

基于改进 PointPillars 的点云车辆目标检测

喻佳祺 杨洪刚 王阳
(上海电机学院机械学院 上海 201306)

摘要:针对自动驾驶领域中,激光雷达获取的点云数据存在稀疏性以及边缘噪点误检等问题,提出一种基于改进 PointPillars 的点云车辆目标检测方法。首先,基于 SimAM 注意力机制改进体素化特征输入,使得网络特征提取阶段能更加关注关键信息,提高特征学习的全局性。其次,基于卷积块注意力模块(CBAM)改进骨干网络结构,提出全新的轻量化通道注意力模块 Tiny-CAM 和可变形空间注意力模块 Deformable-SAM,构建 Multi-CBAM 骨干网络,提升网络特征提取及特征融合能力。在 KITTI 数据集以及非公开车库点云数据集上进行验证,实验结果表明,与原网络相比,改进 PointPillars 方法具有更高的检测精度,平均检测精度提升 2.98%,针对遮挡小于 30% 的点云车辆目标检测精度提升 6.51%,证明了该方法的有效性。

关键词:自动驾驶;目标检测;激光雷达;点云;注意力机制

中图分类号: TP391.4 文献标识码:A 国家标准学科分类代码: 510

Point cloud vehicle target detection based on improved PointPillars

Yu Jiaqi Yang Honggang Wang Yang
(School of Mechanical Engineering, Shanghai Dianji University, Shanghai 201306, China)

Abstract: In the field of autonomous driving, the point cloud data obtained by lidar has problems such as sparsity and edge noise false detection. This paper proposes a point cloud vehicle target detection method based on improved PointPillars. Firstly, the voxelized feature input is improved based on the SimAM attention mechanism, so that the network feature extraction stage can pay more attention to the key information and improve the globality of feature learning. Secondly, based on the improved backbone network structure of CBAM, a new lightweight channel attention module Tiny-CAM and a deformable spatial attention module Deformable-SAM are proposed to construct the Multi-CBAM backbone network and improve the network feature extraction and feature fusion ability. The KITTI data set and the non-public garage point cloud data set are verified. The experimental results show that the method adopted in this paper has higher detection accuracy. Compared with the original network, the average detection accuracy is improved by 2.98 %, and the detection accuracy of point cloud vehicle targets with occlusion less than 30 % is improved by 6.51 %, which proves the effectiveness of the method.

Keywords: autonomous driving; target detection; laser radar; point cloud; attention mechanism

0 引言

随着自动驾驶领域的不断发展,智能感知系统^[1]受到越来越多研究者的关注。点云目标检测^[2]是实现智能车辆感知系统的研究重点。近年来,相对于一般的车载传感器和工业相机,激光雷达的性能逐步提升,具有探测距离远,精度高,鲁棒性强等优点。目前激光雷达数据已广泛应用于自动驾驶、三维重建^[3]、地图构建^[4]、点云配

准^[5]等领域。

近年来,针对三维点云目标检测的研究已成为计算机视觉领域的热点问题之一,吸引众多国际学者的关注。国外研究者在三维点云处理方面取得了一系列创新性成果。通过卷积神经网络(convolutional neural network,CNN)自动学习点云特征表示是早期针对点云数据的代表性工作,由于直接对点云进行处理,有效解决了点云的无序性问题。随后,基于图神经网络(graph neural network,

GNN)的方法被提出来,用于更好地捕捉点云中的局部结构信息。

与传统图像目标检测相比,点云目标检测受限于点云本身的稀疏性和离散性等特点,在检测难度上高于传统图像。但点云具有丰富的拓扑信息和空间深度信息,在执行目标检测任务时可以更加精准地把握空间信息。因此,点云目标检测在自动驾驶领域有着广阔的应用前景。

目前国内外针对点云三维目标检测的算法研究主要分为3种方向。1)基于自取图(bird's eye view, BEV)视图的方法,Li等^[6]提出VeloFCN算法,采用二维检测器缓解了点云中存在遮挡的问题。Beltran等^[7]提出BirdNet算法,将点云投影到三通道鸟瞰图中,三通道依次为高度、强度和密度,并使用Faster R-CNN^[8]在鸟瞰图实现了目标检测。但由于数据的压缩和降维,会在维度转换的过程中丢失部分信息,从而导致网络在对小样本小目标的检测任务中检测效果较差。2)基于空间点(Points)的方法,包括PointNet^[9]、PointNet++^[10]、PointR-CNN^[11]等算法,目的都是为了从原始点中提取空间特征并投入到提取兴趣区域(region of interest, ROI)的分类和检测中。但空间点的数量过于庞大,其计算成本相对较高,训练和部署的速度慢,难以达到实时检测的效果。3)基于体素(Voxel)的方法,Yan等^[12]提出SECOND算法,运用稀疏卷积代替常规的3D卷积提高训练的速度和效果。Deng等^[13]提出Voxel R-CNN在三维特征上聚合了结构信息,使用粗颗粒体素完成高精度检测。Lang等^[14]提出PointPillars算法,将体素改进为柱状划分,并投影到伪图像进行处理,大大加快了检测速度,但在关键特征的学习上存在不足,导致对于稀疏目标以及边缘点云误检率较大。

