

基于动态特征增强的三维小目标检测算法

李潞洋^{1,2,3}, 闫锦龙^{1,2,3}, 方泽儒^{1,2,3}, 金旗旗^{1,2,3}, 薛红新^{1,2,3}

(1. 中北大学计算机科学与技术学院, 山西 太原 030051; 2. 机器视觉与虚拟现实山西省重点实验室, 山西 太原 030051;
3. 山西省视觉信息处理及智能机器人工程研究中心, 山西 太原 030051)

摘要: 在点云三维目标检测任务中, 点云数据的稀疏性在客观上对小目标检测构成显著挑战。小目标自身有效点数稀少导致结构信息缺失与边界模糊, 而上下文感知能力不足阻碍模型有效利用周围环境信息进行空间推理与语义补全, 进而引发定位偏差, 以及其固有的空间定位困难、通道表达弱和特征易被背景淹没等问题, 共同制约了检测性能的提升。为缓解上述问题对检测精度造成的影响, 提出一种融合动态特征提取与特征增强映射的动态感知三维检测算法。该算法聚焦特征提取与候选框生成两大关键阶段对小目标检测进行优化。首先, 引入动态点特征预测网络(DPPFN), 通过自适应预测补充采样点以强化对小目标的结构感知能力; 其次, 构建特征增强映射网络(FEMN), 对原始特征及动态预测网络生成的特征进行深度融合, 输出富含上下文信息的二维特征图, 以有效弥补上下文缺失并提升小目标的定位精度; 最后, 设计点云特征增强网络(PCFEN), 在通道与空间双维度提升网络对小目标关键区域的聚焦能力。在 nuScenes 数据集上的实验结果表明, 所提算法性能优于当前主流目标检测算法, 与基准模型 CenterPoint 相比, 均值平均精度(mAP)由 56.1% 提升至 59.4%, Nuscenes 检测分数(NDS)由 64.4 提升至 67.4。

关键词: 激光雷达; 动态特征提取; 特征增强; 多尺度特征; 目标检测

中图分类号: TP391.41

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0252879

3D Small Object Detection Algorithm Based on Dynamic Feature Enhancement

LI Luyang^{1,2,3}, YAN Jinlong^{1,2,3}, FANG Zeru^{1,2,3}, JIN Qiqi^{1,2,3}, XUE Hongxin^{1,2,3}

(1. School of Computer Science and Technology, North University of China, Taiyuan 030051, Shanxi, China;

2. Shanxi Key Laboratory of Machine Vision & Virtual Reality, Taiyuan 030051, Shanxi, China;

3. Shanxi Vision Information Processing and Intelligent Robot Engineering Research Center, Taiyuan 030051, Shanxi, China)

【Abstract】 In 3D object detection from point clouds, the inherent sparsity of Light Detection And Ranging (LiDAR) data poses pronounced challenges for small objects. A small number of effective points lead to weak structural cues and blurry boundaries; limited contextual awareness hinders spatial reasoning and semantic completion, causing localization bias; and the difficulty of precise spatial localization, weak channel expressiveness, and background dominance constrain accuracy. To mitigate the impact of the aforementioned issues on detection accuracy, a dynamic-aware 3D detector is proposed that integrates dynamic feature extraction with feature-enhancement mapping, targeting two critical stages of small-object detection: feature extraction and candidate generation. Specifically, a Dynamic Point Feature Prediction Network (DPPFN) that adaptively predicts and supplements sampling points to strengthen structural perception of small objects is introduced. Subsequently, a Feature Enhancement Mapping Network (FEMN) is built that deeply fuses the original features with those produced by the dynamic module to yield context-rich 2D feature maps, thereby compensating for contextual deficiency and improving localization. Finally, a Point Cloud Feature Enhancement Network (PCFEN) module is designed to sharpen focus on key small-object regions along both channel and spatial dimensions. Experiments on the nuScenes dataset demonstrate that the proposed approach performs better than mainstream detectors. Relative to the CenterPoint baseline, the mean Average Precision (mAP) increases from 56.1% to 59.4% and the Nuscenes Detection Score (NDS) rises from 64.4 to 67.4.

【Key words】 Light Detection And Ranging (LiDAR); dynamic feature extraction; feature enhancement; multi-scale features; object detection

0 引言

近年来,随着自动驾驶技术的飞速发展,三维目

标检测作为实现环境感知的核心技术之一,其重要性日益凸显^[1]。该技术旨在从点云数据中准确识别物体类别并定位目标位置,为自动驾驶决策提供关

基金项目: 国家自然科学基金(62272426);山西省基础研究计划项目(202303021212372);机器视觉与虚拟现实山西省重点实验室研究基金(447-110103)。

作者简介: 李潞洋(CCF 专业会员),男,讲师,主研方向为计算机视觉;闫锦龙、方泽儒、金旗旗,硕士研究生;薛红新(通信作者),副教授。

收稿日期: 2025-08-11 **修回日期:** 2025-10-07 **E-mail:** xhxhj2279@nuc.edu.cn

键信息支撑。在实际场景中,小尺度目标(如行人、骑行者、交通锥等)常因尺寸小、点云稀疏,导致特征不明显、边界模糊,其检测精度对驾驶安全至关重要。因此,提升小目标检测能力对于保障自动驾驶系统的安全性与可靠性具有重大意义^[2-3]。针对这一问题,本文提出一种融合动态特征提取与特征增强映射的点云三维目标检测方法,从特征提取、候选框生成与注意力聚焦 3 个层面提升小目标的检测精度。该方法首先通过动态点特征预测网络(DPFPN)自适应生成关键采样点以强化小目标的结构感知;随后利用特征增强映射网络(FEMN)融合多尺度语义与几何信息,突出关键区域特征;最后构建点云特征增强网络(PCFEN)注意力模块,在通道与空间维度进一步聚焦小目标区域,实现更稳定准确的检测效果。

1 相关研究

基于点云的三维目标检测技术近年来发展迅速,主流方法可归纳为基于点和基于体素两大类^[4-5]。基于点的方法(如 PointNet^[6]、PointNet++^[7]、PointRCNN^[8])直接处理点云数据,能够保留原始信息以获得较高的检测精度。然而,其所需的反复点采样与特征聚合操作带来了显著的计算开销。相比之下,基于体素的方法(如 VoxelNet^[9]、PointPillars^[10]、SECOND^[11]、VoxelRCNN^[12]、CenterPoint^[13])将不规则点云转换为规则的三维网格(体素),利用三维卷积神经网络进行检测,因其计算效率高、结构规整且易于优化迁移等优势而被广泛应用。

在目标框生成机制上,无锚框(Anchor-free)方法因避免预设大量锚框而提升了计算效率与灵活性^[14],成为近年来的研究热点,该方法主要分为

两类:基于关键点的方法(如 CornerNet^[15]、ExtremeNet^[16])通过定位目标关键点(如角点、极值点)来组合预测目标框;基于中心点的方法(如 CenterNet^[17]、FCOS^[18])则聚焦于目标中心点的检测与属性回归。这类方法更符合点云数据的几何特性,尤其在处理尺度变化较大的目标时表现优异。

CenterPoint 是融合了基于体系结构与基于中心点思想的代表性算法。其第 1 阶段将点云体素化后预测目标中心点并回归生成初始 3D 候选框;第 2 阶段则对这些候选框进行进一步优化和细化以提升精度。

