

面向夜间场景的激光雷达与相机融合的三维目标检测算法

陈泽彬

广东工业大学自动化学院, 广东 广州

收稿日期: 2025年6月20日; 录用日期: 2025年7月21日; 发布日期: 2025年7月31日

摘要

基于激光雷达与相机融合的目标检测算法成功提高了自动驾驶系统的感知性能。然而, 以往的多模态算法都是面向白天场景进行检测, 在夜间场景下会因为夜间环境亮度较低以及由灯光照射引起的过曝问题导致检测性能的下降; 为此, 本文研究了一种改进的夜间多模态检测算法。首先引入了基于统计的过曝区域像素掩码模块, 通过分析像素饱和度特征来抑制过曝区域的干扰; 其次引入了基于权重归一化的通道加权融合模块, 采用逐通道动态权重分配机制来优化不同模态特征的融合。实验采用NuScenes数据集全集及其夜间子集验证模型性能, 实验结果表明改进模型平均类别精度达74.5%, 对比基线方法实现了性能上的提升, 算法可靠性得到验证, 该研究成果展现出较好的工程价值。

关键词

三维目标检测, 夜间场景, 多模态融合, 深度学习

A 3D Object Aetection Algorithm Based on LiDAR and Camera Fusion for Night Scenes

Zebin Chen

School of Automation, Guangdong University of Technology, Guangzhou Guangdong

Received: Jun. 20th, 2025; accepted: Jul. 21st, 2025; published: Jul. 31st, 2025

Abstract

The target detection algorithm based on the fusion of LiDAR and camera successfully improved the perception performance of the autonomous driving system. However, previous multimodal algorithms are all for detection in daytime scenes. In nighttime scenes, the detection performance will decrease due to the low brightness of the nighttime environment and the overexposure problem

文章引用: 陈泽彬. 面向夜间场景的激光雷达与相机融合的三维目标检测算法[J]. 人工智能与机器人研究, 2025, 14(4): 1025-1033. DOI: 10.12677/airr.2025.144097

caused by lighting. Therefore, this paper studies an improved nighttime multimodal detection algorithm. First, a statistical overexposed area pixel mask module is introduced to suppress the interference of the overexposed area by analyzing the pixel saturation characteristics; secondly, a channel weighted fusion module based on weight normalization is introduced, and a channel-by-channel dynamic weight allocation mechanism is used to optimize the fusion of different modal features. The experiment uses the full set of NuScenes dataset and its nighttime subset to verify the model performance. The experimental results show that the average category accuracy of the improved model reaches 74.5%, which is a performance improvement compared with the baseline method. The reliability of the algorithm is verified, and the research results show good engineering value.

Keywords

3D Object Detection, Night Scene, Multimodal Fusion, Deep Learning

Copyright © 2025 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

交通安全问题一直是困扰现代社会的重大挑战。统计数据显示，全球每年因人为操作失误导致的交通事故造成了超过百万人员伤亡，更带来难以估量的经济损失[1]。自动驾驶指的是在没有人类参与的环境下实现车辆自主行驶，其核心价值在于通过车辆进行自主决策和利用车联网协同从根本上消除人为因素引发的交通事故[2]。

自动驾驶系统的感知单元作为其核心子系统，相机与激光雷达是感知单元最常用的两种传感器。基于激光雷达与相机融合的多模态三维目标检测算法感知性能明显优于基于一个传感器的检测算法，且已经在多个基准数据集上证明了其优异的性能[3]。如 PointPainting [4]通过将 2D 图像的语义分割信息投影到激光雷达点云上，增强原始点云的特征表达。以极低的计算开销换取较大的性能提升。BEVFusion [5]通过将不同的数据统一地转换到鸟瞰图(Bird's Eye View, BEV)视角下进行融合，在 NuScenes 数据集上展现出了优异的性能。

然而，现有的多模态检测算法主要针对日间场景进行设计，在夜间环境下相机特征的劣质化会影响算法整体的性能。首先，由于夜间光照不均匀的特性，图像在暗部区域特征不足，这些区域在特征融合过程中会产生特征稀释效应，从而降低检测精度[6]。其次，夜间复杂的光照环境容易造成局部过曝现象，导致图像特征严重失真，这种失真特征经过网络处理后会产生异常响应，进一步干扰检测结果的准确性。这些问题的存在影响了算法在夜间场景下的性能发挥[7]。

