN: enhancing channel information for object detection[J]. Multimedia Tools and Applications, 2022, 81(21): 30685-30704.
- [13] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures[J]. Journal of Real-Time Image Processing, 2024, 21(3): 62.
- [14] ZHENG W, TANG W L, JIANG L, et al. SE-SSD: self-ensembling single-stage object detector from point cloud[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2021: 14489-14498.
- [15] MIAO T, ZENG H C, YANG W, et al. An improved lightweight RetinaNet for ship detection in SAR images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 4667-4679.
- [16] MAITY M, BANERJEE S, CHAUDHURI S S. Faster R-CNN and YOLO based vehicle detection: a survey[C]//The 5th International Conference on Computing Methodologies and Communication. New York: IEEE Press, 2021: 1442-1447.
- [17] CHAI B S, NIE X, ZHOU Q F, et al. Enhanced cascade R-CNN for multiscale object detection in dense scenes from SAR images[J]. IEEE Sensors Journal, 2024, 24(12): 20143-20153.
- [18] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2016: 779-788.
- [19] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 6517-6525.
- [20] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 936-944.
- [21] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. [2025-04-16]. <https://arxiv.org/abs/2004.10934>.
- [22] NELSON J, SOLAWETZ J. YOLOv5 is here: state-of-the-art object detection at 140 FPS[EB/OL]. (2020-06-10) [2025-04-16]. <https://blog.roboflow.com/yolov5-is-here/>.
- [23] LI C Y, LI L L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. [2025-04-16]. <https://arxiv.org/abs/2209.02976>.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[EB/OL]. [2025-04-16]. <https://arxiv.org/pdf/2207.02696.pdf>.
- [25] Ultralytics. YOLOv8(8.0)[EB/OL]. [2025-04-16]. <https://github.com/ultralytics/ultralytics>.
- [26] 杨锦辉, 李鸿, 杜芸彦, 等. 基于改进 YOLOv5s 的轻量化目标检测算法[J]. 电光与控制, 2023, 30(2): 24-30.
- YANG J H, LI H, DU Y Y, et al. A lightweight object detection algorithm based on improved YOLOv5s[J]. Electronics Optics & Control, 2023, 30(2): 24-30 (in Chinese).
- [27] YU B Y, LI Z X, CAO Y, et al. YOLO-MPAM: efficient real-time neural networks based on multi-channel feature fusion[J]. Expert Systems with Applications, 2024, 252: 124282.
- [28] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: unifying object detection heads with attentions[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2021: 7369-7378.
- [29] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: attention-centric real-time object detectors[EB/OL]. [2025-04-16]. <https://arxiv.org/abs/2502.12524>.
- [30] WEN L Y, DU D W, CAI Z W, et al. UA-DETRAC: a new benchmark and protocol for multi-object detection and tracking[J]. Computer Vision and Image Understanding, 2020, 193: 102907.
- [31] YANG G Y, LEI J, ZHU Z K, et al. AFPN: asymptotic feature pyramid network for object detection[C]//2023 IEEE International Conference on Systems, Man, and Cybernetics. New York: IEEE Press, 2023: 2184-2189.