尽管 CenterPoint 在大中型目标检测上表现优异,但其在小目标检测中存在明显局限:首先,在特征提取阶段,小目标点云数量有限且分布稀疏,导致边界和结构特征难以被充分表征;其次,在候选框生成与细化阶段,CenterPoint 依赖于初始候选框的特征进行优化,而小目标上下文信息不足,使得候选框难以有效校准初始定位偏差;最后,小目标的通道表达能力很弱,其特征容易在多尺度融合中被背景淹没。上述不足共同制约了 CenterPoint 在小目标检测任务中的性能提升。

如图 1 所示,CenterPoint 在第 1 阶段的候选框优化通常依赖目标表面(如 4 个侧面)中心位置的特征。对于点云稀疏、尺寸较小的小目标,这些有限的固定位置点特征不足以精确表征目标结构,容易导致初始定位偏差。在第 2 阶段,小目标自身包含的有效点云信息稀少,其上下文特征感知能力本身较弱,细化过程难以有效弥补初始偏差,最终进一步限制小目标检测精度的提升。因此,亟需在特征提取阶段增强对稀疏小目标结构信息的有效捕捉能力,并在候选框生成阶段引入更精细的定位机制。

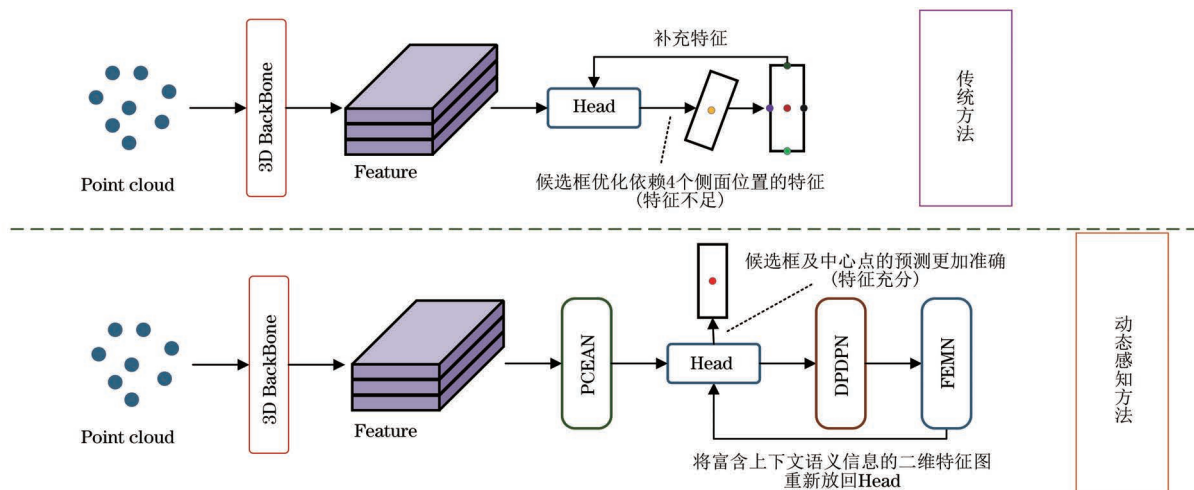


图 1 本文方法与传统方法的比较

Fig. 1 Comparison between the proposed method and traditional method

基于上述问题,本文旨在突破现有方法在小目标检测上的局限,提出了一种融合动态特征提取与特征增强映射的动态感知三维检测(DynA3D)算法,从特征提取、候选框生成和注意力聚焦 3 个层面进行优化,以全面提升小目标检测精度,如图 2 所示。首先,利用动态点特征预测网络(DPFPN)自适应地生成多个采样点,显著补充和强化小目标的边界与结构特征信息。其次,

构建特征增强映射网络(FEMN),对原始特征及 DPFPN 生成的动态特征进行深度融合与映射,输出富含上下文语义信息的二维特征图,进一步提升候选框定位的精准性。最后,在体素化后的点云基础上,引入点云特征增强网络(PCFEN),通过多尺度特征提取融合与跨通道和空间维度的注意力机制,增强网络对小目标关键特征区域的聚焦能力。

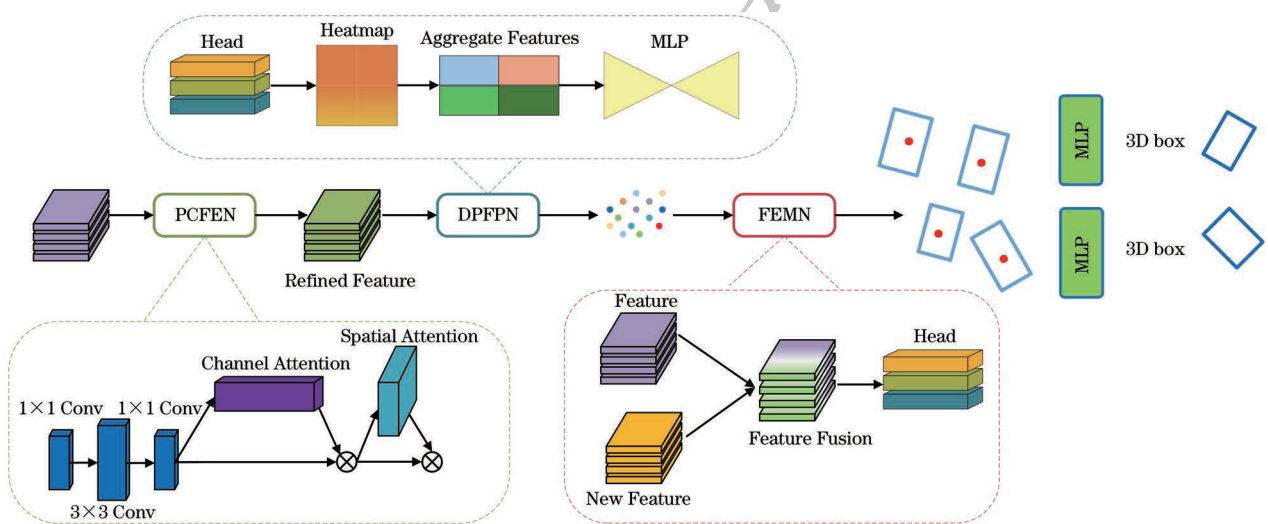


图 2 本文方法整体框架

Fig.2 Overall framework of the proposed method

2 动态感知三维检测算法

三维目标检测流程通常包含点云预处理、特征提取与候选框生成等关键阶段。在小目标检测任务中,由于其点云数量少、边界结构模糊且上下文信息有限,在特征提取阶段因有效点数稀少,难以提取出稳定的语义特征,在候选框生成阶段又因上下文信息缺失而导致目标定位偏差,这些问题共同造成了小目标检测性能下降。针对上述问题,本文提出一种面向小目标的动态感知三维检测算法,该算法从 3 个层面对点云小目标检测任务进行优化,具体包括:

1) 边界与结构建模强化。在特征提取阶段针对小目标因点数稀疏导致边界模糊、结构信息缺失的问题,设计动态点特征预测网络(DPFPN),通过热力图引导的方法动态生成多采样点,强化小目标的边界建模与结构感知能力。

2) 上下文特征融合增强。在候选框生成阶段,针对小目标由于体积小、有效点稀疏且分布不均,易造成语义特征表达不足与上下文信息缺失的问题,导致候选框生成不准确,构建了特征增强映射网络(FEMN),将动态提取的细粒度特征显式融合至原始特征图中,缓解小目标特征的上下文缺失问题。

3) 关键区域注意力聚焦。在特征提取阶段,针对小目标点云语义表达不足、空间信息易被背景干扰的问题,设计点云特征增强网络(PCFEN),从通道与空间两个维度聚焦小目标区域,提高检测器对小目标关键区域的响应能力。

