

基于改进 RT-DETR 模型的雨天环境 车辆目标检测研究



刘阳¹, 黎鹏², 胡玛莉¹, 隆伟², 廖雪花², 朱洲森^{1,*}

¹ 四川师范大学物理与电子工程学院, 四川成都 610101

² 四川师范大学计算机科学学院, 四川成都 610101

摘要: 针对雨天环境下车辆检测易受能见度降低、雨滴遮挡及对比度不足等因素影响而导致检测精度下降的问题, 本文提出一种基于改进 RT-DETR-R18 模型的雨天车辆检测方法 (R-RTDETR)。采用 Conv 与 C2f-wConv 组合网络替换原主干网络, 通过加权卷积实现特征自适应调节, 以增强关键区域信息表达; 在编码器, 引入全局-局部空间注意力模块 (Global-Local Spatial Attention, GLSA) 以实现全局语义与局部细节的高效融合; 在解码器, 引入 P2 浅层特征分支, 优化多尺度融合路径, 强化车辆感知能力。实验结果表明, R-RTDETR 在测试集上的 mAP 较基线 RT-DETR-R18 提升 1.52%, 召回率提升 1.25%; 在 ACDC 雨天数据集上的 mAP@0.5 提升 2.33%, 在保证计算效率的同时提高了雨天车辆检测的稳定性与鲁棒性, 验证了其在智能交通场景中的应用潜力。

关键词: 目标检测; RT-DETR; 雨天车辆检测; 注意力机制; 深度学习

DOI: 10.57237/j.cst.2026.02.001

Research on Vehicle Target Detection in Rainy Environment Based on Improved RT-DETR Model

Liu Yang¹, Li Peng², Hu Mali¹, Long Wei², Liao Xuehua², Zhu Zhousen^{1,*}

¹ College of Physics and Electronic Engineering, Sichuan Normal University, Chengdu 610101, China

² College of Computer Science, Sichuan Normal University, Chengdu 610101, China

Abstract: To address the issue of decreased vehicle detection accuracy in rainy conditions due to reduced visibility, raindrop obstruction, and insufficient contrast, this paper proposes a rain-based vehicle detection method (R-RTDETR) based on an improved RT-DETR-R18 model. The original backbone network is replaced with a combined Conv and C2f-wConv network, and weighted convolutions are used to adaptively adjust features, enhancing the representation of key region information. In the encoder, a Global-Local Spatial Attention (GLSA) module is introduced to achieve efficient fusion of global semantics and local details. In the decoder, a P2 shallow feature branch is introduced to optimize the multi-scale fusion path and strengthen vehicle perception capabilities. Experimental results show that R-RTDETR improves mAP by 1.52% and recall by 1.25% on the test set compared to the baseline RT-DETR-R18; on the ACDC rainy dataset, mAP@0.5 is improved by 2.33%, demonstrating improved stability and robustness of rain-based vehicle detection while maintaining computational efficiency, validating its application potential in intelligent transportation scenarios.

基金项目: 国家社会科学基金一般项目 (20BMZ092).

*通信作者: 朱洲森, zhuzhousen@sicnu.edu.cn

收稿日期: 2026-02-11; 接受日期: 2026-03-04; 在线出版日期: 2026-03-24

<http://www.computscitech.com>

Keywords: Object Detection; RT-DETR; Vehicle Detection in Rainy Weather; Attention Mechanism; Deep Learning

1 引言

计算机视觉技术在自动驾驶与智能交通系统中应用广泛,但雨、雾、雪等恶劣天气会导致图像质量下降,显著降低车辆检测精度,进而威胁行车安全。研究表明,低能见度天气下交通事故发生率明显高于正常条件[1]。在机动车保有量持续增长的背景下,暴雨等复杂气象进一步加剧了追尾、侧滑等风险[2]。因此,提升雨天环境下车辆检测的精度与鲁棒性具有重要现实意义。

近年来,深度学习目标检测算法快速发展,主要包括双阶段与单阶段方法。双阶段检测算法首先生成候选区域,再完成目标的分类与定位,典型代表包括区域卷积神经网络(Region-based Convolutional Neural Network, R-CNN)[3]及其演进版本 Faster R-CNN(Faster Region-based Convolutional Neural Network)[4];单阶段检测算法则无需候选区域生成步骤,将目标检测任务直接建模为回归问题,代表性算法包括 YOLO(You Only Look Once)系列[5]、SSD(Single Shot MultiBox Detector)[6]、DETR(Detection Transformer)[7]以及其高效实时版本 RT-DETR(Real-Time Detection Transformer)[8]等,在速度与精度平衡上更具优势。

现有方法在常规场景下表现良好,但在雨天环境中,雨滴遮挡、图像模糊、对比度低等问题仍会导致检测性能下降。此外,恶劣天气数据采集困难、标注成本高,也制约了模型鲁棒性提升。针对雨天车辆检测问题,现有研究已开展多项改进:胡待方等[9]通过渐进式去雨提升 YOLOv5 抗干扰能力;陈婷等[10]融合去雨网络与 YOLOv4,但泛化性不足;Ogino 等[11]基于 YOLOv3 增强图像质量,但对目标检测保持能力有限;Jing 等[12]在 YOLOv8 中引入 Swin Transformer 模块和高效多尺度注意力(EMA)机制可提升雨天检测精度,但在极端雨量、强遮挡等场景的泛化能力仍有限,且计算复杂度较高;Tabassum 等[13]提出多模态车辆检测网络(MVDNet),通过多头注意力增强传感器融合效果,但部署成本较高;Soheil Gharatappeh 等[14]基于 RT-DETR 优化恶劣天气检测,但对浅层细节特征利用不足且计算量复杂。