图像增强和去噪方法常被用于改善夜间检测性能。典型的低光图像增强方法包括 Zero-DCE [8]、EnlightenGAN [9]等，这些方法能够显著提升图像亮度与细节表达。然而，将这些图像增强方法直接应用于多模态三维检测任务时，可能存在增强后的图像与点云信息不一致的问题，导致融合特征失真。因此，如何在三维检测框架中合理引入图像增强或去噪策略，仍是一个亟需解决的重要研究方向。此外，一些方法也通过调整特征融合来优化夜间场景下的检测性能，ContextualFusion [7]通过引入上下文感知的特征融合模块，有效增强多模态信息之间的互补性，显著提升在夜间场景下的三维目标检测性能。然而，其上下文感知模块通过仅仅是通过对场景的简单分辨对多模态融合过程进行优化，其对夜间场景下的优化是静态的，不能很好地应对夜间复杂多变的环境。

针对上述挑战, 本文研究了一种面向夜间场景的多模态目标检测算法, 旨在解决现有方法在夜间条件下的性能下降问题。首先通过引入基于统计的过曝区域像素掩码(Statistical-Based Pixel Masking Module, SPM)模块, 有效消除过曝区域的干扰, 从而提升特征表达的鲁棒性。其次引入基于动态加权归一化的融合(Channel-Wise Normalized Weighting, CNW)模块, 通过自适应地评估各通道特征的可靠性, 对无效特征进行动态抑制, 从而优化多模态特征的融合质量。

2. 方法

2.1. 基于统计的过曝区域像素掩码

夜间车辆在行驶过程中获取的图像经常会因为对向车辆的灯光照射, 或者反射物对光线的反射而产生局部过曝区域, 这些区域的图像特征信息保存过少甚至完全丢失, 过曝区域的存在会导致模型在特征提取的过程中出现异常响应。而激光雷达通过主动发射激光脉冲获取距离信息, 在不同环境光照条件下均能保持稳定的测量精度。若直接将这种存在质量差异的不同模态特征进行简单融合, 会降低融合特征的质量, 严重影响自动驾驶系统在夜间场景下的决策可靠性。核心挑战在于如何有效消除过曝区域对多模态特征融合的负面影响。掩码策略的优势在于计算高效且实现简单, 既不会为模型引入过多的参数, 增加额外的计算量, 又不需要大量的数据专门对其进行训练。这一思路在 MonoDepth2 [10]中已得到验证。

为了消除过曝区域对算法检测性能的影响, 本文引入了 SPM [11]模块, 其核心是通过统计图像逐像素亮度值, 得到异常区域与正常区域的像素值边界 a 与 b , 后通过生成逐像素掩码图 M_{uc} 实现对过曝区域的隔离。掩码图的生成公式如(1)所示:

$$M_{uc} = \begin{cases} \frac{1}{1 + p^2(x-a)^2}, & x_{\min} \leq x \leq a \\ 1, & a \leq x \leq b \\ \frac{1}{1 + q^2(x-b)^2}, & b < x \leq x_{\max} \end{cases} \quad (1)$$

其中, x 代表的是图像像素的像素值, p 与 q 为衰减系数, 用于控制 SPM 模块对过曝区域的抑制程度。在生成掩码图后, 对过曝区域的遮蔽如公式(2)所示:

$$img_mask = M_{uc} * img \quad (2)$$

其中, img_mask 表示掩码后图像, img 表示原图像, “*” 表示逐元素乘法。掩码的原理如图 1 所示, 其类似一个桥型函数, 光照强度处于 a 与 b 之间的特征信息被完整保留下来, 过曝部分的信息则被有效遮蔽。