为更直观地展示本文提出的动态感知三维小目标检测算法的总体结构与各模块间的特征流向,本文构建了算法总体框架,如图 3 所示。整体流程包含输入层、注意力聚焦层、特征提取与融合阶段以及候选框生成阶段 4 个部分。首先,将输入的 LiDAR 点云数据输入到点云特征增强网络(PCFEN)中,从通道与空间维度对关键区域进行特征聚焦;随后,通过动态特征预测网络(DPFPN)进行边界与结构信息建模,生成增强的动态特征;接着,特征增强映射网络(FEMN)将多尺度细粒度特征与原始特征显式融合,从而提升小目标的上下文表达能力;最后,候选框生成模块根据融合特征完成目标定位与检测结果输出。

2.1 动态点特征预测网络

在三维目标检测任务中,特征提取阶段的关键在于提取有效的特征信息来为候选框阶段候选框的生成提供准确、丰富的特征信息;然而,对于小目标

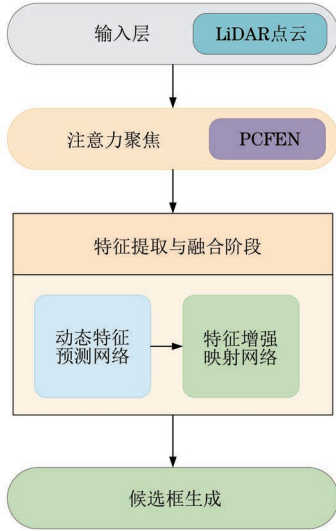


图 3 各模块间的特征流向

Fig.3 The characteristic flow direction among each module

而言,由于其点云数量少、形状轮廓不明显等客观原因,使得小目标的边界结构难以被清晰刻画,这种边界模糊性会影响后续候选框的定位精度,导致定位

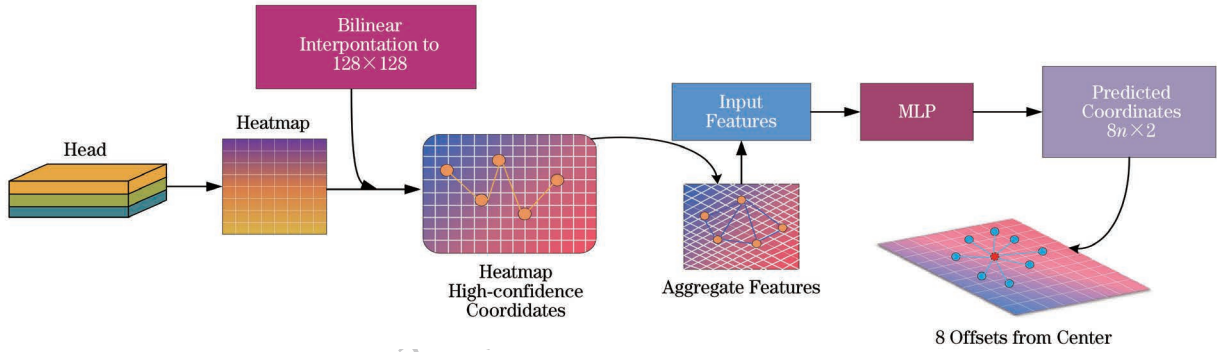


图 4 动态点特征预测网络

Fig.4 Dynamic point feature prediction network

随后,利用 Sigmoid 激活函数对热力图进行归一化,提取高置信度区域的位置:

$$M = \sigma(H_{down}), M_{b,c,h,w} > 0.9 \quad (2)$$

从所有通道中选取任一通道高于阈值的位置,得到置信度掩码,并提取其空间坐标集合 $P = (b_i, h_i, w_i)$ 。在这些高置信度点上,从原始输入特征图 $X \in \mathbb{R}^{[B,C,H,W]}$ 中提取其对应位置的特征,并在批次维度上进行聚合:

$$f_b = \sum_{(b,h,w) \in p} X_{b,.,h,w} \quad (3)$$

每个聚合后的特征向量 $f_b \in \mathbb{R}^C$ 被输入至一个多层感知机(MLP),用于动态预测多个坐标偏移。假设每个中心点预测 K 个偏移点(本文实验中 $K = 8$),每个点包含二维坐标 $(\Delta x, \Delta y)$,则最终输出为:

$$O_b = \text{MLP}(f_b) \in \mathbb{R}^{[K \cdot n_b \cdot 2]} \quad (4)$$

式中: n_b 为第 b 个 batch 中的高置信度点个数;输出

偏移;另外,由于小目标体素内点分布稀疏,很多关键点在特征图中难以被表达出来,使得模型难以正确感知目标的几何结构,生成的候选框难以精确覆盖目标最佳轮廓。

为缓解小目标点云边界模糊在候选框生成阶段引发的定位偏移与覆盖偏移问题,本文提出并引入一种基于热力图引导的动态点特征预测模块。该模块采用热力图引导机制,动态生成多个采样点以精准捕捉小目标的结构与边界信息,从而准确地确定其定位信息。所预测的多组偏移坐标能够有效应对体素稀疏问题,特别是小目标的边界模糊问题,显著增强边界特征信息的丰富度,避免因边界模糊导致的定位偏差问题。

如图 4 所示,输入特征经过检测头后生成初步的热力图 $H \in \mathbb{R}^{[B,C,H,W]}$ 。为了与后续模块保持一致,该热力图通过双线性插值下采样至 $H \in \mathbb{R}^{[B,C,128,128]}$ 的空间尺寸:

$$H_{down} = \text{BilinearInterp}(H, s_{size} = 128 \times 128) \quad (1)$$

为坐标集合 $O_b \in \mathbb{R}^{[K \cdot n_b \cdot 2]}$,用于定位更丰富的结构信息。

然而,如果采样点偏移过大,可能会超出真实目标边界,影响候选框的精度。为此,在 DPFPN 中设计了三重约束机制。首先,利用热力图置信度约束,仅在高置信度区域生成偏移点,从源头上减少无关区域干扰。其次,采用边界一致性约束,通过在训练过程中对偏移点加入边界正则项,使预测的偏移点尽可能保持在真实标注框内,避免了采样点过度偏移到目标外部区域:

$$L_{boundary} = \frac{1}{N} \sum_{i=1}^N \max(0, d(p_i, B_{gt}) - \epsilon) \quad (5)$$

式中: N 表示高置信度点的个数; $d(p_i, B_{gt})$ 表示采样点 p_i 到真实边界框 B_{gt} 的最小距离; ϵ 为允许的偏移范围。若采样点超出阈值范围,将受到惩罚,从而抑制过度偏移。最后,在推理阶段引入

坐标裁剪策略,对超出候选框范围的采样点进行裁剪,使最终的动态采样结果始终紧密围绕目标边界分布。

最后,引入坐标裁剪策略,对超出目标候选框范围的采样点进行限制,确保最终的动态采样结果紧密围绕真实目标边界分布,其公式表示为:

$$\hat{p}_i = \text{Clip}(p_i, B) \quad (6)$$

式中: \hat{p} 表示裁剪后的采样点坐标;候选框范围 $B = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$; $\text{Clip}(\cdot)$ 操作将采样点限制 B 内部,若预测点超出边界,则被裁剪到最近的边界位置,若在范围内则保持不变。通过该策略,推理阶段的采样点能够始终分布在候选框附近,避免偏移点落入无关区域。