针对PointPillars模型,詹为钦等^[15]提出采用SSD算法改进检测头模块,达到提高检测速度的目的。田枫等^[16]针对Pillars特征网络提出新的编码模块,提高了每个Pillars的特征表示能力。虽然在算法优化上取得一定成果,但仍然存在若干问题有待解决,其中关键在于保证整体算法轻量的同时,增强对稀疏点云对象的检测效果,同时消除环境中多余噪点的干扰。

针对目前已有的改进思路的不足,进一步降低边缘噪点的误检率和提高稀疏点云数据检测的精度,本文对PointPillars模型的体素化特征输入部分进行改进,通过SimAM注意力机制^[17]使网络在不增加额外参数量的前提下,进一步优化模型计算量,提高整体网络对全局特征的学习效率。同时,基于卷积块注意力模块(CBAM)^[18]提出全新的骨干网络Multi-CBAM,使得网络在提取特征时更加注重点云的关键特征,降低无效信息的学习以及冗余特征的获取。

1 PointPillars 算法

1.1 算法框架及原理

PointPillars算法模型的整体流程为数据增强,点云数据立柱体划分,特征伪图像化,2D特征提取网络,检测头预测。

算法框架主体是基于伪图像处理思想,对点云划分为立柱体后进行特征编码,并输出每一个立柱体中最具代表性的点,将这些特征点构成伪图像,然后使用2D检测框架进行特征提取并对其进行特征框的预测,最后通过训练可获取权重文件用于三维目标检测任务。PointPillars的整体框架流程如图1所示。

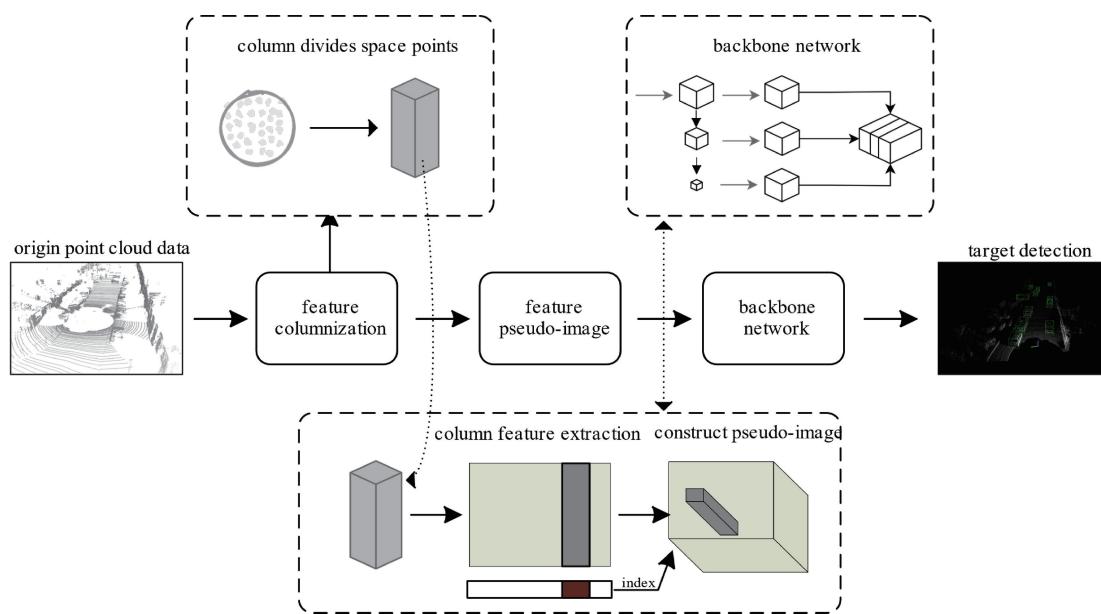


图1 PointPillars 整体框架流程
Fig. 1 PointPillars overall framework process

1.2 损失函数

PointPillars 算法的损失函数需要针对类别分类以及先验框方向分类进行分别计算,同时该算法中每个 GT (ground truth) 框都包含 $(x, y, z, w, l, h, \theta)$ 7 个参数,分别代表长,宽,高,检测框中心点在三维空间的坐标以及检测框的转角。