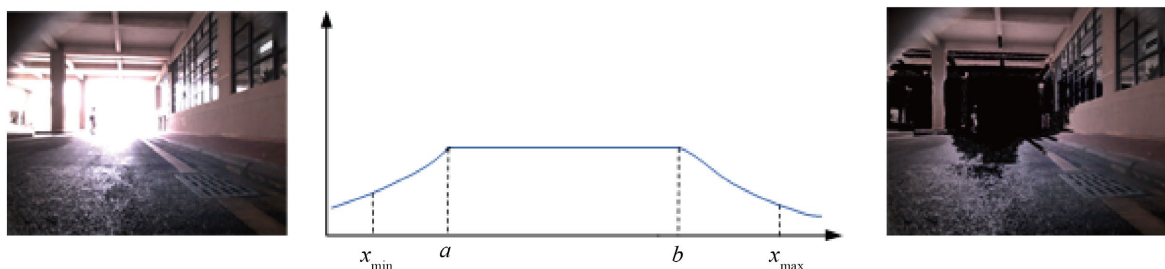


Figure 1. Principle of SPM module

图 1. SPM 模块原理

2.2. 基于权重归一化的通道加权融合

在夜间或低光照条件下，相机传感器捕获的图像数据会出现图像有效特征减少、噪声增加等问题，而激光雷达在夜间环境中也能持续获取高质量的环境特征。直接融合低质量的图像特征与高质量的激光雷达特征降低融合后的特征的质量，进而直接影响算法的检测性能。为了在传感器数据质量下降时，保持融合特征的稳定性，本文引入了 CNW [12] 模块，该模块通过为来自不同传感器的特征学习通道级的权重向量，抑制了噪声或者减少特征通道对融合后特征的影响，提高了融合后特征的质量。CNW 的原理如图 2 所示，该模块首先将相机特征 F_C^{BEV} 与激光雷达特征 F_L^{BEV} 通过权重生成网络，并将输出进行归一化，得到相机逐通道权重 A_C 与激光雷达逐通道权重 A_L ，而后执行最终的特征融合。特征融合的计算如公式 (3) 所示：

$$F_{fuse}^{BEV} = CNW(F_C^{BEV}, F_L^{BEV}) = F_C^{BEV} \odot A_C + F_L^{BEV} \odot A_L \quad (3)$$

其中， F_{fuse}^{BEV} 代表融合后的特征，“ $\odot$ ”代表逐通道乘法，“+”代表特征之间的融合，本文中采用卷积神经网络的方式实现两种特征的融合。该方法能确保在特征融合过程中，信息丰富的通道获得更高权重，而噪声或信息缺失的通道被自动抑制，从而提升融合特征的质量。

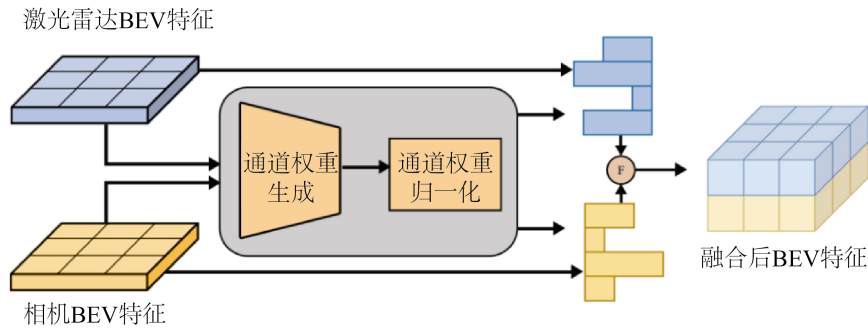


Figure 2. CNW module
图 2. CNW 模块

2.3. 网络模型

本文算法基于 BEVFusion 架构进行改进，BEVFusion 的核心创新在于将不同传感器采集的数据统一映射到 BEV 视角下进行融合，减少了融合时信息损失的同时显著提升算法检测性能。本文针对夜间场景对模型进行改进，通过引入 SPM 模块与 CNW 模块优化模型在夜间的检测性能，改进后的网络结构如图 3 所示。

具体而言，针对相机分支，模型首先使用 SPM 模块遮蔽输入图像的过曝区域，而后使用 ResNet [13] 网络与特征金字塔网络 (Feature Pyramid Network, FPN) 模块提取图像的多尺度特征并进行融合，提升图像特征的表征能力，之后通过基于深度估计的视图变换模块，将提取到的图像特征从图像视角转换到 BEV 视角下，得到图像的 BEV 特征。对于雷达分支，我们参考 VoxelNet [14]，采用基于体素的特征提取网络提取雷达特征，并通过 Z 轴展平操作将其映射至 BEV 视角，得到激光雷达的 BEV 特征。当获得不同模态的 BEV 视角特征后，我们引入基于权重归一化的融合模块 CNW，将不同模态的 BEV 特征进行融合。融合后的特征进一步传递至网络的检测头，该检测头参考 TransFusion [15]，采用了一种基于类别特定中心热力图的检测头结构，融合后的特征经过检测头进行检测，最终得到感兴趣对象的 3D 边界框的中心位置、范围等信息。