动态点特征预测网络通过热力图引导在高置信度区域生成多个偏移采样点,有效地捕捉小目标的结构与边界细节,缓解了因点云稀疏导致的小目标边界模糊等问题。结合提出的三重约束机制(热力图置信度约束、边界一致性约束与坐标裁剪策略),DPFPN 能够有效保证采样点分布的合理性与边界贴合性,避免偏移问题对检测精度的不利影响,从而

为候选框生成阶段提供更为精准的信息,并进一步提升了小目标检测的定位精度。

2.2 特征增强映射网络

由于小目标点云本身稀疏,其在主干网络中生成的特征往往缺乏足够的语义信息,使得在候选框生成阶段难以对其进行准确的建模,存在语义表达能力弱和上下文信息缺失等问题,因此通过引入由动态点特征预测网络(DPFPN)生成的特征信息,对原始特征图进行特征增强,使得小目标区域获得更多的语义支撑,将增强后的特征图输入候选框预测头,提升网络对小目标的空间感知与结构理解能力,进一步提升小目标物体的候选框生成质量,从而提高小目标的检测精度。

如图 5 所示,通过对原始特征与动态特征聚合网络(DPFPN)生成的特征进行深度融合,得到具有丰富特征信息的二维特征图,当得到动态特征聚合网络提取的点坐标特征后,首先创建一个与输入特征图形状相同的副本:

$$X_{\text{modified}} = x \quad (7)$$

式中: x 为输入特征图; X_{modified} 为创建的副本。

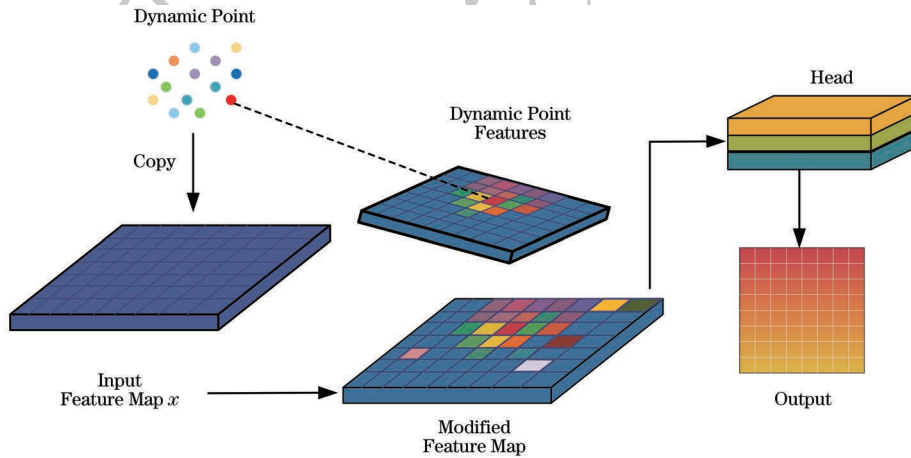


图 5 特征增强映射网络

Fig.5 Feature enhancement mapping network

接下来,遍历每个高置信度坐标,提取预测坐标位置的特征,并将这些特征累加到输入特征图的对应位置:

$$X_{\text{modified}_{b, :, h, w}} = X_{\text{modified}_{b, :, h, w}} + F_{\text{pred}_{b, :, h, w}} \quad (8)$$

式中: F_{pred} 为预测坐标位置的特征。通过将预测坐标位置的特征值加到输入特征图的对应位置,特征增强映射网络能够有效融合额外点特征信息与原始特征图信息,从而增强特征图的表达能力,最终生成经过特殊处理的特征图 X_{modified} 。这样就实现了动态点特征的特征信息与原特征图中的特征信息进行拼接,然后将完整的特征信息图重新放回 Head 头中进行预测信息:

$$P_{\text{predictions}} = H(X_{\text{modified}}) \quad (9)$$

式中: H 为 Head 头的预测函数。通过这种方式,特征增强映射网络能够更充分地利用额外点的特征信息,与原始特征图形成更完整的特征表示,从而细化第一阶段生成的 3D 边界框,提高模型的检测精度。

特征增强映射网络(FEMN)通过动态融合 DPFPN 生成的增强特征和原始特征图信息,大幅减少了小目标特征信息的损失。

2.3 点云特征增强网络

在三维目标检测任务中,尤其在对体积较小的目标,如行人、自行车、摩托车和交通锥等而言,由于其包含的点数稀少,通道特征表达能力弱,在特征图

上易被压制,特别是存在多个大目标时,小目标的通道响应被稀释,导致其识别与分类困难。同时,小目标在空间位置上占据的点云范围较小,特别是在稀疏点云中与背景界限模糊,易被背景干扰。为此,针对小目标物体在点云检测中存在的通道表达弱、空间定位难、特征易被背景淹没等问题,本文提出点云

特征增强网络(PCFEN)如图 6 所示,通过通道和空间两个维度的注意力机制,对小目标区域进行精准地聚焦,增强模型对关键特征区域的感知能力。通过引入注意力机制,使得模型能够更有效地捕捉小目标的特征信息,提升其在检测过程中的准确性。

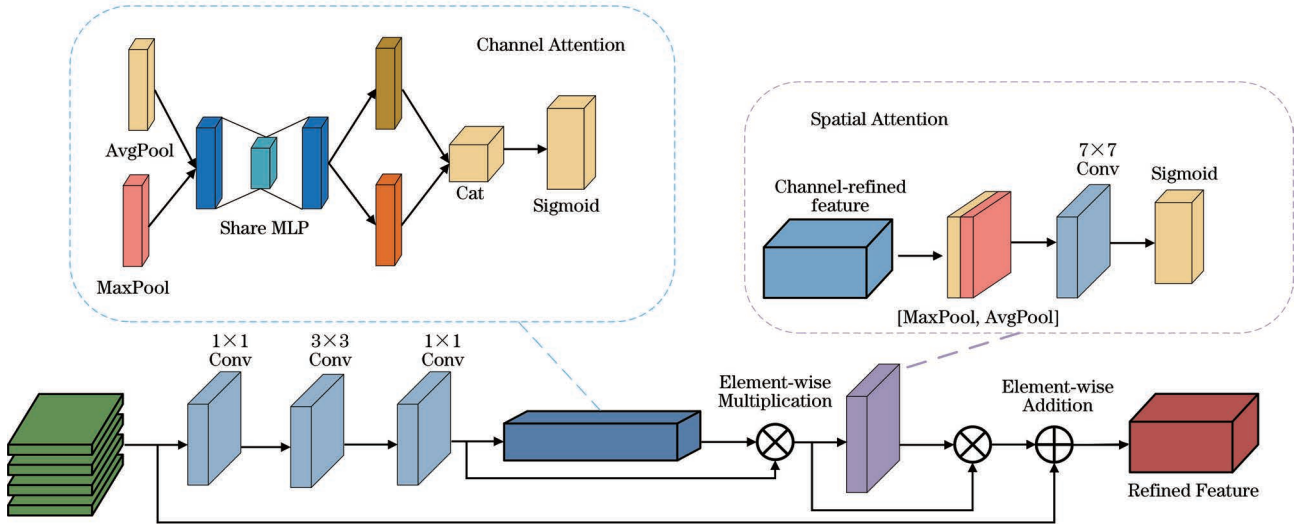


图 6 点云特征增强网络

Fig. 6 Point cloud feature enhancement network

具体来说,PCFEN 通过通道和空间的顺序生成注意力图,加入的通道注意力模块有效地解决了通道信息弱化的问题,使模型能够更精准地捕捉小目标物体的特征信息,提升小目标物体的检测效果。空间注意力模块通过聚焦小目标所在的关键区域,抑制背景噪声和无关信息,使模型能够更有效地提取小目标的空间特征,提升了小目标物体的定位精度和检测精度。由于 PCFEN 是轻量级的通用模块,因此其计算开销可以忽略不计。