将参与损失计算的回归残差定义如下:

$$\Delta x = \frac{x^{\text{gt}} - x^{\text{a}}}{d^{\text{a}}}, \Delta y = \frac{y^{\text{gt}} - y^{\text{a}}}{d^{\text{a}}}, \Delta z = \frac{z^{\text{gt}} - z^{\text{a}}}{h^{\text{a}}} \quad (1)$$

$$\Delta w = \log \frac{w^{\text{gt}}}{w^{\text{a}}}, \Delta l = \log \frac{l^{\text{gt}}}{l^{\text{a}}}, \Delta h = \log \frac{h^{\text{gt}}}{h^{\text{a}}} \quad (2)$$

$$\Delta \theta = \sin(\theta^{\text{gt}} - \theta^{\text{a}}) \quad (3)$$

式中: $x^{\text{gt}}, y^{\text{gt}}, z^{\text{gt}}$ 表示标注框的 x, y, z 轴长度; $x^{\text{a}}, y^{\text{a}}, z^{\text{a}}$ 表示先验框的 x, y, z 轴的长度; d^{a} 表示先验框对角线的长度距离。

考虑到上述回归残差得出总回归损失,其计算公式为:

$$L_{\text{loc}} = \sum_{b \in (x, y, z, w, l, h, \theta)} \text{SmoothL1}(\Delta b) \quad (4)$$

针对先验框类别分类任务,该算法采用了 Focal Loss,用于完成正负样本的协调。其计算公式为:

$$L_{\text{cls}} = -\alpha_a(1 - p^a)^{\gamma} \log p^a \quad (5)$$

式中: α 参数和 γ 参数是两个动态参数,此处分别设置为 0.25 和 2。

针对先验框方向分类任务,该算法采用了 Softmax 函数进行方向类别预测。其计算公式为:

$$L = \frac{1}{N_{\text{pos}}} (\beta_{\text{loc}} L_{\text{loc}} + \beta_{\text{cls}} L_{\text{cls}} + \beta_{\text{dir}} L_{\text{dir}}) \quad (6)$$

式中: $\beta_{\text{loc}}, \beta_{\text{cls}}, \beta_{\text{dir}}$ 均为损失函数系数,分别为 2, 1 和 0.2。

2 改进 PointPillars 算法

2.1 改进 PointPillars 模型框架

PointPillars 算法模型沿用 VoxelNet 算法^[19]以及 SECOND 算法相同的数据预处理方式,在数据增强部分和骨干网络部分未对稀疏点云数据采用差异化的处理方案,从而导致模型对于稀疏点云的学习存在困难。而且,在伪图像特征处理部分采用较为传统的二维 CNN 进行上采样及下采样操作,部署效率高但存在检测精度丢失,物体遮挡程度大于 50% 时容易出现对边缘环境噪点的误检。

针对边缘环境噪点误检问题,改进 PointPillars 对数据输入和骨干网络子模块两方面进行完善,模型中每一个 Pillar 特征的输入均通过 SimAM 注意力机制模块,使得网络能更大程度上关注到有价值的输入特征。同时,基于 CBAM 注意力机制将原骨干网络 RPN 改进为 Multi-CBAM 骨干网络,采用轻量型通道注意力模块 Tiny-CAM (Tiny-channel attention module) 和可变形空间注意力模块 Deformable-SAM(deformable spatial attention module)。模型框架示意图如图 2 所示。

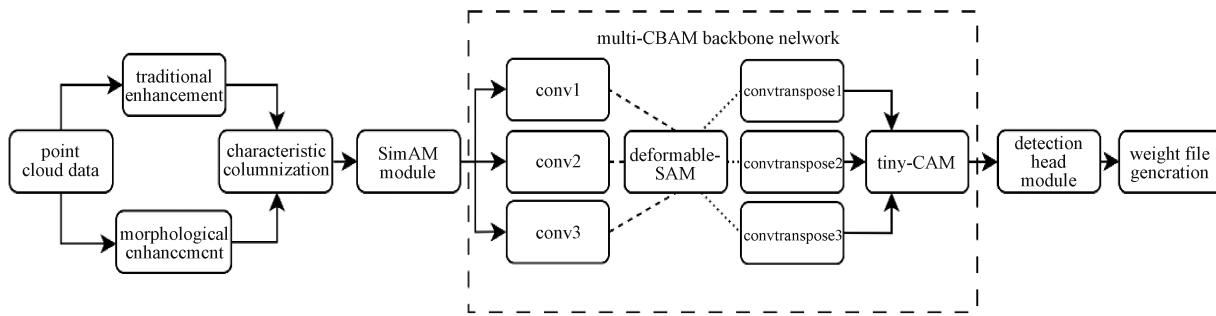


图 2 改进 PointPillars 模型框架示意图

Fig. 2 The improved PointPillars model framework diagram

2.2 SimAM 注意力模块

SimAM 注意力模块是一种基于优化能量函数提出的无参 3D 注意力模块。在网络层中通过能量函数推理出 3D 注意力权重,无需向原始网络增加额外参数,相比于传统低维的注意力模块在计算推理速度上有着较大的提升。同时,SimAM 注意力模块通过 3D 注意力权重对每一个特征点进行函数计算并附上单独的标量,使得最终结果具有全局性和灵活性等优势。

