

面向复杂光照环境的改进CLOCs跨模态目标融合算法

殷紫微

天津职业技术师范大学汽车与交通学院, 天津

收稿日期: 2026年3月26日; 录用日期: 2026年4月23日; 发布日期: 2026年4月30日

摘要

在自动驾驶领域, 复杂光照条件(如弱光、过曝、眩光及空间亮度不均等)会导致视觉特征退化, 从而显著降低目标检测的准确性与可靠性, 尤其对小尺度及部分遮挡目标影响更为突出。针对上述问题, 本文提出了一种基于改进CLOCs的视觉-LiDAR多传感器融合目标检测方法。在统一融合框架下, 联合利用二维图像的语义信息与三维点云的结构信息, 通过对CLOCs算法进行改进, 实现跨模态特征的有效互补, 从而提升复杂场景下的检测性能。同时, 针对光照变化带来的特征不稳定问题, 引入光照鲁棒性优化策略, 以增强模型在低光、眩光及过曝等环境中的适应能力。实验结果表明, 所提方法在mAP@0.5:0.95和mAP@0.5指标上分别达到55.6%和75.3%, 相较于多种主流方法均取得了一定提升。进一步分析表明, 该方法在不同尺度目标上均具有良好的检测性能, 尤其在中小目标检测任务中表现更为优越。本文方法在检测精度、环境适应性及鲁棒性方面均取得了较好的效果, 具有一定的实际应用价值。

关键词

多传感器融合, 目标检测, 复杂光照, CLOCs改进

Improved CLOCs Cross-Modal Target Fusion Algorithm for Complex Lighting Environments

Ziwei Yin

School of Automotive and Transportation Engineering, Tianjin University of Technology and Education, Tianjin

Received: March 26, 2026; accepted: April 23, 2026; published: April 30, 2026

Abstract

In the field of autonomous driving, challenging lighting conditions (such as low light, overexposure,

glare, and spatially non-uniform brightness) can lead to degradation of visual features, thereby significantly reducing the accuracy and reliability of object detection, especially for small-scale and partially occluded targets. To address these issues, this paper proposes a vision-LiDAR multi-sensor fusion object detection method based on an improved CLOCs framework. Within a unified fusion framework, the semantic information from 2D images and the structural information from 3D point clouds are jointly utilized. By improving the CLOCs algorithm, effective cross-modal feature complementarity is achieved, thereby enhancing detection performance in complex scenarios. In addition, to mitigate feature instability caused by illumination variations, an illumination-robust optimization strategy is introduced to improve the model's adaptability under low-light, glare, and overexposed conditions. Experimental results demonstrate that the proposed method achieves 55.6% and 75.3% on the mAP@0.5:0.95 and mAP@0.5 metrics, respectively, outperforming several mainstream methods to varying degrees. Further analysis shows that the method maintains strong detection performance across objects of different scales, with particularly notable advantages in detecting small and medium-sized targets. The proposed method exhibits strong performance in terms of detection accuracy, environmental adaptability, and robustness, indicating its potential for practical applications.

Keywords

Multi-Sensor Fusion, Object Detection, Challenging Lighting, Improved CLOCs

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

近年来, 基于深度学习的目标检测方法在自动驾驶领域取得了显著进展, 通过大规模数据训练能够在常见类别上实现较高的识别精度。然而, 在复杂场景或未知环境中, 其泛化能力仍然有限, 尤其在目标被遮挡或光照条件异常时, 容易出现漏检与误检等问题。这主要是由于视觉检测模型高度依赖图像显著特征, 而遮挡及光照变化会导致关键信息缺失或特征退化, 从而削弱模型的感知能力[1]。常见的复杂光照场景下, 单纯视觉检测无法准确识别的情形, 包括太阳炫光、逆光遮挡、低光图像模糊以及夜间炫光干扰等。在这些情况下, 视觉目标检测模型往往因显著特征被遮挡或退化而产生漏检和定位不稳定的问题, 即使目标在物理空间中存在, 也可能在图像平面上不可观测, 形成光照引起的“视觉遮挡”[2]。相比之下, 激光雷达不受外界光照变化影响, 能够稳定获取目标的三维空间位置与几何结构信息, 为环境提供可靠的距离和形状描述。因此, 将视觉语义信息与激光雷达几何信息进行多模态融合[3], 能够在极端光照和遮挡场景下实现互补增强, 从而构建兼具高查全率与高稳定性的融合感知体系。