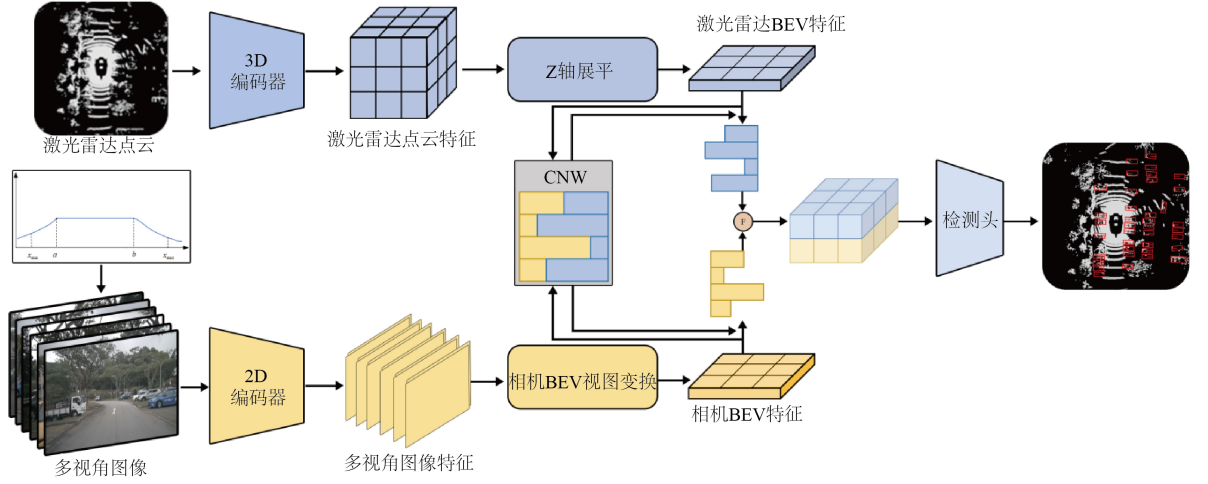


Figure 3. Overall framework diagram of the model
图 3. 模型整体框架图

本算法的损失函数 L_{total} 由三部分组成，具体计算方式如公式(4)所示：

$$L_{total} = L_{cls} + L_{heatmap} + L_{bbox} \quad (4)$$

其中， L_{cls} 为分类损失， L_{cls} 采用的是 FocalLoss 函数进行计算，FocalLoss 计算公式如(5)所示：

$$L_{cls}(p_t) = \alpha_t (1 - p_t)^\gamma \log(p_t) \quad (5)$$

其中， α_t 表示类别权重，用于平衡正负样本的损失贡献， γ 是聚焦参数，用于控制难分类样本的权重， p_t 为模型预测的目标类别概率， p_t 的计算公式如(6)所示：

$$p_t = \begin{cases} p, & \text{if } y = 1 \\ 1 - p, & \text{otherwise} \end{cases} \quad (6)$$

其中，其中 p 为模型实际输出的概率， y 为真实标签， y 为 1 表示正样本，0 则表示负样本。

热力图损失 $L_{heatmap}$ 使用 GaussianFocalLoss 作为损失函数，GaussianFocalLoss 的计算如公式(7)所示：

$$L_{heatmap}(p_t) = -\alpha_t (1 - p_t)^\gamma e^{-\beta d^2} \log(p_t) \quad (7)$$

其中， α_t 表示类别权重， γ 是聚焦参数， p_t 为模型预测的目标类别概率。 β 是高斯权重参数，控制高斯分布的衰减速度， d 是样本点到目标中心点的归一化距离。

损失函数的最后一部分回归损失 L_{bbox} 使用 L1Loss 函数进行计算，L1Loss 计算公式如(8)所示：

$$L_{bbox}(y, g) = \frac{1}{N} \sum_{i=1}^N |y_i - g_i| \quad (8)$$

其中， y_i 是真实值， g_i 表示模型预测值， N 表示样本数量。

3. 模型训练与评估

3.1. 数据集

在数据集选择方面，我们采用了 NuScenes [16]数据集及其夜间场景子集。该数据集是由 Aptiv 公司于 2019 年发布的一个大规模、多模态自动驾驶数据集。其具有多种传感器，在文中我们选择一个激光雷达以及六个摄像头数据作为算法的输入。具体而言，全集中我们选择了 700 个场景用于训练，150 个场景