通道注意力通过全局平均池化和最大池化捕捉通道的全局信息,并通过共享的多层感知机(MLP)学习通道之间的相互关系。最终,使用 Sigmoid 激活函数生成注意力权重,对输入特征图进行加权。这样可以自适应地强调重要通道,抑制不重要通道,从而提高模型对关键信息的捕捉能力,在小目标物体检测中,能够突出小目标物体的关键信息特征,避免其特征被背景或大目标物体掩盖,显著提升小目标物体的检测精度。通道注意力模块的输出公式如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F) + \text{MLP}(\text{MaxPool}(F)))) \quad (10)$$

式中: F 表示输入特征图;AvgPool 和 MaxPool 分别表示全局平均池化和最大池化操作;MLP 表示多

层感知机; σ 表示 Sigmoid 激活函数。

空间注意力模块要接收经过通道注意力模块处理后的特征图,即已经加权了通道重要性的特征图。随后,这些特征图通过一个 7×7 的卷积层,该卷积层采用大小为 7×7 的卷积核在空间维度上滑动,以捕捉局部区域的空间特征信息。卷积层的输出经过激活函数生成每个空间位置的注意力权重图,最终与输入特征图相结合,实现空间加权。空间注意力模块的输出公式如下:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (11)$$

式中: $f^{7 \times 7}$ 表示一个 7×7 的卷积操作; $[\text{AvgPool}(F); \text{MaxPool}(F)]$ 表示将平均池化和最大池化结果沿通道轴拼接起来。在小目标检测任务中,目标物体通常尺寸较小、纹理信息有限,易在多尺度特征融合过程中被背景信息淹没。引入空间注意力机制有助于显著增强模型对目标区域的聚焦能力,改善目标特征的表达质量。特别是在稀疏点云数据中,空间注意力模块能够有效捕捉细粒度的空间位置信息,提升小目标的定位精度。

综合上述机制,为保持梯度流畅并保留原始特征信息,采用残差连接设计整个 PCFEN 模块。在通道-空间注意力机制前引入了三层卷积 $1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1$ 作为轻量特征抽取模块来获取空间注

意图,因为当前 BEV 特征图的分辨率与体素大小共同决定了 3×3 在不增加明显计算量的前提下即可覆盖小目标周围的有效区域,能够更好地保留边界与几何细节:

$$F_{\text{out}} = F + M_s(F) \cdot (M_c(F) \cdot \text{Conv}_{1 \times 1, 3 \times 3, 1 \times 1}(F)) \quad (12)$$

式中: F_{out} 为经过空间和通道注意力机制增强后的输出特征图; $M_s(F)$ 为空间注意力模块生成的空间注意力权重图; $M_c(F)$ 为通道注意力模块生成的通道注意力权重图。

小目标在 BEV 特征图上的空间占用通常仅为数个栅格。过大的卷积核或强膨胀会将较多背景上下文引入注意力图,导致小目标边界出现“晕化”与对比度下降,进而削弱定位与分类的判别性。基于这一点,本文在空间注意力中采用 3×3 的卷积,以获得最小必要的局部感受野与更强的空间选择性:一方面可避免大核造成的背景干扰,另一方面与上游特征金字塔提供的跨尺度上下文形成互补,从而在维持实时性与数值稳定性的同时,增强检测器对小目标关键区域的响应。

在不改变整体框架的前提下,为了验证多尺度卷积相比于单尺度卷积对小目标检测效果是否带来收益,本文实现一个多尺度变体用于对比,记为 PCFEN-MS,在原有 3×3 分支基础上,引入 5×5 与膨胀 3×3 ($d=2$) 两个并行空间分支,通过轻量化融合获得多尺度但受控的空间聚合能力。设输入特征为 $X \in \mathbb{R}^{C \times H \times W}$ 。先进行通道压缩的全局统计聚合:

$$P = [\text{AvgPool}(X); \text{MaxPool}(X)] \in \mathbb{R}^{2 \times H \times W} \quad (13)$$

在 P 上并联三路空间分支:

$$\begin{aligned} A_s^{(3)} &= \text{DWConv}_{3 \times 3}(P) \\ A_s^{(5)} &= \text{DWConv}_{5 \times 5}(P) \\ A_s^{(d_2)} &= \text{DWConv}_{3 \times 3, d_{il}=2}(P) \end{aligned} \quad (14)$$

在三路输出上按通道级联,并经卷积 1×1 与 Sigmoid 得到多尺度空间注意力权重:

$$A_s^{\text{MS}} = \sigma(\text{Conv}_{1 \times 1}(\text{Concat}(A_s^{(3)}, A_s^{(5)}, A_s^{(d_2)}))) \quad (15)$$

最终输出为:

$$Y = X \odot A_s^{\text{MS}} \quad (16)$$

式中: \odot 表示逐元素相乘。多尺度并未优于单尺度,由于小目标占格极少,多分支聚合更容易把远距背景和噪声一并引入,注意力峰值被摊薄,边界

对比度下降,小目标 mAP 略低,同时额外分支带来一定的延迟和显存开销。综合准确率与开销,本文最终保留 3×3 单尺度空间注意力作为默认实现。

PCFEN 模块以通道-空间联合注意力机制为核心,从语义和空间两个维度增强点云特征表达能力。模块先通过轻量卷积提取初始特征,继而引入通道与空间注意力引导网络关注通道信息与关键空间区域,从而突出小目标在稀疏背景中的显著性,该机制强化了小目标特征的表达,显著改善了行人、自行车、交通锥等小目标特征的检测性能。

3 实验结果与分析

3.1 数据集

nuScenes 数据集^[19]包含了波士顿和新加坡两个城市采集的自动驾驶车辆传感器数据,数据集包括 1 000 个驾驶场景,每个场景大约 20 s,总时长约为 28 h,包括 LiDAR 点云、摄像头图像等多种传感器数据,并且还提供了详细的 3D 目标标注,具体包括汽车、行人、公交车、障碍物、交通锥、卡车、拖车、摩托车、建筑车辆和自行车 10 类目标,总共有约 140 万个 3D 标注框。

3.2 实验设置

本文实验使用 mmdetection3D 目标检测框架和 PyTorch 深度学习框架构建,本文实验的硬件环境为: Intel Core i7-14700F, NVIDIA A10。软件环境为: Ubuntu 24.04.1 LTS, Cuda 11.6 版本。使用 nuScenes 数据集^[19]进行训练和测试,并使用 AdamW^[20]优化器进行训练。设置初始学习率为 1×10^{-4} ,权重衰减值为 0.01,最大迭代次数为 40。X 轴、Y 轴、Z 轴的检测范围分别为 $[-54 \text{ m}, +54 \text{ m}]$ 、 $[-54 \text{ m}, +54 \text{ m}]$ 、 $[-5 \text{ m}, +3 \text{ m}]$ 内,体素大小为 $(0.075 \text{ m}, 0.075 \text{ m}, 0.2 \text{ m})$ 。

3.3 评价指标

为了评估模型的性能,使用均值平均精度 (mAP, P_{mAP}) 和 Nuscenes 检测分数 (NDS, N_{NDS}) 作为评估指标,其中 mAP 的公式为:

$$P_{\text{mAP}} = \frac{1}{|C| |D|} \sum_{c \in C} \sum_{d \in D} P_{\text{AP}, c, d} \quad (17)$$