基于上述背景, 本文开展了系统研究, 主要贡献如下:

提出了一种面向复杂光照场景的改进 CLOCs 融合框架。在保留原有几何一致性约束的基础上, 对二维与三维候选框配对机制进行优化, 引入视觉失效标记, 在 ONCE 数据集上的实验结果表明, 所提方法在 mAP@0.5:0.95 指标上达到 55.6%, 有效减少无效匹配对融合过程的干扰。

提出了一种光照退化感知机制。通过统计三维候选框对应的有效二维检测支持程度, 从全局角度刻画视觉模态在当前场景下的可靠性, 实现对视觉信息质量的显式建模。

设计了一种自适应加权融合策略。根据光照退化感知系数, 对融合得分与三维检测结果进行动态加权, 在光照良好时充分利用跨模态信息, 在光照退化严重时增强激光雷达主导作用, 从而提升检测结果

的鲁棒性与稳定性。

2. 相关工作

2.1. 多模态目标检测方法概述

随着自动驾驶的发展,单一传感器在复杂环境下的感知局限性逐渐显现,多传感器融合已成为提升目标检测性能的重要方向。根据融合阶段的不同,现有方法可分为早期融合、中期融合和后期融合三类。早期融合在数据层直接整合多模态信息,但对标定精度要求较高;中期融合在特征层进行交互与对齐,在性能与效率之间取得平衡;后期融合基于检测结果进行关联,结构简单但依赖跨模态匹配策略的有效性[4]。

CLOCs 提出了一种典型的后期融合框架,通过构建二维与三维候选框的稀疏关联表示,实现高效的跨模态融合判别。然而,该方法主要依赖局部候选对信息,未考虑复杂光照条件下视觉信息整体可靠性的变化,导致在低光、过曝及强反射等场景中性能受限。因此,本文在 CLOCs 框架基础上进行改进,引入光照退化感知机制,对融合过程进行自适应调节,从而提升复杂光照环境下的检测鲁棒性与精度。

2.2. CLOCs 算法网络结构

不同传感器由于成像机理与数据获取方式的差异,对同一场景的表征存在显著不同。相机能够提供稠密的纹理与语义信息,但易受光照与遮挡影响;激光雷达则能够获取精确的三维几何结构与距离信息,但数据稀疏且语义表达能力有限。尽管两者存在差异,但对同一目标的观测本质上属于不同模态表征,因此应同时满足几何一致性与语义一致性,这为多传感器融合提供了基础[5]。基于此,CLOCs 方法通过建立二维与三维候选框之间的对应关系,利用几何投影将三维检测框映射至图像平面,并与二维检测框进行匹配,从而实现跨模态融合[6]。其中,几何一致性通过相机标定关系建模,可表示为:

$$b_j^{2D} = P \cdot B_j^{3D} \quad (2.1)$$

其中 b_j^{2D} 为 3D 框在图像平面的投影框。几何一致性通常以投影后的 3D 框与二维检测框的交并比(IoU, Intersection over Union)进行量化:

$$IoU_{i,j} = \frac{\text{area}(B_i^{2D} \cap b_j^{2D})}{\text{area}(B_i^{2D} \cup b_j^{2D})} \quad (2.2)$$

在满足几何一致性的前提下,语义一致性用于约束不同模态检测结果在类别层面的匹配。设二维检测器预测类别为 c_i^{2D} , 三维检测器预测类别为 c_j^{3D} , 则语义一致性可表示为:

$$S_{i,j} = \begin{cases} 1, & c_i^{2D} = c_j^{3D} \\ 0, & c_i^{2D} \neq c_j^{3D} \end{cases} \quad (2.3)$$