用于测试；在夜间子集中，我们从 NuScenes 数据集中精心挑选了 99 个夜间场景，其中 72 个场景用于训练，27 个场景用于测试。将其命名为 Night NuScenes。该子集专门用于测试模型在夜间环境下的性能表现，从而更准确地衡量模型在夜间场景下的检测能力。在 NuScenes 全集中，我们选择了 10 个类别进行目标检测，这些类别包括轿车(car)、卡车(truck)、工程车辆(construction_vehicle)、公交车(bus)、拖车(trailer)、路障(barrier)、摩托车(motocycle)、自行车(bicycle)、行人(pedestrian)以及交通锥(traffic_cone)。这些类别涵盖了多种常见交通参与者和障碍物，能够全面评估模型在复杂场景下的检测性能。然而，在夜间场景中，由于部分类别对象的出现频率较低，出于样本均衡考虑，我们对检测类别进行了调整。因此，在 Night NuScenes 子集中，我们专注于检测卡车、轿车、行人、摩托车、公交车和拖车这六个类别对象。这些类别在夜间场景中具有较高的代表性，能够有效反映模型在低光照条件下的检测能力。

3.2. 评价指标

本研究采用目标检测领域广泛使用的平均精度均值(mean Average Precision, mAP)作为核心评价指标。它通过计算所有类别的平均精度(Average Precision, AP)的均值获得。AP 则是通过对精确率 - 召回率曲线进行积分获得。精确率与召回率是计算各个类别的平均精度的重要参数，二者的计算公式如(9)所示：

$$\begin{aligned} Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \end{aligned} \quad (9)$$

Precision 表示精确率；*Recall* 表示召回率；*TP* 为真阳性，表示模型正确识别的正样本数量；*FP* 为假阳性，指模型将负样本误判为正样本的数量；*FN* 为假阴性，指模型未能识别的正样本数量。基于精确率和召回率的量化分析，可以绘制精确率 - 召回率曲线，进而计算各个类别的 *AP*。*mAP* 的计算公式如(10)所示：

$$mAP = \frac{1}{M} \sum_{i=1}^M AP_i \quad (10)$$

其中， AP_i 表示第 i 个类别的平均精确率， M 是类别数量。

3.3. 实验环境

本文的实验训练环境配置如表 1 所示。

Table 1. Experimental environment

表 1. 实验环境

硬件环境		软件环境	
CPU	Intel Xeon Gold 6133 CPU	操作系统	Ubuntu 20.04.1 LTS
GPU	NVIDIA RTX4090D	Cuda 版本	11.7
内存	32G	Pytorch 版本	1.10

3.4. 模型超参数设置

在模型训练之前，我们对关键超参数进行了系统性配置以优化模型性能。点云截取范围方面，统一设置点云的有效范围为(-51.2 m, -51.2 m, -5.0 m, 51.2 m, 51.2 m, 3.0 m)；点云体素化时的体素大小统一设置为(0.1 m, 0.1 m, 0.2 m)；优化器选择 AdamW 优化器并设置初始学习率为 $2e-4$ ，并实施同时实施学习率衰减策略，学习率随训练过程均匀衰减；批处理大小(batch size)设置为 8，模型训练轮次(epochs)设定

为 20 轮。本文算法实现均基于 MMDetection3D 工具箱进行搭建。

3.5. 模型评估

为了深入评估改进后算法的性能，我们将其与当前主流的多模态目标检测算法 BEVFusion [5]、UniBEV [12]和 ContextualFusion [7]进行了全面的对比实验。实验分别在 NuScenes 数据集的完整测试集以及专门划分的夜间子集 Night-NuScenes 上展开。NuScenes 数据集全集主要聚焦于白天场景下的目标检测，而 Night-NuScenes 则专注于夜间场景，能够精准地衡量算法在夜间场景下的表现能力。实验过程使用各个类别的精度 AP 与总的平均精度 mAP 作为评价指标。所有的算法均基于表 1 所展示硬软件环境进行重新训练，在 NuScenes 全集上的训练过程由于随机种子设置以及训练硬件均与各算法原文所用不同，故而实验结果与原文存在一定差异。