目前主流雨天车辆检测方法多基于 YOLO 系列,但其检测性能在一定程度上依赖非极大值抑制(NMS),

对检测速度与精度形成约束[15]。相比之下,端到端目标检测框架 DETR 摒弃 NMS 操作,利用 Transformer 实现全局关系建模,在复杂雨天干扰场景下具有更强的鲁棒性。然而,DETR 存在计算复杂度高、收敛速度慢等问题,难以满足实时应用需求[8]。RT-DETR 在 DETR 基础上进行轻量化改进,其中 RT-DETR-18 版本在保证检测精度的同时显著降低了参数规模与推理开销,为本文雨天车辆检测方法提供了合适的基准模型。

此外,模型性能提升依赖高质量训练数据。针对现有数据集中雨天样本不足的问题,本文自采雨天图像并且整合 BDD100K、DAWN 公开数据集,构建多源融合雨天车辆检测数据集,并结合图像增强技术提升数据多样性与模型泛化能力。

在此基础上,本文以 RT-DETR-R18 为基准进行改进:首先,将主干网络替换为 Conv 与 C2f-wConv 组合,通过加权卷积自适应强化关键区域特征;其次,在编码器中引入 GLSA 模块,实现全局与局部特征高效融合;最后,在解码器增加 P2 浅层分支并优化 FPN/PAN 融合路径,缓解雨天条件下车辆细节信息退化。上述改进在不显著增加模型复杂度的前提下,提升雨天车辆检测的精度与稳定性,为复杂气象条件下智能交通系统的安全应用提供了技术支撑。

2 RT-DETR-R18 网络模型

主干网络作为特征提取模块,由卷积单元(ConvNormLayer)、残差块(BasicBlock)和最大池化层构成,用于提取并下采样图像特征,最终输出 S3、S4、S5 三个尺度特征图(下采样倍数分别为 8、16 和 32),为编码器提供多尺度输入。

高效混合编码器为模型核心,由基于注意力机制的同尺度特征交互模块(Attention-based intra-scale feature interaction, AIFI)和基于卷积神经网络的跨尺度特征融合模块(Cross-scale feature fusion module, CCFM)组成。其中, AIFI 作用于语义信息最丰富的 S5 特征图,以增强全局上下文建模能力;CCFM 通过卷积与 RepC3 模块对多尺度特征进行对齐与融合,并与 AIFI 输出共同实现跨尺度信息整合,从而提升不同尺度目标的检