结合几何一致性与语义一致性,可以构建跨模态匹配的综合判别指标:

$$T_{i,j} = IoU_{i,j} \cdot S_{i,j} \quad (2.4)$$

当 $T_{i,j} \geq \tau$ (综合阈值)时,认为二维和三维候选框为匹配对,可用于后续融合判别。通过这种方式,不仅能够剔除假阳性,还能保证融合结果在位置和类别上的一致性,从而提高跨模态检测的可靠性和精度。

基于几何一致性与语义一致性,CLOCs 将 k 个二维候选框与 n 个三维候选框编码为一个稀疏张量,用于跨模态融合判别。具体表示为:

$$T_{i,j} = (IoU_{i,j}, s_i^{2D}, s_j^{3D}, d_j) \tag{2.5}$$

其中 $IoU_{i,j}$ 表示第 i 个 2D 候选框与第 j 个投影到图像平面的 3D 候选框的交并比； s_i^{2D} 和 s_j^{3D} 分别为对应候选框的置信度分数； d_j 为第 j 个 3D 框至激光雷达传感器的归一化距离。当 IoU 为 0 时，视为不满足几何一致性，对应候选对被忽略；对于仅由激光雷达检测到的目标，则将其 IoU 与二维置信度设为 -1 进行标记。随后，将所有有效候选对编码为稠密张量输入卷积网络进行特征学习。该网络由多层卷积、BatchNorm 与 ReLU 组成，并通过最大池化实现特征聚合。最终，经映射与后处理得到每个三维候选框的融合得分，从而生成更精确的三维检测结果[7]。CLOCs 网络架构图如图 1 所示。

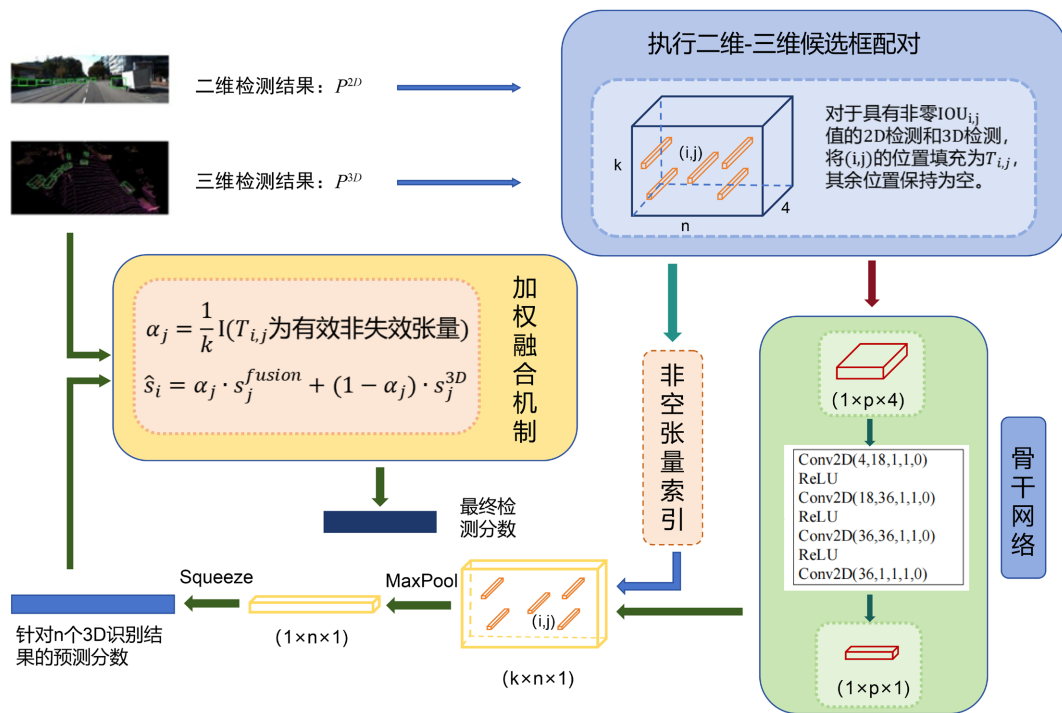


Figure 1. Architecture diagram of the CLOCs network
图 1. CLOCs 网络架构图

3. 所提方法

本文沿用 CLOCs 基于几何投影的一致性约束策略，在二维与三维候选框配对阶段，将三维候选框投影至图像平面，并与二维候选框计算交并比 $IoU_{i,j}$ 。当 $IoU_{i,j} = 0$ 时，认为该候选对不满足几何一致性约束，并将其对应的二维置信度及 IoU 统一设为 -1，以标记视觉信息不可用。由此，候选框对可统一表示为：