式中: mAP 与 2D 目标检测中用于衡量精度和召回率的 mAP 类似,但在 3D 目标检测中,其计算基于地平面上的中心距离,而非三维交并比 (IoU),用于评估预测结果与实际目标之间的匹配程度, C 表示类别数量, D 指地平面的 2D 中心距离。NDS 的公

式为:

$$N_{\text{NDS}} = \frac{1}{10} \left[5P_{\text{mAP}} + \sum_{P_{\text{mTP}} \in P_{\text{TP}}} (1 - \min(1, P_{\text{mTP}})) \right] \quad (18)$$

式中: P_{mTP} 表示平均真阳性中心点距离误差。

此外, 本文将行人、摩托车、自行车、交通锥 4 类的 AP 取算术平均, 记为 small-mAP, 用于衡量小目标整体检测性能。

3.4 对比试验

在 nuScenes 数据集上将本文方法与先进的三维目标检测方法进行对比实验, 对比算法包括 WYSIWYG^[21]、PointPillars^[10]、CVCNet^[22]、PointPainting^[23]、PMPNet^[24]、SSN^[25]、CBGS^[26] 以及 CenterPoint^[13]。

实验结果如表 1 所示, 改进后的算法在 mAP 和 NDS 指标上均取得了显著提升, 表明了改进算法在整体检测性能上具有明显优势。如表 1 所示, 改进后的算法在 mAP 和 NDS 指标上均优于基线网络 CenterPoint, 其中 mAP 从 56.1% 提升至 59.4%, NDS 从 64.4% 提升至 67.4%, 分别提升了 3.3 和 3.0 百分点, 表明改进算法在整体性能检测上具有显著优势。此外, 在小目标类别(行人、摩托车、自行车和交通锥)检测上, 改进方法的提升尤为显著。

表 1 不同算法在 nuScenes 数据集上的指标值对比

Table 1 Comparison of metric values for different algorithms on nuScenes dataset

算法	出版刊物	mAP/% ↑	NDS ↑
WYSIWYG ^[21]	CVPR	35.0	41.9
PointPillars ^[10]	CVPR	40.1	55.0
CVCNet ^[22]	CVPR	55.3	64.4
PointPainting ^[23]	CVPR	46.4	58.1
PMPNet ^[24]	ECCV	45.4	53.1
SSN ^[25]	ECCV	46.3	56.9
CBGS ^[26]	arXiv	52.8	63.3
CenterPoint ^[13]	CVPR	56.1	64.4
Ours	—	59.4	67.4

如表 2 所示, 与 CenterPoint 相比, 改进算法行人类别的 mAP 从 84.5% 提升至 85.8%, 摩托车类别的 mAP 从 54.0% 提升至 61.1%, 自行车类别从 35.1% 提升至 45.0%, 交通锥类别从 66.7% 提升至 70.5%。行人、摩托车、自行车和交通锥目标较小、特征较少, 且当提取特征时无法提取到有用的特征信息, 导致提取到的特征信息不足以支撑算法进行准确的检测。所提算法利用动态点特征提取网络和特征增强映射网络解决了小目标在特征提取中特征信息丢失的问题, 利用点云特征增强网络(PCFEN)提升了算法对小目标的感知能力。

表 2 不同算法在各目标上的检测性能比较

Table 2 Detection performance comparison of different algorithms on various objects

算法	汽车	卡车	公交	拖车	施工车辆	行人	摩托车	自行车	交通锥	路障
WYSIWYG ^[21]	79.1	30.4	46.6	40.1	7.1	65.0	18.2	0.1	28.8	34.7
PointPillars ^[10]	68.4	23.0	28.2	23.4	4.1	59.7	27.5	1.1	30.8	38.9
PointPainting ^[23]	77.9	35.8	36.2	37.3	15.8	73.3	41.5	24.1	62.4	60.2
CVCNet ^[22]	82.7	46.1	46.6	49.4	22.6	79.8	59.1	31.4	65.6	69.6
PMPNet ^[24]	79.7	33.6	47.1	43.1	18.1	76.5	40.7	7.9	58.8	48.8
SSN ^[25]	80.7	37.5	39.9	43.9	14.6	72.3	43.7	20.1	54.2	56.3
CBGS ^[26]	81.1	48.5	54.9	42.9	10.5	80.1	51.5	22.3	70.9	65.7
CenterPoint ^[13]	84.2	53.1	68.0	33.3	13.0	84.5	54.0	35.1	66.7	69.9
Ours	85.0	57.6	69.8	36.4	17.9	85.8	61.1	45.0	70.5	65.3

本文所提算法主要解决了小目标特征提取不充分导致小目标检测精度低的问题。因此, 对小目标的检测提升效果较为明显。此外, 在大目标类别(汽车、卡车、公交和拖车)上, 改进方法同样表现出色, 如表 2 所示, 汽车类别的 mAP 从 84.2% 提升至 85.0%, 卡车类别从 53.1% 提升至 57.6%, 公交类别从 68.0% 提升至 69.8%, 拖车类别从 33.3% 提升至 36.4%。这表明了本文的改进方法不仅在小目标检测上具有优势, 同时在大目标检测上也保持了稳定的高性能。

为了验证模型设计的有效性, 本文在 nuScenes 数据集上与基线网络进行了对比实验, 如表 3 所示。实验指标涵盖了均值平均 Translation 误差(mATE)、均值平均尺度误差(mASE)、均值平均朝向误差(mAOE)、均值平均速度误差(mAVE)、均值平均加速度误差(mAEE)等多个维度的误差评估, 以及均值平均精度(mAP)和 Nuscenes 检测分数(NDS)两项关键性能指标。从实验结果来看, 模型在多个关键指标上均有显著改善。其中, mAP 从 56.1% 提升至 59.4% (+3.3 百分点), NDS 从 64.6

提升至 67.4(+3.0), 表明了整体的检测精度有明显增长。均值平均位置误差(mATE)从 0.28 下降至 0.27, 虽然变化幅度较小, 但在稀疏点云场景中已能体现模型对目标空间位置估计的稳定性; mAOE 从 0.33 下降至 0.27, 说明模型在目标方向

预测上的精度明显提升, 改善了目标定位和朝向估计, 特别是对长条状和不对称的小目标尤为关键; mAVE 从 0.28 下降至 0.24, 反映了模型在运动目标速度建模方向能力的增强, 提升了对动态目标的预测精度。

表 3 提出方法与基线方法的性能比较

Table 3 Performance comparison between the proposed method and the baseline method

方法	mATE↓	mASE↓	mAOE↓	mAVE↓	mAAE↓	mAP/%↑	NDS↑
baseline	0.28	0.25	0.33	0.28	0.18	56.1	64.6
Ours	0.27	0.25	0.27	0.24	0.18	59.4	67.4

3.5 小目标专用对比与分析

为了系统评估本文方法对小目标的检测能力, 在 nuScenes 上与近年来代表性的方法进行对比, 包括 SASA^[27]、3DSSD^[28]、PVT-SSD^[29]、SSN^[25]、PointPainting^[23]、CenterPoint^[13] 以及本文方法。统计 4 个小目标类别, 即行人、摩托车、自行车、交通锥的 AP, 并以算术平均作为 small-mAP。性能对比结果如表 4 所示, 本文方法在 small-mAP 上达到 65.4%, 相较当前最强基线