根据优化能量函数的定义,得出最小能量函数为:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{u})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (7)$$

式中: t 为输入特性目标神经元; λ 为函数超参数; 参数 $\hat{\sigma}$ 与 \hat{u} 由式(8)和(9)求得。

$$\hat{u} = \frac{1}{M} \sum_{i=1}^M x_i \quad (8)$$

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{u})^2 \quad (9)$$

式中: x_i 为输入特性目标神经元的相邻神经元。

根据式(7)~(9)得出,特征神经元能量越低,神经元 t 与相邻神经元之间的区别越大,重要性越强。最终通过式(10)对特征进行增强处理。

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (10)$$

为了使模型在保证检测精度及训练速度的同时骨干网络能更有效的学习关键点特征,令骨干网络的数据输入优先通过 SimAM 模块后再送入 Multi-CBAM 骨干网络,提升对目标的检测精度。

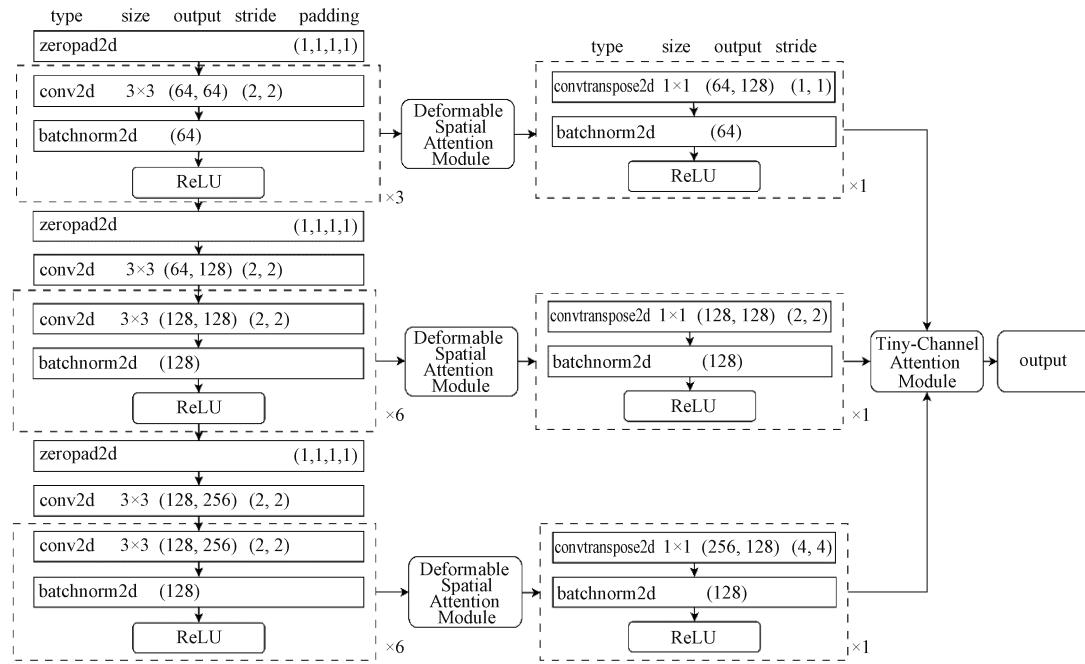


图 3 Multi-CBAM 骨干网络框架

Fig. 3 Multi-CBAM backbone network framework

Multi-CBAM 骨干网络在上采样与下采样之间加入 Deformable-SAM 模块,在保证网络计算量不膨胀的同时,利用可变形卷积在一定程度上缓解因物体形变引起的位置偏移的问题,在特征提取过程中能更加准确地捕捉到目标物体的特征信息。经过骨干网络下采样后,将 3 个 256 维的特征通道叠加送入 Tiny-CAM 模块,最终以一维特征数组的形式用于检测头预测模块。

相比原骨干网络,整体网络的泛化性大幅度提升,网络中自适应池化层和卷积层对不同的特征进行加权计算,使得模型在训练时能捕捉更多有价值的特征信息,从

2.3 Multi-CBAM 骨干网络

为了提高模型对点云车辆目标的检测精度,对二维骨干网络的特征提取部分进行改进,构建 Multi-CBAM 骨干网络,Multi-CBAM 骨干网络框架图如图 3 所示。

而针对各种形态的点云车辆都可以取得较好的特征提取表现。

1) Deformable-SAM 模块

Deformable-SAM 模块将输入特征平均送入全局最大值池化层和全局平均池化层,得到两个 $H \times W \times 1$ 特征层后进行 Concat 拼接,将全部的 7×7 的卷积改为可变性卷积(Deformable Conv),采用可变形卷积策略针对点云目标的形状变化进行自适应,大幅度改善了传统卷积的局限性,最终得到一维空间注意力特征。Deformable-SAM 模块结构如图 4 所示。