$$T_{i,j} = (IoU_{i,j}, s_i^{2D}, s_j^{3D}, d_j) \tag{3.1}$$

其中， s_j^{3D} 表示三维检测置信度， d_j 为目标归一化距离。该编码方式实现了对候选框对的初步筛选，将其划分为有效配对、无效配对及纯三维候选框[7]。

然而，上述方法仅基于局部几何关系进行判别，未考虑复杂光照条件下视觉信息整体可靠性的变化。在复杂光照退化场景(如低光照、过曝及强反射)下，视觉检测性能显著下降，而激光雷达仍能提供稳定的几何信息[8]。此时，若仍采用固定权重融合策略，易引入不可靠视觉信息，进而影响整体检测性能。

为此，本文在融合决策阶段引入光照退化感知机制，从全局角度评估视觉模态的有效性[9]。通过统计与同一三维候选框满足几何一致性且置信度较高的二维候选框比例，刻画视觉信息对该目标的支持程度，从而为后续自适应融合提供依据。

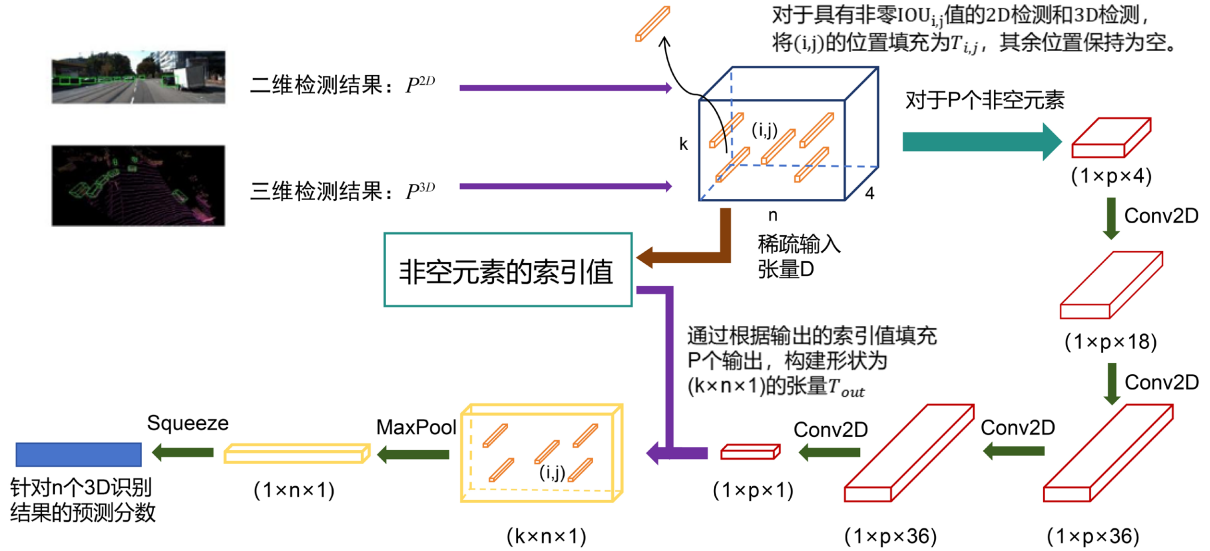


Figure 2. Architecture diagram of the CLOCs network
图 2. 改进的 CLOCs 网络架构图

具体而言，对于第 j 个三维候选框，在所有二维候选框中统计与其满足几何一致性且置信度高于阈值的有效配对数量，并计算其占比，用以刻画视觉模态对该目标的支持程度。基于此，定义光照退化感知系数为：

$$\alpha_j = \frac{1}{k} \sum_{i=1}^k I(T_{i,j} \text{ 为有效非失效张量}) \quad (3.2)$$

其中， k 为二维候选框总数， $I(\cdot)$ 为指示函数。当光照条件不利、二维信息大范围失效时， α_j 取值较小，表明视觉信息对该目标的支持程度有限。

在最终预测阶段，本文采用加权融合方式，将 CLOCs 网络输出的融合得分 s_j^{fusion} 与原始三维检测器得分 s_j^{3D} 进行组合：

$$\hat{s}_j = \alpha_j \cdot s_j^{fusion} + (1 - \alpha_j) \cdot s_j^{3D} \quad (3.3)$$