测能力。

头部解码器采用多层 Transformer 架构，并创新性引入 IoU 感知的查询选择机制以筛选高质量候选查询，

同时设置辅助预测头以提升训练稳定性与检测精度。

RT-DETR-R18 模型的网路结构如下图 1 所示。

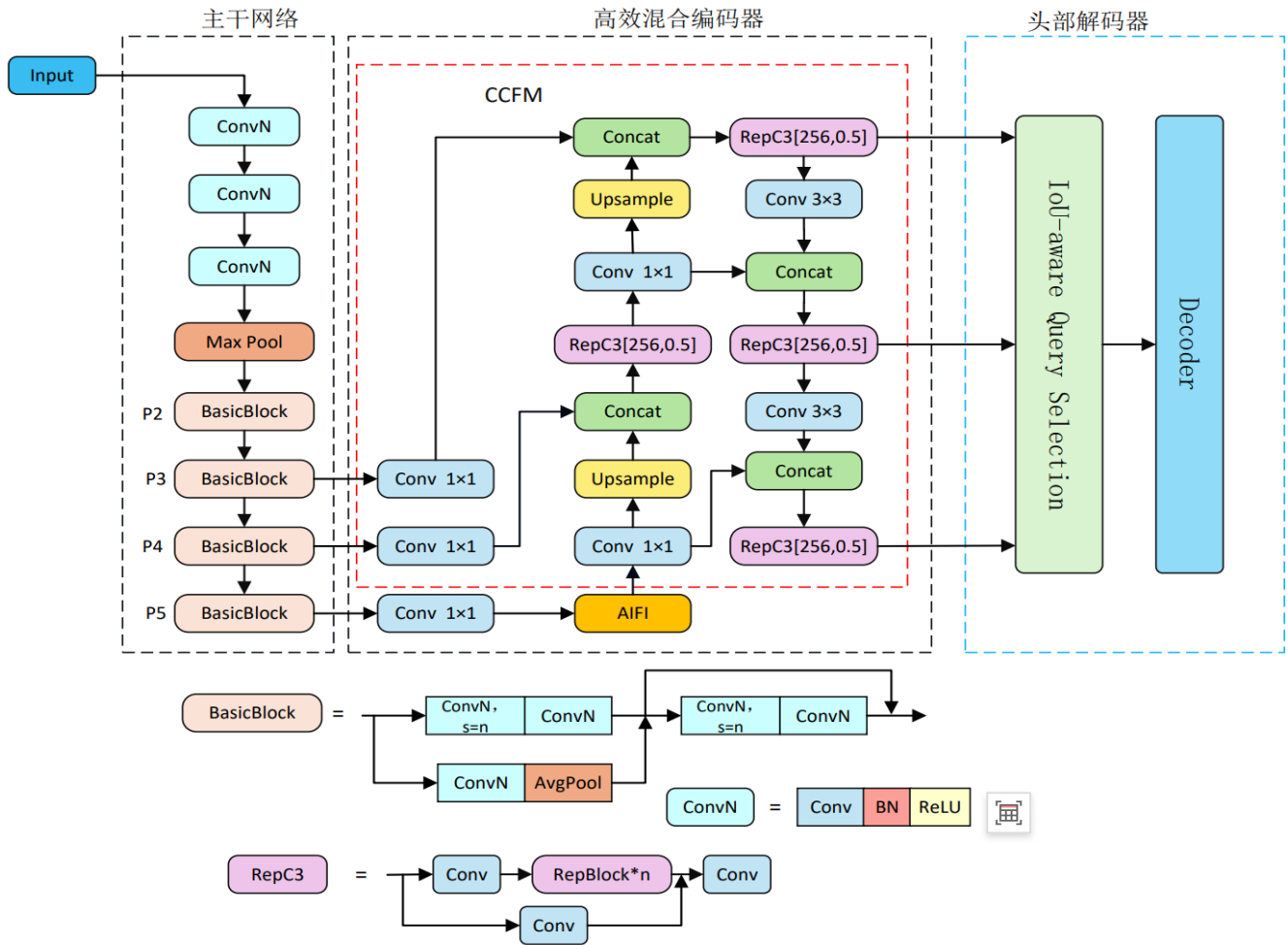


图 1 RT-DETR-R18 网络结构图

3 改进的 RT-DETR-R18 网络结构

3.1 改进网络结构

针对雨天场景中雨滴干扰、图像模糊及车辆特征弱化等问题，本文在 RT-DETR-R18 基础上提出一种结构改进方法，整体框架如图 2 所示。

在主干网络中，引入 Conv 与 C2f-wConv 组合模块替代原有结构。C2f-wConv 通过加权卷积机制对不同空间位置赋予差异化权重，突出关键区域特征并抑制雨滴噪声，从而增强模型在低对比度和模糊条件下的特

征提取能力。

在编码器部分，引入全局-局部空间注意力（GLSA）模块替代 1×1 卷积。GLSA 通过全局分支建模长程依赖关系，并结合局部分支强化边缘与细节特征，使模型在复杂雨天场景中兼顾全局语义与局部信息表达。

在解码器部分，引入 P2 特征分支，并对原有 FPN/PAN 融合路径进行调整，将主干网络中较高分辨率的 P2 特征引入解码阶段参与预测。通过通道压缩与特征增强操作，在保证计算开销可控的前提下补充细节信息；同时，利用下采样方式将 P2 特征与高层特征进行融合，实现语义层级对齐，从而缓解雨天条件下车辆边缘模糊和细节退化对检测性能的影响。

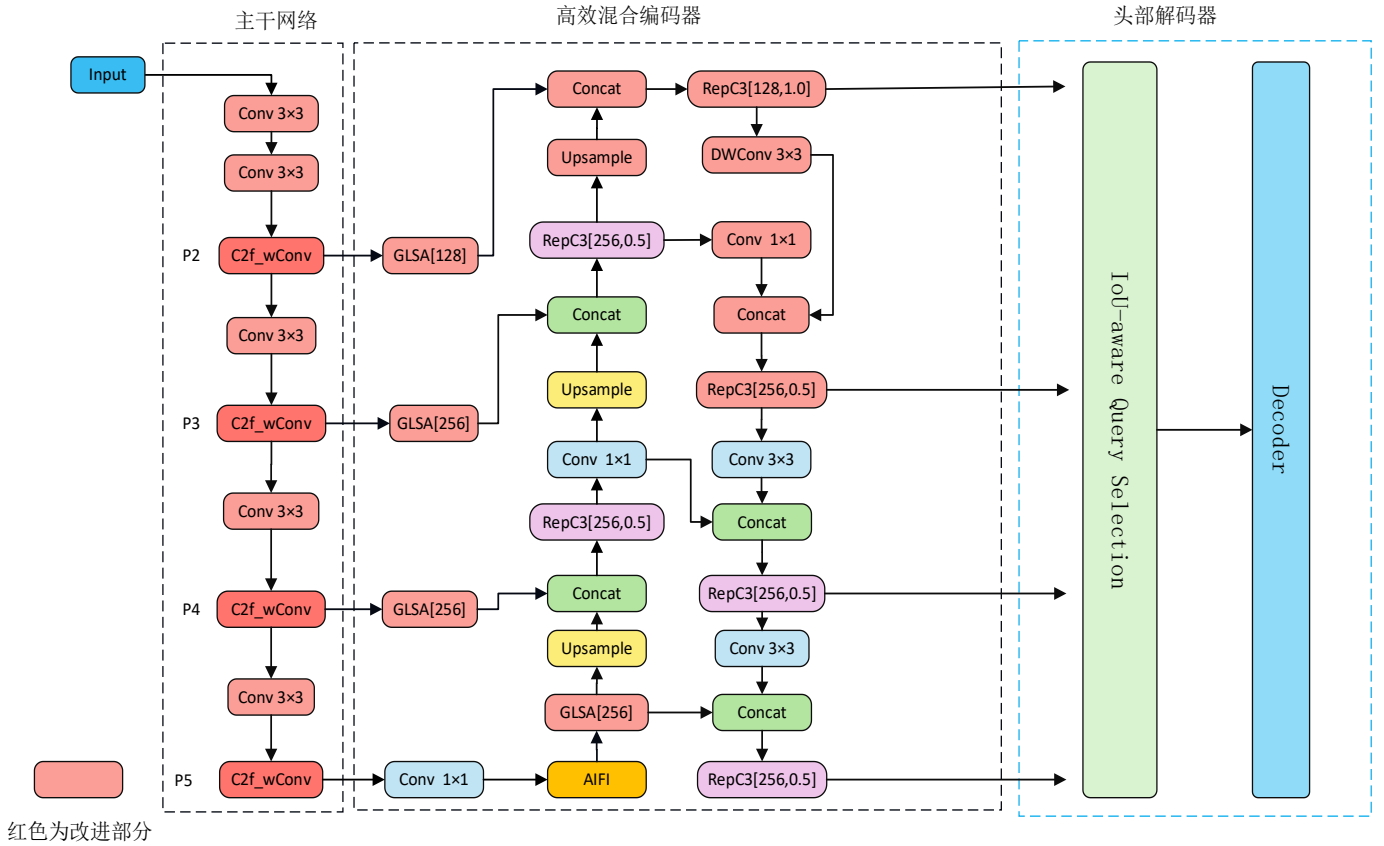


图2 改进的 RT-DETR-R18 网络结构图

3.2 基于 C2f-wConv 的主干网络改进

RT-DETR-R18 主干网络由多层卷积与 BasicBlock 残差模块堆叠而成，在常规场景下兼顾特征提取与计算效率。然而，BasicBlock 的标准卷积对空间位置等权分配，在雨天易受雨滴遮挡和成像模糊干扰，使低对比度区域与背景噪声混淆，削弱关键特征表达，导致检测性能下降。为此，本文将主干网络改为 Conv 与 C2f-wConv 交替堆叠，以增强雨天条件下的特征判别能力与整体鲁棒性。