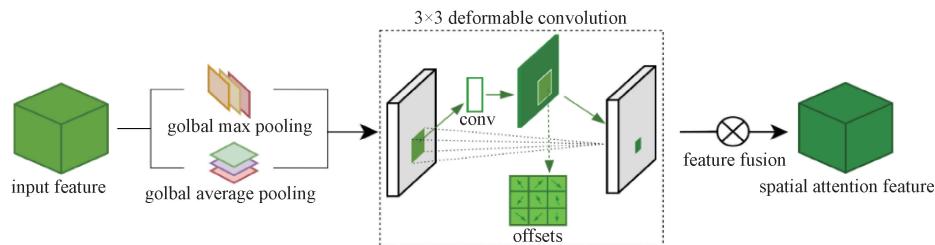


图 4 Deformable-SAM 模块结构

Fig. 4 Deformable-SAM module structure diagram

2) Tiny-CAM 模块

Tiny-CAM 模块将骨干网络输入特征送入全局平

均池化层,将多层感知机(multi-layer perception, MLP)改为 1×1 卷积层并将卷积步长设置为 2,利用卷积方

式大幅度减少了后续网络的计算量,其中激活函数改为替代数值型 ReLU,为网络引入非线性。将卷积层特征送入全局平均池化层和全局最大池化层,最后将两

种池化层通道叠加,使最后结果与原输入特征保持维度一致得到通道注意力特征。Tiny-CAM 模块结构如图 5 所示。

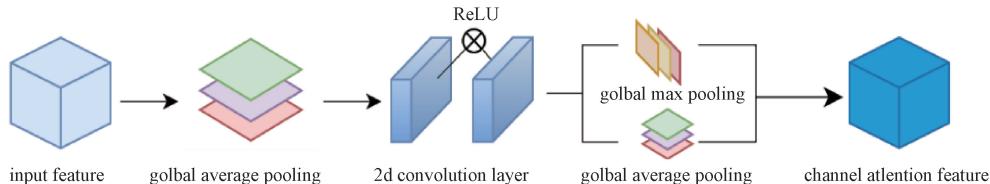


图 5 Tiny-CAM 模块结构

Fig. 5 Tiny-CAM module structure

3 实验结果与分析

3.1 实验数据集

采用两种数据集进行实验分析,数据集 1 为大型公开数据集 KITTI^[20],用于自动驾驶领域研究。该数据集共含点云数据 7 481 帧,其中 3 712 帧用于训练模型,3 769 帧用于模型验证评估。数据集 2 为车库场景下私有数据集,使用镭神智能 C16 传感器进行数据采集,在车库的 6 个不同位置采集 230 份点云数据用于算法模型的可视化验证。

大型公开数据集参考 KITTI 官方测试服务器上的检测指标,检测任务中有 3 种对应难度分别为 Easy、Moderate、Hard,难度越高相应的数据遮挡程度以及稀疏影响则越大,遮挡的最大截断分别为 15%、30%、50%。模型效果通过 4 种评判指标进行综合评价,分别为 2D 检测框精度(bbox),BEV 视图检测框精度(bev),3D 检测框精度(3d)以及检测目标旋转角精度(aos)。

实验环境基于 Ubuntu20.04 操作系统,采用 Pytorch 以及 Openpcdet 框架进行模型调试与编写。CPU 为 I5-12600HF, GPU 为 RTX3060, 内存为 16 GB。模型采用 adam_onecycle 优化器,其在传统 adam 优化器的基础上加入 onecycle 学习率调整策略,从而将学习率控制在合理的区间范围防止训练结果的过拟合。表 1 为训练模型的优化器固定参数。

算法模型采用相同的 adam_onecycle 优化器条件下,对改进后的 PointPillars 模型进行损失函数以及体素大小的调整,最终采用 BinaryCrossEntropy 作为 CLSLoss,

表 1 优化器固定参数

Table 1 Fixed parameters of optimizer

Parameter	Content
BATCH_SIZE_PER_GPU	4
NUM_EPOCHS	80
LR	0.003
WEIGHT_DECAY	0.01
MOMENTUM	0.9
LR_DECAY	0.1
LR_CLIP	1×10^7

SmoothL1 作为 REGLoss。同时分析体素尺寸对算法精度的影响,若将体素尺寸扩大至 [0.32, 0.32, 4],虽然提升了模型的训练速度,但由于点云车辆目标实际尺寸较小,若体素尺寸过大不利于精度优化,为了兼顾模型训练速度和精度两方面,体素体积最终选取 [0.16, 0.16, 3]。后续的消融实验以及实验分析部分验证了该体素尺寸的正确性。

3.2 消融实验

为验证各个模块的有效性,设置消融实验,每组模型在 KITTI 数据集以及私有地库数据集上进行 250 轮训练,使用 Easy 以及 Moderate 难度 R40 标准的 AP_{3D} 作为评价模型的性能衡量指标。所有实验均为采用预训练权重且保证优化器以及损失函数采用相同的类别。实验编号以及实验结果如表 2 所示,其中“√”表示本组实验已使用该模块。