具体各个算法在不同数据集上的 mAP 如表 2 所示。实验结果表明，改进后的算法在不同的数据集上均展现出优越性。在 NuScenes 全集的测试中，改进后的算法与原算法 BEVFusion 的 mAP 稍有降低，但是总体而言，引入 SPM 模块和 CNW 模块并不会过分削弱模型在常规场景下的检测性能。相反，即便在融合了这些新模块之后，模型在常规场景中的检测性能依旧稳定。在 Night-NuScenes 数据集上，本文的算法与其他算法相比在夜间场景下的检测精度得到提高，达到了 74.5%。改进模型取得了最优表现。这一结果证明了本文引入的 SPM 模块和 CNW 模块对于提升夜间场景检测性能的有效性。

Table 2. mAP of each model on two datasets

表 2. 各模型在两个数据集上的 mAP

	Full NuScenes	Night NuScenes
BEVFusion [5]	0.676	0.657
UniBEV [12]	0.641	0.64
ContextualFusion [7]	0.669	0.728
Ours	0.673	0.745

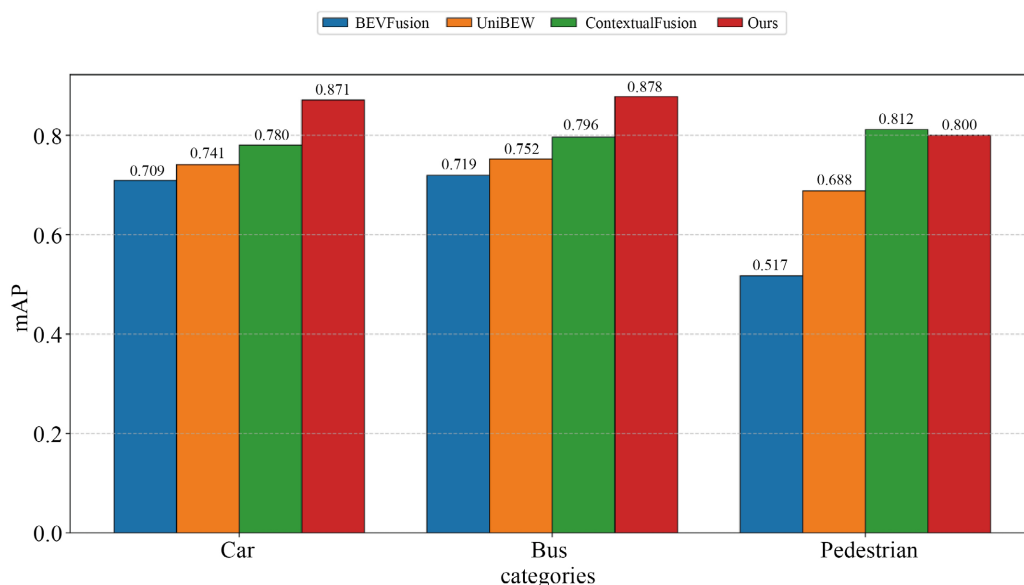


Figure 4. Detection accuracy AP of each algorithm in main categories in night scenes

图 4. 夜间场景下各个算法在主要类别上的检测精度 AP

在夜间数据集中, 轿车、行人和公交车的样本数量较多, 但这些类别易受过曝区域和暗光区域的影响。因此, 我们重点对比了各算法对这三个类别的检测结果 AP, 具体的结果如图 4 所示。可视化结果显示, 改进后的算法在各类别上的检测精度均表现优异, 且明显优于改进前的算法 BEVFusion。实验结果证明了这两个模块的引入提升了算法在夜间复杂场景下的检测能力, 使算法能够更准确地识别和定位目标。

4. 结论

本文研究了夜间场景下的基于激光雷达与相机融合的多模态目标检测算法。首先, 针对夜间由于车灯照射或者反射物引起的图像过曝, 引入 SPM 模块, 采取掩码策略对图像过曝区域进行了遮蔽, 成功规避了夜间图像过曝部分对模型检测性能的影响; 接着, 通过引入 CNW 模块优化了不同模态数据在视融合, 避免了图像暗光部分特征不足以及图像噪声对融合后特征的影响, 提升了模型的整体性能。基于 NuScenes 数据集及其夜间子集的实验证明, 改进后的算法不仅在常规场景下有着较好的检测性能, 且在夜间的检测性能优于其他算法。实验证明了本文算法的优秀性能, 展示了其在实际应用中的潜力和优势。