C2f-wConv 模块基于 C2f 结构，将标准卷积替换为加权卷积 (Weighted Convolution, wConv) [16]。不同于传统卷积中的均匀权重分布，wConv 引入空间密度函数，对卷积核内部不同位置赋予差异化权重系数，从而实现对关键区域的自适应增强。在连续形式下，传统卷积表示为公式(1)，wConv 表示为公式(2)，其离散实现为公式(3)：

$$(f * g)(t) = \int_{\Omega} f(\tau)g(t-\tau)d\tau \quad (1)$$

$$(f * g_{\varphi})(t) = \int_{\Omega} f(\tau)(\varphi(t-\tau)g(t-\tau))d\tau \quad (2)$$

$$(W_{\varphi})_{ij} = \Phi_{ij} \cdot w_{ij} \quad (3)$$

其中， $(W_{\varphi})_{ij}$ 表示传统卷积核的权重参数， Φ_{ij} 表示位置处 (i, j) 的密度系数， w_{ij} 则是经过加权处理后的卷积核元素。该机制使卷积核能够自适应聚焦车辆关键区域特征，并有效抑制雨滴噪声及模糊背景等非关键信息的干扰。

在具体实现中，C2f-wConv 采用 3×3 卷积核，设置密度参数 $[0.9, 0.9, 0.9]$ ，通过动态调节权重强化目标关键区域响应、抑制噪声，从而显著提升雨天环境下的特征表达能力。此外，由于加权机制仅作用于卷积核内部参数，不引入额外计算开销，因此能够在维持原有推理效率的基础上，提升模型在复杂雨天场景中的判别性能。

3.3 基于 GLSA 的特征增强

原始 RT-DETR-R18 编码器特征投影采用 1×1 卷积线性映射，雨天场景中车辆边缘受雨滴、反光等干扰，

1×1 卷积难以突出关键特征，易引入噪声，削弱车辆目标检测性能。为此，本文在编码器的特征投影层中引入全局-局部空间注意力（Global-Local Spatial Attention, GLSA）模块[17]，替代该 1×1 卷积。该模块通过双支路结构同时建模全局依赖与局部细节，提升雨雾环境下的特征判别能力，结构如图 3 所示。

GLSA 模块[17]，替代该 1×1 卷积。该模块通过双支路结构同时建模全局依赖与局部细节，提升雨雾环境下的特征判别能力，结构如图 3 所示。

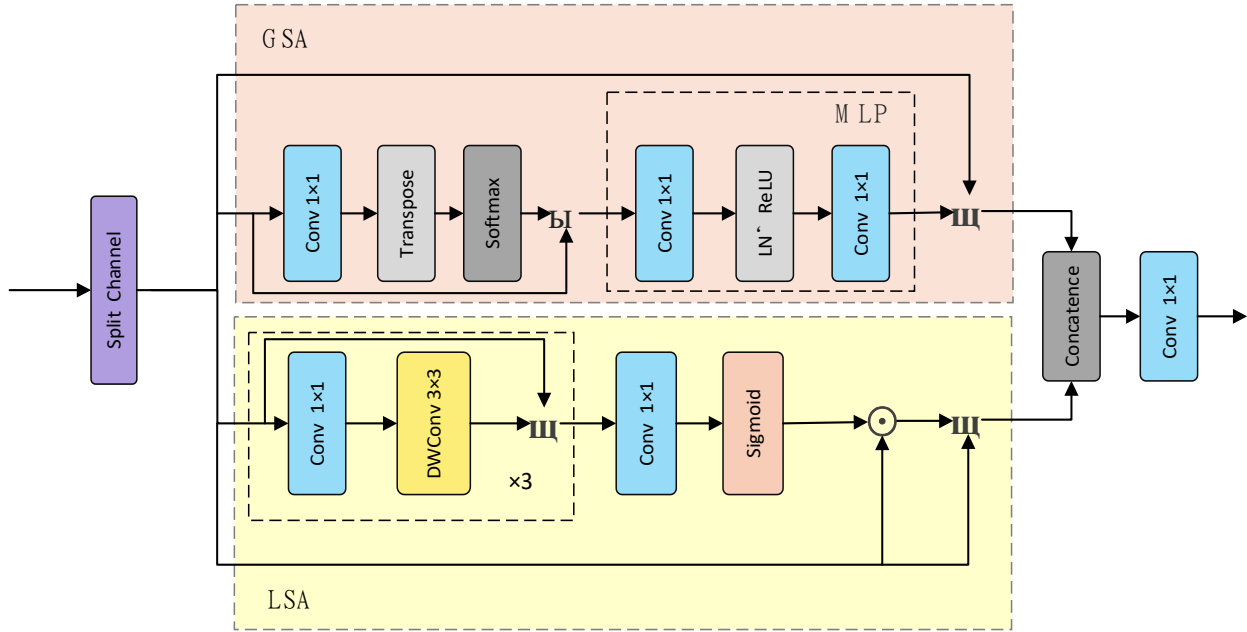


图 3 GLSA 模型结构图

GLSA 由 GSA（全局注意力分支）与 LSA（局部注意力分支）组成，输入特征沿通道均分后分别送入两支路。GSA 分支用于建模像素长程空间依赖，经 1×1 卷积通道映射、转置操作及 Softmax 生成全局注意力权重矩阵，捕获长程依赖并聚合全局上下文信息，再通过多层感知机（MLP）非线性变换，输出增强后的全局语义特征；LSA 分支侧重增强局部感兴趣区域响应，经卷积与 Sigmoid 生成局部注意力图，结合残差连接保留原始细节，同时确保梯度稳定与训练收敛。两支路输出拼接融合后，经 1×1 卷积整合降维生成高判别性特征，控制参数数量的同时提升雨天检测鲁棒性。整体的过程如图式(4)~式(9)所示。

$$F'_i = \text{Conv}_{1 \times 1}(\text{Concat}(G_{sa}(F_i^1), L_{sa}(F_i^2))) \quad (9)$$