当视觉信息可靠时，融合结果充分利用多模态信息；当视觉退化严重时，模型自动降低视觉分支权重，使检测结果更多依赖激光雷达，从而保证在复杂光照条件下的稳定性。

综上，本文在 CLOCs 框架基础上，通过在候选框配对阶段引入视觉失效标记，并在融合阶段设计光照退化感知的自适应加权机制，实现了对视觉退化与真实几何不一致的有效区分，如图 2 所示。在良好光照条件下提升融合精度，在复杂光照环境下增强检测鲁棒性。

4. 实验结果及分析

4.1. 融合模块数据集构建

为验证所提激光雷达感知算法在真实道路环境下的适用性与鲁棒性，本文基于 ONCE 数据集开展实

验评估。该数据集涵盖白天与夜晚、晴天与雨天以及城市与郊区等多种场景，提供了丰富的点云与图像标注，能够有效支撑多模态感知算法的训练与性能验证。

在此基础上，本文针对复杂光照条件对原始数据进行筛选，构建了包含三类典型光照异常的测试子集：低光、炫光与过曝[10]。其中，低光场景表现为整体亮度不足导致目标可见性下降；炫光场景存在局部强光干扰；过曝场景则导致图像细节丢失。这些场景均是视觉感知易失效的关键情形，有助于评估所提方法在复杂环境下的鲁棒性与适应能力。

在数据集构建阶段，先基于定量图像统计方法对原始样本进行初步筛选，以识别低光照及非理想曝光图像。具体而言，综合分析图像的全局亮度分布[11]、颜色特征以及直方图相关统计度量，对潜在异常光照样本进行候选提取。

针对低光照场景，本文通过计算图像全局灰度均值 μ ，对 $\mu < 70$ 的样本进行初筛，并结合亮度方差对局部阴影区域进行校验，以避免短时遮挡或局部暗区造成的误判。对于炫光与过曝场景，则基于亮度方差(>2500)或高亮饱和像素比例($>2\%$)进行筛选，从而覆盖强光干扰下的典型光照情况。

需要指出的是，本文对光照异常的划分并非基于严格的成像物理模型，而是面向目标检测任务，从光照对目标可见性与检测性能的影响出发，采用基于亮度统计与空间分布特征的判别方法进行定量划分，以保证筛选过程的一致性与可复现性。

在样本筛选后，所有数据均进行统一标注校准，包括：边界框优化以提升轮廓精度、类别标签一致性校验，以及按 BDD100K 标准规范化标注格式。同时，通过定期抽样(每 200 帧抽检 20 帧)进行人工复核，对误分类样本进行修正与重新分配。最终，本文采用“自动筛选 + 人工复核”的两阶段流程构建极端光照测试子集，如图 3 所示，有效保证了数据的代表性与可靠性，为后续算法评估提供了坚实基础。



(a) 炫光遮挡场景



(b) 低光遮挡场景



(c) 炫光遮挡场景

Figure 3. Lighting test subset samples

图 3. 光照测试子集样本

为保证数据质量与一致性,构建过程中确保每幅图像仅归属于单一场景,并完整保留原始类别分布,从而提升标注一致性与数据可靠性。基于上述策略,最终构建的复杂光照子数据集能够真实反映光照退化特性,并具备良好的实验可用性[12]。该子数据集共包含 2500 帧图像,其中低光、炫光和过曝场景分别为 1020 帧、720 帧和 760 帧,整体分布相对均衡。在保留原始数据多样性的基础上,对典型光照异常场景进行针对性筛选,使数据能够充分覆盖视觉退化情形。同时,适中的数据规模既满足定量评估与可视化分析需求,又保证了人工复核的可行性。该数据集可有效揭示视觉方法在复杂光照条件下的性能局限,并为多模态融合方法的鲁棒性验证提供可靠支撑。