CenterPoint^[13] 的 60.1% 提升了 5.3 个百分点; 在行人、摩托测、自行车和交通锥小目标检测上取得了最高 AP, 分别为 85.8%、61.1%、45.0% 和 70.5%, 与其他代表性方法相比, 本文方法的 small-mAP 分别较 SASA^[27]、3DSSD^[28]、PVT-SSD^[29]、SSN^[25]、PointPainting^[23]、CenterPoint^[13] 提升了 26.5、28.9、8.4、17.8、15.1、5.3 个百分点。在整体上, 本文在不牺牲整体精度的前提下, 对多类小目标均表现出稳定优势。

表 4 不同方法在 nuScenes 数据集小目标上的性能对比

Table 4 Performance comparison of different methods on nuScenes dataset small targets

方法	行人	摩托车	自行车	交通锥	small-mAP	%
SASA ^[27]	69.1	39.6	16.9	29.9	38.9	
3DSSD ^[28]	70.2	36.0	8.6	31.1	36.5	
PVT-SSD ^[29]	79.8	53.4	38.2	56.6	57.0	
SSN ^[25]	72.3	43.7	20.1	54.2	47.6	
PointPainting ^[23]	73.3	41.5	24.1	62.4	50.3	
CenterPoint ^[13]	84.5	54.0	35.1	66.7	60.1	
Ours	85.8	61.1	45.0	70.5	65.4	

3.6 消融实验

为了验证动态点特征预测网络(DPFEN)、特征增强映射网络(FEMN)和点云增强网络(PCFEN)对小目标检测精度的影响, 在 nuScenes 数据集上对小目标的行人、摩托车、自行车以交通锥进行消融实验, 基准模型为 CenterPoint。实验结果如表 5 所示。在加入点云特征增强注意力(PCFEN)后, 模型对行人、摩托车、自行车及交通锥的检测效果分别提升了 0.5、1.3、1.6、0.1 个百分点, 结果表明点云特征增强网络(PCFEN)能够提升模型的局部信息提取能力, 使得模型对小目标对象的检测效果有小幅提升。在同时加入动态点特征预测网络(DPFEN)和特征增强映射网络(FEMN)时, 模型在处理点云数据稀疏和对小目标特征表达能力方面得到了显著提升。模型对行人、摩托车、自行车以及交通锥的检

测效果分别提升了 1.3、6.5、7.7、3.3 个百分点, 结果表明, DPFEN 通过动态预测关键点的位置, 能够更全面地捕捉小目标的特征信息, 而 FEMN 通过深度融合动态特征和原始特征, 进一步增强了模型对小目标特征的表达能, 使得模型对小目标的检测精度大幅度提升。在同时加入 PCFEN、DPFEN 以及 FEMN 3 个模块之后, 模型对行人、摩托车、自行车

表 5 不同方法在小目标检测上的性能比较

Table 5 Performance comparison of different methods

方法	行人	摩托车	自行车	交通锥	%
baseline	84.5	54.0	35.1	66.7	
+PCFEN	85.0	55.3	36.7	66.6	
+DPFEN+FEMN	85.8	60.5	42.8	70.0	
Ours	85.8	61.1	45.0	70.5	

以交通锥的检测效果分别提升了 1.3、7.1、9.9、3.8 个百分点,实验结果表明,采用本文提出的方法后,小目标物体的检测精度显著提升,充分验证了所提方法的有效性。

为了验证各模块对模型性能的贡献,本文进行了消融实验,实验结果如表 6 所示。以 CenterPoint 为基线网络,逐步加入点云特征增强网络(PCFEN)、动态点特征预测网络(DPFEN)和特征增强映射网络(FEMN)对模型的 mAP 和 NDS 指标进行了评估。当仅加入 PCFEN 时,mAP 从 56.1% 提升至 56.9%,NDS 从 64.4 提升至 65.2,表明 PCFEN 通过结合通道和空间注意力机制,增强了模型对稀疏点云数据的特征捕捉能力,提升了整体检测性能。在同时加入 DPFEN 和 FEMN 后,mAP 进一步提升至 59.0%,NDS 提升至 66.9,说明 DPFEN 通过动态预测关键点的位置,能够更全面地捕捉目标的特征信息,而 FEMN 通过深度融合动态特征和原始特征,进一步增强了模型对目标特征的表达能力。

表 6 不同模块对模型性能的影响

Table 6 Impact of different modules on model performance

网络模型	mAP/% ↑	NDS ↑
baseline	56.1	64.4
baseline+PCFEN	56.9	65.2
baseline+DPFEN+FEMN	59.0	66.9
Ours	59.4	67.4

结合所有改进模块的改进 CenterPoint 模型在 mAP 和 NDS 上分别达到了 59.4% 和 67.4,相较于基线网络分别提升了 3.3 个百分点和 3.0。这表明,本文提出的各模块在解决小目标检测精度低等问题上具有显著效果,充分验证了所提方法的有效性和优越性。

为了进一步验证通道-空间注意力机制中的卷积结构设计的合理性,本文将所采用的瓶颈卷积结构与多尺度卷积 PCFEN-MS 进行了消融对比实验。实验结果如表 7 所示,从实验结果可以看出,多尺度卷积在整体性能上较基线模型有所提升,但在小目标类别的行人、摩托车、自行车和交通锥的检测精度上,仍不及本文提出的瓶颈卷积结果。

表 7 不同卷积结构在小目标检测上的性能比较

Table 7 Performance comparison of different convolutional structures on small object detection %

方法	mAP	行人	摩托车	自行车	交通锥
多尺度	58.7	85.0	55.3	36.7	66.6
单尺度	59.4	85.8	61.1	45.0	70.5

4 可视化实验

4.1 可视化实验说明

为了验证本文方法在小目标上的有效性,在 nuScenes 数据集中选取具有代表性的三帧场景进行成对可视化展示,如图 7 所示(彩色效果见《计算机工程》官网 HTML 版)。其中,每一行对应同一场景的对比,左列为基线 CenterPoint,右列为本文方法,红色框为检测结果,黄色圆圈标出仅由本文方法检出的真实小目标。

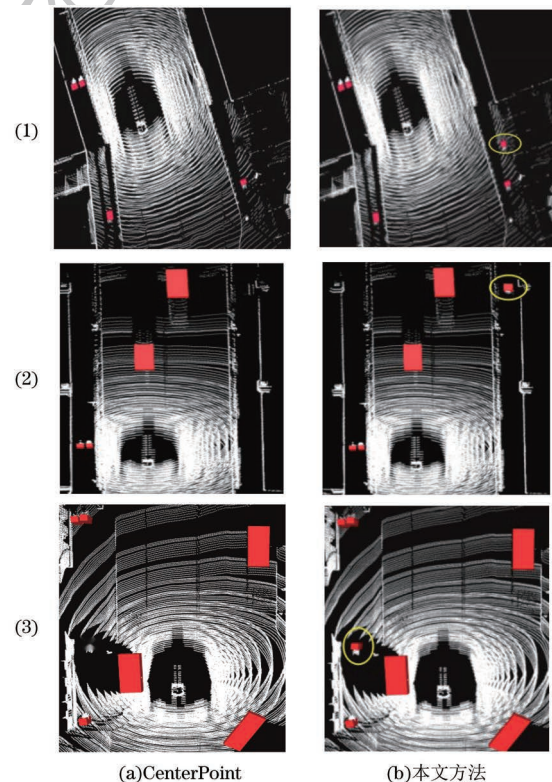


图 7 点云特征增强网络可视化

Fig. 7 Visualization of point cloud feature enhancement network

4.2 可视化结果分析

如图 7 所示,第 1 行展示了远距离且点云较稀疏的场景,道路右侧存在尺寸极小的目标。基线在该位置没有形成稳定候选,右列中本文方法能够给出明显且边界贴合的检测框,黄色圆圈标识的即为基线方法未检测出而本文方法检测出的小目标物体。第 2 行可视化基线易受背景干扰而漏检,本文方法通过对关键区域的特征加强与聚焦,仍能稳定检出并保留下来。第 3 行可视化为近处小尺寸且多目标共存的场景,在大目标干扰下基线对小目标的响应较弱,而本文方法对小目标的框定位更完整,未出现显著的额外误检。总体来看,这 3 组可视化表明,本文融合“动态特征提取+特征增强映射”的动