其中，将输入特征 F_i 沿通道维度均分为两个子特征 F_i^1 和 F_i^2 ，分别送入 GSA 与 LSA 分支；式(5)~(6)为 GSA 运算， $\text{Att}_G(F_i^1)$ 是全局注意力权重矩阵， $\otimes$ 表示矩阵乘法，经加权融合、MLP 变换与残差连接，输出全局增强特征 $G_{sa}(F_i^1)$ ；式(7)~(8)为 LSA 运算， $\text{Att}_L(F_i^2)$ 是局部注意力权重图， $\square$ 表示逐元素乘法，经逐元素加权与残差连接，输出局部增强特征 $L_{sa}(F_i^2)$ ；式(9)拼接双分支输出，通过 1×1 卷积整合通道，最终得到与输入 F_i 同维度的增强特征 F'_i 。

$$F_i^1, F_i^2 = \text{Split}(F_i) \quad (4)$$

$$\text{Att}_G(F_i^1) = \text{Softmax}(\text{Transpose}(\text{Conv}_{1 \times 1}(F_i^1))) \quad (5)$$

$$G_{sa}(F_i^1) = \text{MLP}(\text{Att}_G(F_i^1) \otimes F_i^1) + F_i^1 \quad (6)$$

$$\text{Att}_L(F_i^2) = \sigma(\text{Conv}_{1 \times 1}(F_c(F_i^2) + F_i^2)) \quad (7)$$

$$L_{sa}(F_i^2) = \text{Att}_L(F_i^2) \square F_i^2 + F_i^2 \quad (8)$$

3.4 目标检测改进方法

3.4.1 基于 P2 浅层特征的检测分支构建

在图像特征提取中，原始 RT-DETR-R18 对远处或遮挡车辆等依赖高分辨率细节的目标存在局限：多次下采样导致纹理和边缘信息丢失，增加漏检和定位偏差风险[18]。尽管 RT-DETR-R18 凭借 Transformer 架构具备优异的全局建模能力，但对浅层高分辨率特征利

用不足,限制了复杂场景下的特征表达[19]。原检测头仅使用 P3、P4、P5 特征,其中最小尺度 P3 (1/8 下采样)语义弱、表达能力有限,经多次卷积后目标响应衰减,影响检测性能。

为增强空间细节感知,本文在检测头中引入 P2 浅层特征 (1/4 下采样)。具体流程为:通过 1×1 卷积将

P2 特征通道压缩至 128,与上采样至相同尺度的 P3 特征拼接,实现浅层细节与深层语义融合;随后输入 RepC3 模块进行非线性增强,其中该分支 RepC3 重复因子设为 1.0 (其余分支为 0.5),强化浅层特征表达。改进后结构如图 4 所示。

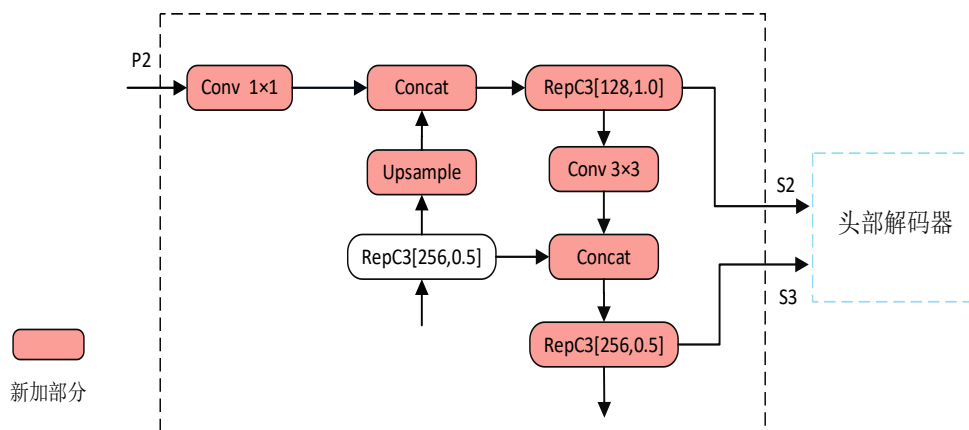


图 4 融合 P2 分支的 FPN/PAN 结构图

3.4.2 基于深度可分离卷积的特征融合优化

为进一步提升 P2 浅层特征的利用,本文对 FPN/PAN 路径中 P2→P3 的下采样与特征融合过程进行优化。核心采用深度可分离卷积 (Depthwise Separable Convolution, DSConv) 替代标准卷积进行 P2→P3 下采样。DSConv 将卷积分解为逐通道空间卷积和逐点卷积,既降低计算量,又能保留 P2 的边缘与纹理细节,同时方便与高层特征融合。

具体流程为:首先,将 P3 特征经 1×1 卷积降维,使其通道数与通过 3×3 DWConv 下采样后的 P2 特征一致;随后在通道维度拼接,形成 256 维融合特征。DWConv 对各通道独立执行空间卷积,在尺度变换过程中保留浅层细粒度信息;Conv 1×1 则负责通道整合与映射,保持局部结构完整。最终,融合特征经 RepC3 模块非线性增强,提升浅层特征的判别能力。优化后的结构如图 5 所示。