4.2. 实验环境

本研究在统一的实验环境与评价标准下开展。多模态输入包括 RGB 图像与激光雷达点云,视觉分支采用改进的 RT-DETR 模型,点云分支基于端到端点云特征提取网络,并在特征层进行融合以输出最终检测结果。实验基于 ONCE 数据集进行训练,以两分支生成的候选框作为输入。训练过程中,迭代次数设为 5000, epoch 为 80,批大小为 64,初始学习率为 3×10^{-3} ,并每 10 个 epoch 衰减 10 倍。所有实验均在 NVIDIA RTX 4090 GPU 上运行,基于 PyTorch 框架实现。软件环境为 Ubuntu 18.04 (64-bit), CUDA 11.7, Python 3.7。上述设置保证了实验结果的可复现性与可比性。

4.3. 评价指标分析

为验证所提融合感知方法在复杂光照与遮挡条件下的有效性与鲁棒性,本文选取具有代表性的单模态与多模态目标检测算法作为对比方法。单模态方法包括基于激光雷达的 PointPillars 和 CenterPoint,以及基于视觉的 BEVFormer;多模态方法选取 BEVFusion 作为对比基线[13]。上述方法覆盖不同传感信息处理范式,可为性能评估提供可靠参考。

(1) 改进模型性能评估

实验采用 mAP@0.5 和 mAP@0.5:0.95 作为主要评价指标,分别衡量模型在宽松与严格 IoU 条件下的检测能力。同时,引入大、中、小目标检测精度(API、APm、APs),以评估模型在多尺度场景下的适应性。

表 1 给出了各方法在测试集上的定量结果。可以看出,本文方法在整体检测性能及不同尺度目标上均取得了较优表现。

Table 1. Average precision table of different models

表 1. 不同模型的平均精度表

模型	mAP@0.5:0.95/%	API/%	APm/%	APs/%	mAP0.5/%
PointPillars	48.3	83.5	60.9	42.1	69.8
CenterPoint	50.1	84.1	61.2	42.7	70.3
BEVFormer	31.4	74.2	40.8	30.1	42.4
BEVfusion	48.5	86.7	62.5	43.9	72.7
CLOCs	52.5	87.4	63.8	47.2	73.6
Ours	55.6	89.6	68.9	50.3	75.3

在整体检测性能方面,本文方法在 mAP@0.5:0.95 上达到 55.6%,较 PointPillars (48.3%)、CenterPoint (50.1%)、BEVFusion (48.5%)和 CLOCs (52.5%)分别提升 7.3%、5.5%、7.1%和 3.1%;在 mAP@0.5 上取

得 75.3% 的最高结果，分别提升 5.5%、5.0%、2.6% 和 1.7%。两项指标的同步提升表明，所提方法在降低漏检率的同时，能够有效提升目标定位精度与检测稳定性。

从不同尺度目标来看，本文方法在大、中、小目标上均表现出稳定优势。其中，大目标(API)达到 89.6%，在结构信息完整的场景中具有更高的检测可靠性；中等目标(APm)为 78.9%，体现出较强的中距离检测能力；小目标(APs)达到 50.3%，相较各对比方法均有明显提升，表明所提方法在复杂光照条件下能够有效缓解小目标易漏检问题。

综上，本文方法在保证整体检测精度的同时，实现了对不同尺度目标的稳定适应，在复杂光照场景下表现出更强的鲁棒性与泛化能力。

(2) 平均误差分析

为量化不同算法在复杂光照及遮挡条件下的目标检测性能表现，本文选取了 mATE (目标位置预测精度)、mASE (目标尺度预测误差) 以及 mAOE (目标朝向预测误差) 等关键评价指标，对模型在三维目标检测任务中的预测精度进行定量分析 [14] [15]。其中，mATE 主要用于衡量预测目标与真实目标在空间位置上的偏差，反映模型的定位精度；mASE 用于评估目标尺度预测的准确性，刻画检测框尺寸回归的稳定程度；mAOE 则用于衡量目标朝向预测误差，反映模型对目标姿态信息的建模能力。实验结果如表 2 所示。