表 2 消融实验结果数据

Table 2 Ablation experimental results data

实验编号	PointPillars	SimAM	Deformable-SAM	Tiny-CAM	Easy	Moderate
A	√				90.72	89.53
B	√	√			95.03	90.26
C	√		√		90.77	88.47
D	√			√	93.97	89.89
E	√		√	√	95.34	90.42
F	√	√	√	√	96.68	91.15

对比消融实验 A 与 B 可知,若单独采用 SimAM 模块对数据输入阶段进行单独优化,同时保留原 PointPillars 的骨干网络部分,验证无参 3D 注意力模块 SimAM 模块有效提升模型的 3D 检测性能;对比消融实验 A、C 与 D 可知,单独采用结构优化后的 Deformable-SAM 模块与 Tiny-CAM 模块分别投入骨干网络中,提升了 Easy 难度下的检测精度但在 Moderate 难度下效果一般,考虑环境遮挡对模块的效果造成了一定影响。对比消融实验 A 与 E 可知,采用子模块融合而成的 Multi-CBAM 骨干网络代替原有骨干网络使得模型泛化性大幅度提升,提升了车辆点云目标的检测框准确率。对比消融实验 A 与 F 可知,实验组 F 为本文的最终模型,结合 SimAM 模块和 Multi-CBAM 骨干网络的结构优势,大幅优化算法模型的检测性能与特征学习效率。

表 3 改进 PointPillars 模型在 KITTI 车辆验证集定量评估
Table 3 Quantitative evaluation of improved PointPillars model in KITTI vehicle validation set (%)

AP _{3d} (R11 standard)			AP _{bev} (R11 standard)			AP _{3d} (R40 standard)			AP _{bev} (R40 standard)		
easy	mod	hard	easy	mod	hard	easy	mod	hard	easy	mod	hard
90.75	89.61	80.92	90.78	89.80	89.11	96.68	91.15	86.03	96.82	93.48	88.44
AP _{bbox}											
AP _{bbox} (R11 standard)			AP _{aos} (R11 standard)			AP _{bbox} (R40 standard)			AP _{aos} (R40 standard)		
easy	mod	hard	easy	mod	hard	easy	mod	hard	easy	mod	hard
90.66	88.86	86.88	90.60	88.54	86.39	96.02	91.73	87.16	95.94	91.38	86.70

2) 算法精确度对比

在 R40 标准下,将算法改进后所得到的 AP_{3d} 值与主流的三维目标检测算法进行对比。在检测难度为简单、中等、困难 3 种情况下,算法模型的检测准确度均有所提升,且在简单难度下对于点云车辆的检测精确度相较于同类模型有着明显领先。相较于原算法模型,在 3 种难度下平均准确度 AP_{3d} 性能分别提升了 6.51%、1.71%、0.72%。图 6 为改进的 PointPillars 模型与 Voxel、Voxel R-CNN、SECOND、PointRCNN、Pointpillars 算法模型的检测精确度对比柱状图。

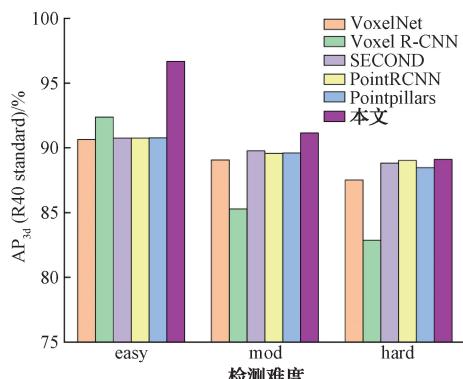


图 6 算法对比柱状图
Fig. 6 Algorithm comparison histogram

3.3 实验分析

1) 算法定量分析

针对改进 PointPillars 模型在 KITTI 数据集以及车场场景下私有数据集进行定量评估与分析。实验评估采用 R11 以及 R40 两种计算标准,其中 R11 标准采用 11 点插值法计算检测精度—召回率曲线的区域面积,R40 标准采用 40 点插值法计算检测精度—召回率曲线的区域面积。以上两种评估标准为 KITTI 数据集针对前向视角点云车辆数据的官方度量标准。改进的 PointPillars 模型在两种评估标准下的 3D 目标检测精度、鸟瞰视角目标检测精度、2D 检测框精度以及检测目标旋转角精度如表 3 所示。本文所有实验设定交并比(intersection of union, IoU)标准如下:Easy、Moderate 以及 Hard 3 种检测难度下 IoU 分别为 0.7、0.5、0.5。该算法在验证集上每帧用时 0.046 s,满足当下点云实时检测的处理标准。