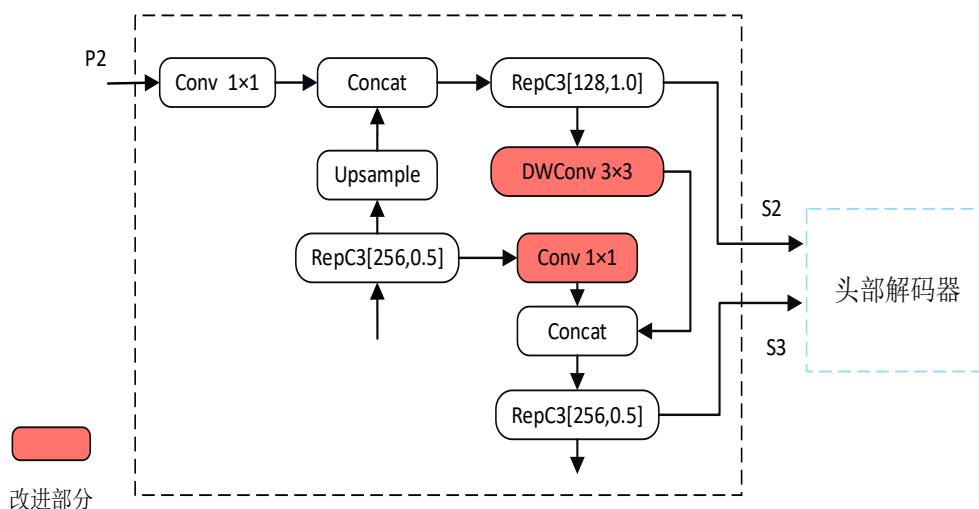


图 5 基于 DSConv 优化的 P2 分支 FPN/PAN 结构图

4 实验与结果分析

4.1 实验数据集收集与处理

针对公开雨天车辆检测数据集规模有限、场景单一的问题,本文构建多源融合雨天数据集,以突破样本不足的制约。实验数据集由公开数据与真实采集数据共同组成,主要来源包括 BDD100K、DAWN 数据集及实地采集的雨天道路图像。

BDD100K 数据集[20]由加州大学伯克利分校 AI 实验室发布,是目前公开的自动驾驶数据集中规模最大、场景最丰富且标注最全面的代表之一,包含晴、雨、雾、雪等多种天气与光照条件。DAWN 数据集[21]包含多种恶劣天气下的真实场景,覆盖城市道路、高速公路及不同交通密度,有助于提升模型鲁棒性。本文选取 car、bus、motorcycle、bicycle 四类车辆目标作为检测对象。

为丰富雨天样本,本文引入 ACM (Adaptive Composite Model) 与 HRM (Heavy Rain Model) 两种加雨方法对 BDD100K 中雨天特征不足的样本进行增强。ACM 通过复合叠加方式模拟不同方向与强度的雨纹,HRM 侧重刻画强降雨下的视觉退化特征,构建涵盖小雨、中雨及强降雨的多样化样本。同时,为提升数据集泛化能力,本文还结合 DAWN 真实雨天车辆数据及自采标注雨天道路图像。

最终构建的多源融合雨天车辆检测数据集由真实样本与合成样本共同组成,按 7:2:1 划分为 5650 张训练集、1620 张验证集和 807 张测试集。其中,测试集全部由真实雨天图像构成,场景覆盖均衡,可保证模型性能评估具备合理性与代表性。所有数据均转换为 YOLO 标注格式,满足 RT-DETR-R18 模型的训练需求。数据增强前后的对比如图 6 所示。



图 6 图像加雨对比展示

4.2 实验环境及参数配置

本研究的实验环境基于 Windows 11 操作系统,配

备 Intel(R) Xeon(R) Gold 5418Y 处理器 (10 核心)、NVIDIA GeForce RTX 4090 显卡及 120 GB 内存。深度学习框架基于 PyTorch 2.3.0 构建,编程语言为

Python 3.12。

训练过程中，输入图像尺寸设为 640×640，优化器采用 AdamW 以确保训练过程的稳定收敛。为增强数据多样性并缓解过拟合，训练阶段引入数据增强策略，其中图像水平翻转概率设为 0.5。具体实验参数设置如表 1 所示。

表 1 实验训练参数配置表

参数名称	参数值
Learning Rate	0.001
Epochs	160
Batch Size	4
Image Size	640×640
Momentum	0.9
Fliplr	0.5

4.3 评估指标

为评估所提出模型在雨天环境下的车辆目标检测性能，本文选取平均精度均值（mean Average Precision, mAP）、精确率（Precision, P）和召回率（Recall, R）作为主要评价指标，并引入模型参数量作为衡量模型复杂度与部署可行性的辅助指标。

平均精度均值（mAP）用于衡量模型在多类别目标检测任务中的整体检测性能。对于单一类别，通过绘制 Precision-Recall（P-R）曲线并计算其曲线下的面积，可得到该类别的平均精度（Average Precision, AP）。对所有类别的 AP 值取平均，即可得到 mAP，其计算公式如下：

$$mAP@0.5 = \frac{1}{c} \sum_{i=1}^c AP_i \tag{10}$$

$$mAP@0.5:0.95 = \frac{1}{T} \sum_{t=0.5}^{0.95} AP_i \tag{11}$$

其中，c 表示目标类别总数，AP_i 是第 i 类目标的平均精度，T 是 IOU 阈值的数量。mAP@0.5 表示在 IoU 阈值为 0.5 时，对各类别 AP 取平均；mAP@0.5:0.95 表示在 IoU 阈值从 0.5 到 0.95（步长为 0.05）范围内计算 AP，并对所有阈值与类别取平均。

准确率（Precision，又称查准率）用于衡量模型预测结果的准确性，其计算公式如下：

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

召回率（Recall，又称查全率）用于衡量模型对目标的检出能力，其计算公式如下：

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

上述两式中，TP 表示正确检测的目标数量，FP 表示误检目标数量，FN 表示漏检目标数量。此外，参数量为模型可参与训练的参数总数，该指标用以衡量模型的复杂度。

4.4 消融实验

为验证本文所提出各改进模块的有效性，本文以 RT-DETR-R18 作为基线模型，在保持训练策略与参数设置一致的条件下开展消融实验，实验结果如表 2 所示。其中，Backbone-C2fw 表示采用 Conv 与 C2f-wConv 组合结构替代原主干网络；GLSA 表示在颈部网络中引入全局-局部空间注意力模块；P2-FPN 表示引入 P2 特征分支并对 FPN/PAN 的特征融合路径进行优化。