参考文献

- [1] 王科俊, 赵彦东, 邢向磊. 深度学习在无人驾驶汽车领域应用的研究进展[J]. 智能系统学报, 2018, 13(1): 55-69.
- [2] 国务院关于印发“十四五”现代综合交通运输体系发展规划的通知[J]. 中华人民共和国国务院公报, 2022(4): 8-28.
- [3] 李昌财, 陈刚, 侯作勋, 等. 自动驾驶中的三维目标检测算法研究综述[J]. 中国图象图形学报, 2024, 29(11): 3238-3264.
- [4] Vora, S., Lang, A.H., Helou, B. and Beijbom, O. (2020) Pointpainting: Sequential Fusion for 3D Object Detection. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 4604-4612. <https://doi.org/10.1109/cvpr42600.2020.00466>
- [5] Liang, T., Yu, K., et al. (2022) Bevfusion: A Simple and Robust Lidar-Camera Fusion Framework. *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, New Orleans, 28 November-9 December 2022, 10421-10434.
- [6] Zhang, C., Wang, H., Cai, Y., Chen, L., Li, Y., Sotelo, M.A., et al. (2022) Robust-Fusionnet: Deep Multimodal Sensor Fusion for 3D Object Detection under Severe Weather Conditions. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-13. <https://doi.org/10.1109/tim.2022.3191724>
- [7] Sural, S., Sahu, N. and Rajkumar, R.R. (2024) Contextualfusion: Context-Based Multi-Sensor Fusion for 3D Object Detection in Adverse Operating Conditions. 2024 *IEEE Intelligent Vehicles Symposium (IV)*, Jeju Island, 2-5 June 2024, 1534-1541. <https://doi.org/10.1109/iv55156.2024.10588584>
- [8] Guo, C., Li, C., Guo, J., Loy, C.C., Hou, J., Kwong, S., et al. (2020) Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 1780-1789. <https://doi.org/10.1109/cvpr42600.2020.00185>
- [9] Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., et al. (2021) Enlightengan: Deep Light Enhancement without Paired Supervision. *IEEE Transactions on Image Processing*, **30**, 2340-2349. <https://doi.org/10.1109/tip.2021.3051462>
- [10] Godard, C., Aodha, O.M., Firman, M. and Brostow, G. (2019) Digging into Self-Supervised Monocular Depth Estimation. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 3828-3838. <https://doi.org/10.1109/iccv.2019.00393>
- [11] Zheng, Y., Zhong, C., Li, P., Gao, H., Zheng, Y., Jin, B., et al. (2023) STEPS: Joint Self-Supervised Nighttime Image Enhancement and Depth Estimation. 2023 *IEEE International Conference on Robotics and Automation (ICRA)*, London, 29 May-2 June 2023, 4916-4923. <https://doi.org/10.1109/icra48891.2023.10160708>
- [12] Wang, S., Caesar, H., Nan, L. and Kooij, J.F.P. (2024) Unibev: Multi-Modal 3D Object Detection with Uniform BEV Encoders for Robustness against Missing Sensor Modalities. 2024 *IEEE Intelligent Vehicles Symposium (IV)*, Jeju Island, 2-5 June 2024, 2776-2783. <https://doi.org/10.1109/iv55156.2024.10588783>
- [13] Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S. and Miao, Y. (2021) Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sensing*, **13**, Article No. 4712. <https://doi.org/10.3390/rs13224712>
- [14] Zhou, Y. and Tuzel, O. (2018) VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4490-4499.

-
- <https://doi.org/10.1109/cvpr.2018.00472>
- [15] Bai, X., Hu, Z., Zhu, X., Huang, Q., Chen, Y., Fu, H., *et al.* (2022) Transfusion: Robust Lidar-Camera Fusion for 3D Object Detection with Transformers. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 1090-1099. <https://doi.org/10.1109/cvpr52688.2022.00116>
- [16] Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., *et al.* (2020) Nuscen: A Multimodal Dataset for Autonomous Driving. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 11621-11631. <https://doi.org/10.1109/cvpr42600.2020.01164>