R-CNN、SECOND 等三维目标检测算法进行比较的结果。

在 R11 标准下,进一步验证本文所提出的算法模型在俯瞰图中的精确度,将改进后的算法模型与同样基于体素思想的三维目标检测算法进行对比。如表 4 所示,改进后的 Pointpillar 模型在 AP_{bev} 精确度指标对比中相较于其他算法模型在中等和困难难度下 AP_{bev} 有着明显的提高,证明改进后的算法模型对 BEV 的感知能力得到进一步提升。

表 4 AP_{bev} 精确度对比
Table 4 Comparison of AP_{bev} accuracy

算法模型	点云车辆对象		
	Easy	Moderate	Hard
MV3D	85.8	77.0	68.9
PointPillars	88.4	86.1	79.8
VoxelNet	89.4	79.3	77.4
SECOND	88.1	79.4	77.9
本文	90.7	89.8	89.1

3) 算法推理时间对比

图 7 所示为 KITTI 数据集每批次网络推理时间的对

比,以评估算法模型的训练效率和轻量化程度。相较于传统体素化算法 VoxelNet,改进后的 PointPillars 模型在控制每批次处理数据的条件下,推理速度提升了 58%,在保证平均精确度和推理速度的同时,整体网络结构依然凸显轻量化和高效率的优势。

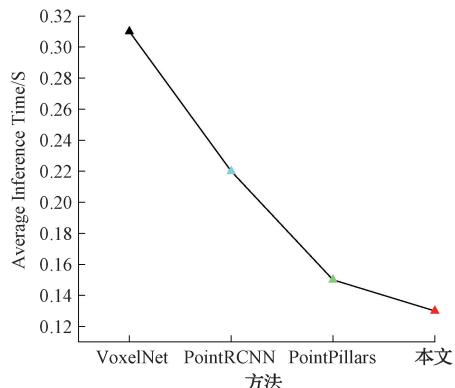


图 7 网络推理时间对比

Fig. 7 Comparison of network inference time

3.4 结果可视化分析

改进 PointPillars 在 KITTI 数据集的检测效果如图 8 所示。目标检测效果图中以绿色 3D 框表示汽车的 3D 检测框,红框表示可视化分析中存在的异常。

从图 8 可以看出,原算法(图 8(a))将距雷达外 50 m 的一处噪点判定为车辆并进行框选,表明原算法在针对远距离的点云对象会产生误判的情况,改进 PointPillars 算法(图 8(b))则具备较好的可视化表现,在准确定位车辆位置的同时避免了对于边缘噪声的误判,验证了改进算法模型在避免噪点影响的同时,检测精度进一步提升。

(a) 原PointPillars模型
(a) The original PointPillars model(b) 改进PointPillars模型
(b) Improved PointPillars model

图 8 KITTI 数据集的检测效果对比

Fig. 8 Comparison of detection effects of KITTI

改进 PointPillars 在车库场景下私有数据集中的检测

效果如图 9 所示。图 9(a)为采用原算法对私有数据集的可视化效果,可以看出未能成功检测出车辆目标,主要原因如下:1)私有数据集采用十六线激光雷达进行数据采集,获取点云数据的稀疏性增加了检测的难度;2)PointPillars 模型特征处理部分采用较为传统的二维 CNN 对于图 9 中稀疏的点云车辆数据极其容易存在漏检情况。由图 9(b)可以看出,改进 PointPillars 算法能够精确识别点云车辆,三维坐标与转角信息没有出现明显的错误偏移,凸显了改进后算法面对稀疏点云数据的检测效果更具优势。

综上所述,改进 PointPillars 在针对较远距离噪点的误检以及稀疏点云数据检测上都具备较好的检测效果,从而提升了整体的检测效果。

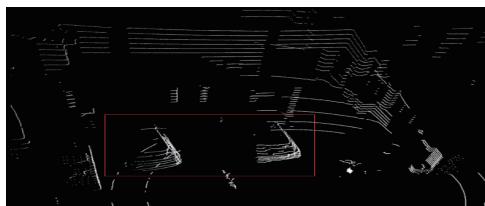
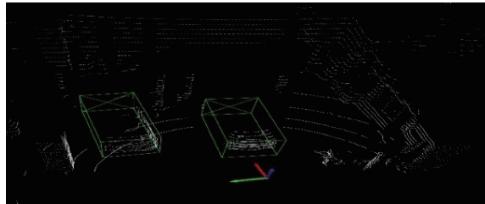
(a) 原PointPillars模型
(a) The original PointPillars model(b) 改进PointPillars模型
(b) Improved PointPillars model