单模块实验结果表明，引入 Backbone-C2fw 后，mAP@0.5 提升 0.56%，参数量减少 5.98M，体现了在降低模型复杂度的同时提升检测性能；引入 GLSA 模块后，mAP@0.5 提升 0.34%，表明注意力机制对雨天车辆检测具有积极作用；引入 P2-FPN 模块后，mAP@0.5 和 Precision 分别提升 0.21%和 1.52%，整体检测性能得到改善。

当多种改进模块组合引入时，模型整体性能进一步提升。其中，同时引入 Backbone-C2fw、GLSA 与 P2-FPN 后，模型参数量减少约 2.29M，mAP@0.5、Precision 与 Recall 分别提升 1.50%、0.97%和 1.25%。在保持较低模型复杂度的前提下，模型在检测精度与目标检出能力方面均表现出较为稳定的提升趋势，表明所提出的各模块在雨天车辆检测任务中具有良好的协同作用与综合有效性。

表 2 消融实验结果

Methods	Backbone-C2fw	GLSA	P2-FPN	mAP50/%	Precision/%	Recall/%	Parameters/10 ⁶
RT-DETR-R18	-	-	-	66.71	78.39	59.17	19.8768
1	√	-	-	67.27	78.07	58.85	13.9013

Methods	Backbone-C2fw	GLSA	P2-FPN	mAP50/%	Precision/%	Recall/%	Parameters/10 ⁶
2	-	√	-	67.05	77.92	60.60	21.9605
3	-	-	√	66.92	79.91	58.43	18.7683
4	√	√	-	68.14	75.70	62.46	16.0177
5	√	-	√	67.36	79.69	60.82	15.2884
6	-	√	√	67.32	76.72	58.87	23.5136
7	√	√	√	68.23	79.36	60.42	17.5831

4.5 对比实验

为验证本文所提出改进模型的检测性能，将其与

Faster R-CNN、SSD 及 YOLO 系列等代表性雨天目标检测方法进行对比实验。所有模型均在相同实验环境与数据集条件下训练，并在统一测试集上进行评估，实验结果如表 3 所示。

表 3 对比实验结果

Methods	mAP50/%	Map50-95/%	Precision/%	Recall/%	Parameters/10 ⁶	(Car) mAP50/%
RT-DETR-R18	66.71	40.52	78.39	59.17	19.8768	78.66
Fast-R-CNN	65.86	39.69	45.73	68.45	28.3121	61.74
SSD	52.73	31.62	83.47	42.14	26.2853	43.69
YOLOv5n	60.99	35.94	78.21	53.29	2.5037	37.00
YOLOv8n	63.13	36.31	75.17	56.25	3.0064	70.18
Yolov10n	59.43	34.91	72.29	51.73	2.2659	68.91
Yolov10s	66.82	41.52	81.02	58.75	7.2195	77.32
Yolov11n	65.73	39.78	76.29	56.82	2.5829	72.17
Yolov12n	60.28	35.51	74.84	52.28	2.5575	67.00
Ours	68.23	40.76	79.36	60.42	17.5831	79.28

由表 3 可知，本文改进模型在 mAP@0.5 指标上达到 68.23%，在所有对比方法中表现最优；其 mAP@0.5:0.95 为 40.76%，与性能较优的 YOLOv10s (41.52%) 接近。同时，改进模型在核心检测目标 (Car) 上的 mAP@0.5 达到 79.28%，优于其他对比方法，表明其在雨天车辆检测任务中具有更稳定的目标识别能力。

在召回率方面，改进模型达到 60.42%，在保持较高检测精度的同时具备良好的目标检出能力；在精确率方面，其 Precision 为 79.36%，综合表现优于多数对比模型。相比参数量较大的两阶段检测方法，本文模型在降低模型规模的同时保持了较高的检测性能；相比 YOLO 系列轻量化模型，本文模型在整体检测精度及车辆类别识别能力方面更具优势。

从模型复杂度角度看，改进模型参数量为 17.58M，较 RT-DETR-R18 基线模型减少约 11.5%，在精度与模型规模之间实现了较为均衡的折中。综上，本文提出的改进模型在雨天车辆检测任务中兼顾检测精度、召回

性能与模型复杂度，具有较好的实际应用潜力。

4.6 可视化分析

为评估本文所提改进方法在实际场景中的应用效果，将 RT-DETR-R18 基线模型与改进模型在雨天车辆检测任务中的输出结果进行可视化对比分析。实验选取白天与夜晚两种典型光照条件下的雨天场景图像，以考察模型在不同光照条件下的检测表现。

图 7 展示了对比结果，其中左侧为原始雨天图像，中间为 RT-DETR-R18 的检测结果，右侧为本文改进模型的检测结果。由图可见，基线模型在雨天场景中仍存在一定的误检与漏检现象，而改进模型能够更准确地识别车辆目标，并在检测置信度方面表现出更为稳定的输出。可视化结果表明，所提出方法在复杂雨天与不同光照条件下具备更好的检测鲁棒性，验证了其实际应用的有效性。



图 7 改进前后模型检测结构可视化对比图

4.7 在 ACDC 数据集上的表现

为验证所提方法的适用性与鲁棒性，本文在专注于恶劣天气场景的 ACDC 数据集[22]的雨天子集上对改进模型进行验证。实验结果表明，改进模型在核心目标 car 上的 mAP@0.5 达到 75.72%，较 RT-DETR-R18 提升 2.33%，具体结果如表 4 所示。该结果表明，所提模型在雨天场景下对车辆目标的分类与定位能力均优于基线模型，在一定程度上缓解了雨水遮挡、路面反光及雨雾干扰等因素对检测性能的影响，验证了模型在不同雨天数据分布下的适用性与鲁棒性。