态感知三维检测算法,在“特征提取”和“候选框生成”两个关键阶段对小目标进行了有效优化,使得在远距、点云稀疏与复杂背景场景下的小目标检出与保留显著优于基线方法。

5 结束语

本文针对激光雷达三维目标检测中小目标检测精度低的问题,以 CenterPoint 为基准网络,提出了动态点特征预测网络和特征增强映射网络,提高了模型对小目标物体的特征提取能力和候选框生成精度,并引入了点云特征增强网络强化模型对稀疏点云特征的提取能力。在 nuScenes 数据集上的实验结果表明,动态感知三维检测算法对小目标物行人、摩托车、自行车以及交通锥的检测效果相较于基线模型提高了 1.3、7.1、9.9、3.8 百分点。实验结果表明了本文所提出的改进模型在实际的三维目标检测任务上的优越性和有效性。未来将开展跨数据集与恶劣天气的一体化评测与训练,即统一评测基线与数据划分,覆盖雨雾夜、强光、远距稀疏等典型场景,实现“换城市、换天气”后仍能稳定输出。

参考文献

- [1] 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.
LI C C, CHEN G, HOU Z X, et al. A review of research on three-dimensional target detection algorithms in autonomous driving [J]. Chinese Journal of Image Graphics, 2024, 29(11): 3238-3264. (in Chinese)
- [2] ZHANG B, WANG H S, YOU S L, et al. A small-size 3D object detection network for analyzing the sparsity of raw LiDAR point cloud[J]. Journal of Russian Laser Research, 2023, 44(6): 646-655.
- [3] WANG J, LIU Y, ZHU Y, et al. 3D point cloud object detection method based on multi-scale dynamic sparse voxelization[J]. Sensors, 2024, 24(6): 1804.
- [4] YANG H H, WANG W X, CHEN M H, et al. PVT-SSD: single-stage 3D object detector with point-voxel transformer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2023: 13476-13487.
- [5] 龙丽叶, 焦世超, 郭磊, 等. 基于紧凑中心的多模态三维模型检索研究[J]. 计算机工程, 2025, 51(2): 322-334.
LONG L Y, JIAO S C, GUO L, et al. Research on multimodal 3D model retrieval based on compact center[J]. Computer Engineering, 2025, 51(2): 322-334. (in Chinese)
- [6] CHARLES R Q, HAO S, MO K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 652-660.
- [7] QI C R, YI L, SU H, et al. Pointnet++: deep hierarchical feature learning on point sets in a metric space [C] // Proceedings of the Advances in Neural Information Processing Systems. Cambridge, USA: MIT Press, 2017: 301-312.
- [8] SHI S S, WANG X G, LI H S. PointRCNN: 3D object proposal generation and detection from point cloud [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE Press, 2019: 770-779.
- [9] ZHOU Y, TUZEL O. VoxelNet: end-to-end learning for point cloud based 3D object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE Press, 2018: 4490-4499.
- [10] LANG A H, VORA S, CAESAR H, et al. PointPillars: fast encoders for object detection from point clouds [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2019: 12697-12705.
- [11] YAN Y, MAO Y, LI B. SECOND: sparsely embedded convolutional detection[J]. Sensors, 2018, 18(10): e3337.
- [12] DENG J J, SHI S S, LI P W, et al. Voxel R-CNN: towards high performance voxel-based 3D object detection [J]. Artificial Intelligence, 2021, 35(2): 1201-1209.
- [13] YIN T, ZHOU X, KRAHENBUHL P. Center-based 3D object detection and tracking[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2021: 11784-11793.
- [14] 孙一杰, 李晓明. 结合动态循环金字塔与任务解耦的无锚框检测[J]. 计算机工程与设计, 2025, 46(4): 1157-1166.
SUN Y J, LI X M. Combining dynamic cyclic pyramid with task decoupling for anchor-free frame detection [J]. Computer Engineering and Design, 2025, 46(4): 1157-1166. (in Chinese)
- [15] LAW H, DENG J. CornerNet: detecting objects as paired keypoints[C] // Proceedings of the European conference on Computer Vision. Berlin, Germany: Springer, 2018: 765-781.
- [16] ZHOU X Y, ZHUO J C, KRAHENBUHL P. Bottom-up object detection by grouping extreme and center points[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE Press, 2019: 850-859.
- [17] ZHOU X, WANG D, KRAHENBUHL P. Objects as points [EB/OL]. [2025-07-10]. <https://arxiv.org/abs/1904.07850>.
- [18] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2019: 9627-9636.
- [19] CAESAR H, BANKITI V, LANG A H, et al. nuScenes: a multimodal dataset for autonomous driving[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE Press, 2020: 11621-11631.
- [20] LOSHCHELOV I, HUTTER F. Decoupled weight decay regularization[EB/OL]. [2025-07-10]. <https://arxiv.org/abs/1711.05101>.
- [21] HU P Y, ZIGLAR J, HELD D, et al. What you see is what you get: exploiting visibility for 3D object detection[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE Press, 2020: 11001-11009.
- [22] CHEN Q, SUN L, CHEUNG E, et al. Every view counts: cross-view consistency in 3d object detection with hybrid-cylindrical-spherical voxelization [C] // Proceedings of the Advances in Neural Information Processing Systems. Cambridge, USA: MIT Press, 2020: 21224-21235.
- [23] VORA S, LANG A H, HELOU B, et al. PointPainting: sequential fusion for 3D object detection[C]//Proceedings of

- the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE Press, 2020: 4604-4612.
- [24] YIN J B, SHEN J B, GUAN C Y, et al. LiDAR-based online 3D video object detection with graph-based message passing and spatiotemporal transformer attention [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2020: 11495-11504.
- [25] ZHU X G, MA Y X, WANG T, et al. SSN: shape signature networks for multi-class object detection from point clouds[C]// Proceedings of the European conference on Computer Vision. Berlin, Germany: Springer, 2020: 581-597.
- [26] ZHU B, JIANG Z, ZHOU X, et al. Class-balanced grouping and sampling for point cloud 3d object detection[EB/OL]. [2025-07-10]. <https://arxiv.org/abs/1908.09492>.
- [27] CHEN C, CHEN Z, ZHANG J, et al. SASA: semantics-augmented set abstraction for point-based 3d object detection[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Washington D. C., USA: IEEE Press, 2022: 221-229.
- [28] YANG Z, SUN Y, LIU S, et al. 3DSSD: point-based 3d single stage object detector[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2020: 11040-11048.
- [29] YANG H, WANG W, CHEN M, et al. PVT-SSD: single-stage 3d object detector with point-voxel transformer[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2023: 13476-13487.

文字编辑 索书志
栏目编辑 赖玉玲

计算机工程
www.ecice06.com