图 9 车库场景下私有数据集检测效果对比

Fig. 9 Comparison of detection effects of private data sets in garage scene

4 结 论

针对激光雷达获取的点云数据存在稀疏性以及边缘噪点误检等问题,本文提出一种改进 PointPillars 的三维目标检测算法。通过引入 SimAM 注意力机制,改进了 PointPillars 中体素化特征输入部分,对每一个投入到骨干网络的特征赋上单独标量,减少参数量,提高特征学习全局性。进一步提出全新的轻量化通道注意力模块 Tiny-CAM 和可变形空间注意力模块 Deformable-SAM,构建 Multi-CBAM 骨干网络替代原有骨干网络,在保证轻量化的同时,提高网络学习关键信息的能力。实验表明,改进 PointPillars 在 Easy、Moderate 和 Hard 3 种难度下,与原算法模型相比 3D 精度分别提升 6.51%、1.71%、0.72%,平均检测精度提升 2.98%,每帧数据检测时间为 0.046 s,满足实时目标检测的处理标准,在点云车辆检测领域具有重要意义。后续研究将进一步针对三维目标检测骨干网络的设计进行深入挖掘,并将模型应用到复杂环境的检测场景中。

参考文献

- [1] 朱浩,倪锐峰.基于点云稀疏语义特征的智能网联汽车协同感知配准算法[J].仪器仪表学报,2023,44(10): 314-324.
ZHU H, NI R F. An intelligent connected vehicle collaborative perception registration algorithm based on sparse semantic features of point cloud [J]. Chinese Journal of Scientific Instrument, 2023, 44(10): 314-324.
- [2] LI Y F, HUANG Y P. Target detection based on the fusion of lidar and camera [J]. Electronic Measurement Technology, 2021, 44(5): 112-117.
- [3] 刘青青,毋涛.基于RealSense人体多视角三维重建方法的研究[J].国外电子测量技术,2021,40(10):77-83.
LIU Q Q, WU T. Research on the method of multi-view 3D reconstruction of human body based on RealSense [J]. Foreign Electronic Measurement Technology, 2021, 40(10): 77-83.
- [4] 王存华,王伟.基于模拟退火优化BP算法的指纹地图构建方法[J].国外电子测量技术,2020,39(3):17-24.
WANG C H, WANG W. Fingerprint map construction method based on simulated annealing optimization BP algorithm [J]. Foreign Electronic Measurement Technology, 2020,39(3): 17-24.
- [5] 傅瑶,陈鹏,郭贵松,等.基于4PCS和SICP的点云配准方法在钢轨磨耗计算中的应用[J].电子测量与仪器学报,2022,36(12): 210-218.
FU Y, CHEN P, GUO G S, et al. Application of point cloud registration method based on 4PCS and SICP in rail wear calculation[J]. Journal of Electronic Measurement and Instruments, 2022, 36(12): 210-218.
- [6] LI B, ZHANG T, XIA T. Vehicle detection from 3d lidar using fully convolutional network [J]. ArXiv preprint arXiv: 1608.07916, 2016.
- [7] BELTRÁN J, GUINDEL C, MORENO F M, et al. Birdnet; A 3D object detection framework from lidar information[C]. 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018: 3517-3523.
- [8] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39 (6): 1137-1149.
- [9] QI C R, SU H, MO K, 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, 2017: 652-660.
- [10] QI C R, YI L, SU H, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space[J]. Advances in Neural Information Processing Systems, 2017, DOI:10.48550/arXiv.1706.02413.
- [11] ZHOU Q, YU C. Point RCNN: An angle-free framework for rotated object detection[J]. Remote Sensing, 2022, 14(11): 2605.
- [12] YAN Y, MAO Y, LI B. Second: Sparsely embedded convolutional detection[J]. Sensors, 2018, 18(10): 3337.
- [13] DENG J, SHI S, LI P, et al. Voxel R-CNN: Towards high performance voxel-based 3d object detection[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(2): 1201-1209.
- [14] 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, 2019: 12697-12705.
- [15] 詹为钦,倪蓉蓉,杨彪.基于注意力机制的PointPillars+三维目标检测[J].江苏大学学报(自然科学版),2020,41(3):268-273.
ZHAN W Q, NI R R, YANG B. PointPillars + 3D object detection based on attention mechanism [J]. Journal of Jiangsu University (Natural Science Edition), 2020,41(3): 268-273.
- [16] 田枫,刘超,刘芳,等.基于改进PointPillars的激光雷达三维目标检测[J].激光与光电子学进展, 2024, 61(8):0812007.
TIAN F, LIU CH, LIU F, et al. Laser radar three-dimensional target detection based on improved PointPillars [J]. Progress in Laser and Optoelectronics, 2024,61(8): 0812007.
- [17] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks [C]. International Conference on Machine Learning. PMLR, 2021: 11863-11874.

- [18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [19] ZHOU Y, TUZEL O. Voxelnet: End-to-end learning for point cloud based 3d object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4490-4499.
- [20] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: The kitti dataset [J]. The International Journal of Robotics Research, 2013, 32(11): 1231-1237.

作者简介

喻佳祺,硕士,主要研究方向为三维点云目标检测。
E-mail: yujiaqifinn@163.com
杨洪刚(通信作者),博士,副教授,硕士生导师,主要研究方向为目标识别与检测等。
E-mail: yanghg@sdju.edu.cn