Table 2. Analysis of mean error

表 2. 平均误差分析

方法	PointPillars	CenterPoint	BEVFormer	BEVfusion	Ours
模态	激光雷达	激光雷达	摄像头	融合	融合
mATE	0.667	0.384	0.452	0.457	0.397
mASE	0.351	0.434	0.313	0.308	0.314
mAOE	0.507	0.434	0.601	0.401	0.418
mAP@0.5	0.698	0.703	0.424	0.727	0.753
mAP@0.5:0.95	0.483	0.501	0.314	0.485	0.556

从目标位置预测精度(mATE)来看，CenterPoint 取得最低误差(0.384)，本文方法为 0.397，略高于 CenterPoint，但明显优于 PointPillars (0.667) 以及 BEVFormer (0.452) 和 BEVFusion (0.457)。这表明在复杂光照条件下，本文方法在融合视觉信息的同时，能够有效保持激光雷达的高精度定位能力。

在目标尺度误差(mASE)方面，本文方法为 0.314，与 BEVFormer (0.313) 和 BEVFusion (0.308) 接近，显著优于 PointPillars (0.351) 与 CenterPoint (0.434)，说明所提方法在多模态融合过程中提升了目标尺寸回归的稳定性。

在目标朝向误差(mAOE)方面，本文方法为 0.418，优于 PointPillars (0.507) 与 BEVFormer (0.601)，与 BEVFusion (0.401) 和 CenterPoint 处于相近水平，表明引入激光雷达几何信息有助于提升复杂光照条件下的朝向预测能力。

综合来看，本文方法在 mATE、mASE 和 mAOE 指标上均取得较优表现，在保证定位精度的同时提升了尺度与方向预测的稳定性，进一步验证了所提融合策略的有效性。

4.4. 可视化结果分析

本研究基于 ONCE 数据集开展实验，其中包括低光场景、炫光场景以及曝光不均场景，三类场景均

覆盖了城市道路中具有代表性的典型交通要素。这些场景在光照分布、目标清晰度及背景复杂度方面具有高度真实性，对视觉感知与融合算法均构成较大挑战。

在三种复杂光照条件下，本文选取 LA-DETR 纯视觉检测算法、BEVFusion 融合算法以及本文提出的多模态融合方法进行对比分析。实验结果展示了各方法在同一帧图像上的检测输出，其中 LA-DETR 仅依赖视觉信息进行目标检测，而 BEVFusion 与本文方法均融合了视觉与激光雷达信息，用以对比不同融合策略在复杂光照条件下的检测性能差异，其场景包括炫光、炫光遮挡、低光、低光遮挡以及曝光不均和曝光不均遮挡几类情况。可视化结果如图 4 至图 5 所示。



Figure 4. BEVFusion target detection results in complex lighting scenarios

图 4. BEVFusion 复杂光照场景目标检测结果

从图 4 与图 5 可以看出，在远距离小目标场景(a)中，BEVFusion 存在明显漏检，而本文方法能够稳定检出目标，表现出更强的鲁棒性。在炫光遮挡场景(b)下，BEVFusion 虽能检测主要目标，但检测置信度存在波动；相比之下，本文方法在相同目标上给出更高且更稳定的置信度(如远处车辆由 0.67 提升至

0.70), 表明在视觉受干扰时, 融合策略能够有效提升目标判别可靠性。在一般复杂场景(c)中, 本文方法同样在检测置信度上优于对比方法(如左侧车辆提升至 0.79), 进一步体现了融合策略的稳定性优势。在低光与阴影遮挡场景(d)以及曝光不均场景(e)中, BEVFusion 的检测结果存在一定波动, 而本文方法仍能保持较稳定的检测性能, 说明引入激光雷达信息有助于缓解光照退化带来的影响。



Figure 5. The proposed method achieves accurate object detection results in complex lighting scenarios