表 4 ACDC 数据集对比实验结果

Methods	(Car) mAP50/%
RT-DETR-R18	73.39
Ours	75.72

5 结论与展望

针对雨天环境下车辆检测过程中易受雨滴遮挡、对比度下降及远距离目标特征弱化等因素影响而导致的漏检与误检问题，本文围绕雨天场景的成像特性与目标特征退化机理，在 RT-DETR-R18 框架基础上提出了一种面向雨天车辆检测的结构优化策略。首先，将原始主干网络替换为 Conv 与 C2f-wConv 交替堆叠的复合模块，C2f-wConv 通过加权卷积突出关键区域特征并抑制雨滴噪声，从而增强模型在低对比度和模糊条件

下的特征提取能力；其次，在编码器中引入全局-局部空间注意力（GLSA）模块，缓解雨天条件下全局语义与局部细节信息易被破坏的问题；最后，在检测头部分引入 P2 浅层特征分支，并基于深度可分离卷积优化 FPN/PAN 特征融合路径，提升模型对低层细节的感知与表达能力。实验结果表明，所提出的结构优化策略在多源雨天数据集上均取得了稳定的性能提升，验证了其在雨天车辆检测任务中的有效性，并为复杂天气下智能交通系统的安全应用提供了技术支撑。

尽管本文方法在检测精度方面取得了一定改进，但在模型参数量与计算复杂度方面仍存在进一步优化空间。未来工作将从两个方面展开：一是探索更加轻量化的结构设计，以提升模型在实际交通场景中的部署可行性；二是考虑将雨天图像增强方法（如 PReNet）与检测模型进行更紧密的协同设计，通过端到端或级联方式进一步提升复杂雨天环境下的车辆检测鲁棒性。

参考文献

[1] SONG Y F, LU Y F. A review of unmanned visual target detection in adverse weather [J]. Electronics, 2025, 14(13): 2582.

[2] PENNELLY C, REUTER G W, TJANDRA S. Effects of weather on traffic collisions in Edmonton, Canada [J]. Atmosphere-Ocean, 2018, 56(5): 362-371.

- [3] LI X, GENG S. Improved traffic sign detection algorithm for YOLOv5s [C] // International Conference on Computer Engineering and Application. Piscataway: IEEE, 2023: 696-699.
- [4] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [6] LYU Z W, JIN H F, ZHEN T, et al. Small object recognition algorithm of grain pests based on SSD feature fusion [J]. IEEE Access, 2021, 9: 43202-43213.
- [7] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C] // Computer Vision—ECCV 2020: 16th European Conference, Proceedings, Part I. Cham: Springer International Publishing, 2020: 213-229.
- [8] ZHAO Y A, LÜ W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection [C] // 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 16965-16974.
- [9] 胡待方, 全秋红, 柴国庆, 等. 雨天车辆检测的两阶段渐进式图像去雨算法 [J]. 激光与光电子学进展, 2023, 60(22): 2215001.
- [10] 陈婷, 姚大春, 高涛, 等. 基于 PReNet 和 YOLOv4 融合的雨天交通目标检测网络 [J]. 交通运输工程学报, 2022, 22(3): 225-237.
- [11] OGINO Y, SHOJI Y, TOIZUMI T, et al. ERUP-YOLO: enhancing object detection robustness for adverse weather condition by unified image-adaptive processing [EB/OL]. arXiv, 2024.
- [12] JING Z Y, LI S, ZHANG Q W, et al. YOLOv8-STE: enhancing object detection performance under adverse weather conditions with deep learning [J]. Electronics, 2024, 13(24): 5049-5063.
- [13] TABASSUM N, EL-SHARKAWY M. Vehicle detection in adverse weather: a multi-head attention approach with multimodal fusion [J]. Journal of Low Power Electronics and Applications, 2024, 14(2): 23.
- [14] GHARATAPPEH S, SEKEH S, DHIMAN V, et al. Weather-aware object detection transformer for domain adaptation [EB/OL]. arXiv, 2025.
- [15] SALSCHIEDER N O. FeatureNMS: non-maximum suppression by learning feature embeddings [C] // Proceedings of the 2020 25th International Conference on Pattern Recognition. Piscataway: IEEE, 2021: 7848-7854.
- [16] CAMMARASANA S, PATANÈ G. Optimal density functions for weighted convolution in learning models [EB/OL]. arXiv, 2025.
- [17] TANG F L, XU Z X, HUANG Q M, et al. DuAT: dual-aggregation transformer network for medical image segmentation [C] // Pattern Recognition and Computer Vision: 6th Chinese Conference, PRCV 2023, Proceedings, Part III. Cham: Springer, 2023: 343-356.
- [18] ZHOU Y, WEI Y. UAV-DETR: an enhanced RT-DETR architecture for efficient small object detection in UAV imagery [J]. Sensors, 2025, 25(15): 4582.
- [19] HUANG J, LI T. Small object detection by DETR via information augmentation and adaptive feature fusion [C] // Proceedings of the 2024 ACM International Conference on Multimedia Retrieval. New York: ACM, 2024: 39-44.
- [20] YU F, CHEN H, WANG X, et al. BDD100K: a diverse driving dataset for heterogeneous multitask learning [EB/OL]. arXiv, 2020.
- [21] KENK M A, HASSABALLAH M. DAWN: vehicle detection in adverse weather nature dataset [EB/OL]. arXiv, 2020.
- [22] SAKARIDIS C, DAI D, VAN GOOL L. ACDC: The Adverse Conditions Dataset with Correspondences for Semantic Driving Scene Understanding [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 10765-10775.

作者简介

刘阳

1998 年生, 女, 研究生, 主要研究领域为大数据应用与计算机视觉。

E-mail: 1102919549@qq.com

朱洲森

1965 年生, 男, 教授, 主要从事大数据存储与优化、云计算和人工智能的研究。

E-mail: zhuzhousen@sicnu.edu.cn