图 5. 本文所提方法在复杂光照场景目标检测结果

综上所述, 在炫光、低光及曝光不均等复杂光照条件下, 本文方法在漏检率、检测置信度及结果稳定性方面均优于对比方法, 验证了所提融合策略在复杂环境下的有效性。

5. 结论与局限性

本文针对单一视觉与激光雷达在复杂道路场景中存在的感知局限性, 提出了一种多传感器融合目标检测方法。通过分析视觉在光照变化与遮挡场景下的不稳定性, 以及激光雷达在语义表达方面的不足,

构建了视觉-LiDAR 协同的融合思路,实现了不同模态信息的优势互补。在方法设计上,基于改进的 CLOCs 融合框架,联合利用二维图像的语义信息与三维点云的空间结构信息,在候选框层面进行有效融合,从而提升复杂场景中的目标检测精度与查全率。实验结果表明,在低光、炫光及过曝等典型复杂光照条件下,所提出方法相比单一传感器检测方法具有更好的稳定性与鲁棒性,能够有效缓解视觉退化带来的性能下降问题。

尽管本文方法在多种复杂场景中取得了较好的检测效果,但仍存在一定局限性。一方面,融合方法对二维与三维检测结果的质量依赖较强,当任一模态检测性能明显下降时,整体融合效果仍可能受到影响;另一方面,融合过程引入了一定的计算开销,在实时性方面仍有优化空间。此外,在极端光照退化或点云极度稀疏的情况下,目标信息不足仍可能影响最终检测性能。未来可从自适应融合策略、弱模态补偿机制以及轻量化模型设计等方向进一步改进。

参考文献

- [1] 秦晨曦,于洋. 基于卷积神经网络的遮挡目标图像检测识别技术研究[J]. 电脑编程技巧与维护, 2023(2): 153-155.
- [2] 李亚南. 自动驾驶环境感知技术研究现状及发展方向[J]. 汽车与新动力, 2025, 8(5): 1-5.
- [3] Hu, R., Liu, Y., Gu, K., Min, X. and Zhai, G. (2023) Toward a No-Reference Quality Metric for Camera-Captured Images. *IEEE Transactions on Cybernetics*, **53**, 3651-3664. <https://doi.org/10.1109/tcyb.2021.3128023>
- [4] Shi, S., Guo, C., Jiang, L., Wang, Z., Shi, J., Wang, X., et al. (2020) PV-RCNN: Point-Voxel Feature Set Abstraction for 3D Object Detection. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 10526-10535. <https://doi.org/10.1109/cvpr42600.2020.01054>
- [5] Chen, X., Ma, H., Wan, J., Li, B. and Xia, T. (2017) Multi-View 3D Object Detection Network for Autonomous Driving. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 1907-1915. <https://doi.org/10.1109/cvpr.2017.691>
- [6] 杨壮涛,朱纪洪,刘兆沛. 浅议无人车探测多通道数据同步采集系统[J]. 中国设备工程, 2022(12): 29-32.
- [7] 张亚楠,张浩芳. 电子式电流互感器在线校验技术研究[J]. 电子技术与软件工程, 2015(24): 250.
- [8] 王宝树,李芳社. 基于数据融合技术的多目标跟踪算法研究[J]. 西安电子科技大学学报, 1998(3): 5-8.
- [9] 杨露菁. 多源信息融合理论与应用[M]. 北京: 北京邮电大学出版社, 2011.
- [10] 鲁斌,孙洋,杨振宇. 融合体素图注意力的三维目标检测算法[J]. 智能系统学报, 2024, 19(3): 598-609.
- [11] 郑宝峰,张朝阳,张向清,等. 交通场景中的相机标定方法研究[J]. 科技风, 2018(11): 11.
- [12] 蔡先航,陈胜. 结合空间一致性的高密度口内点云配准方法[J]. 控制工程, 2025, 32(11): 1979-1987.
- [13] 袁善帅,丁雷. 基于多阶段信息增强的 3D 点云目标检测算法[J]. 激光与光电子学进展, 2024, 61(4): 290-296.
- [14] 冯文卿,张永军. 利用多尺度融合进行面向对象的遥感影像变化检测[J]. 测绘学报, 2015, 44(10): 1142-1151.
- [15] 薛培林. 基于激光雷达与视觉融合的城市自主车辆实时感知技术研究[D]: [硕士学位论文]. 南京: 东南大学, 2020.