Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2025, 14(6), 1424-1432
Published Online November 2025 in Hans. https://www.hanspub.org/journal/airr
https://doi.org/10.12677/airr.2025.146133

Hans汉斯

基于路端激光雷达的多目标检测算法研究

许正尧

西华大学汽车与交通学院，四川 成都

收稿日期：2025年10月13日；录用日期：2025年11月6日；发布日期：2025年11月14日

摘　要

为了解决路端点云目标检测任务中复杂场景下远距离检测精度差，内存和计算开销大等问题，现有的基于点的检测器常采用随机采样和最远点采样对输入点云进行下采样，忽略了前景点的重要性。本文提出了一种面向路端的基于点的单阶段三维目标检算法。本文算法采用了一种面向自动驾驶路端检测任务的实例感知下采样策略，分级选择检测对象的前景点。此外，算法引入中心感知模块来进一步估计精确的实例中心，提高检测精度，改善误检问题。为了验证算法的性能，在自动驾驶路端数据集DAIR-V2X-I上进行实验。实验结果表明，所提算法相比其它公开算法检测准确率具有一定的提升，同单阶段检测算法3DSSD相比，在中等检测难度下，对DAIR-V2X-I数据集汽车、行人、骑行者的检测准确率分别提高了0.99%，2.4%和4.7%。

关键词

目标检测，实例感知，误检漏检，中心感知

Research on Multi Object Detection Algorithm Based on Roadside Lidar

Zhengyao Xu

School of Automotive and Transportation, Xihua University, Chengdu Sichuan

Received: October 13, 2025; accepted: November 6, 2025; published: November 14, 2025

Abstract

In order to solve the problems of poor remote detection accuracy, high memory and computational overhead in complex scenes of endpoint cloud object detection tasks, existing point based detectors often use random sampling and farthest point sampling to downsample the input point cloud, ignoring the importance of foreground points. We introduce a novel point-based and single-stage algorithm for 3D object detection at the road end. The key to this algorithm is to use two learnable,

文章引用: 许正尧. 基于路端激光雷达的多目标检测算法研究[J]. 人工智能与机器人研究, 2025, 14(6): 1424-1432.
DOI: 10.12677/airr.2025.146133

task oriented, instance aware downsampling strategies to hierarchically select the foreground points of the object of interest. In addition, a context center aware module is introduced to further estimate the accurate instance center, improve detection accuracy, and address false positives. To validate the performance, we conducted experiments on the DAIR-V2X-I autonomous driving road-side dataset. Results demonstrate that our algorithm achieves superior detection accuracy over other public benchmarks. Compared to the single-stage detector 3DSSD, our method achieves gains of 0.99%, 2.4%, and 4.7% in AP for cars, pedestrians, and cyclists, respectively, on the DAIR-V2X-I dataset under moderate difficulty.

Keywords

Object Detection, Instance Perception, Misdetection and Omission, Central Perception

1. 引言

我国在推动智能网联汽车发展进程中，已将车路协同纳入顶层设计。与之相应，路端感知系统也构成了实现该战略的核心基石，其重要性不言而喻[1]。依托路端感知系统，可实时采集更完整的交通动态信息，这一能力能有效填补单车智能在远距离感知稳定性不足、存在感知盲区等方面的短板。三维目标检测作为环境感知的关键任务，在汽车工业领域和学术领域引起了广泛的关注，各类先进感知算法的提出为高效的智能交通提供了坚实的技术基础。

基于深度学习的三维目标检测方法相较于传统方法具有较强的鲁棒性，能够更好地适应计算量较大的场景，在目标检测领域中已经占据了主导地位[2]。基于深度学习的三维目标检测算法主要分为三类：基于图像、基于点云和基于多模态数据融合。

基于图像的方法首先在二维平面内进行检测，随后通过深度估计等方法将结果映射到三维空间，代表模型为 SMOKE [3]、FCOS3D [4]、FCOS3D++ [5]等。这种方法能够获取丰富的特征信息且成本较低，但容易受到光照条件的影响，且无法适应远距离检测任务。

基于点云的目标检测方法具有较大的感知范围且不受光照条件影响，可以全天候的进行检测任务。现阶段，基于点云的目标检测方法大致可分为两类：一类是基于点的处理方法，另一类是基于体素的方法。对于基于点的检测网络而言，其会直接对原始点云进行特征提取，随后构建出 3D 检测框。PointNet [6]、PointNet++ [7]使用多层感知机学习点云特征学习点云特征为基于点的检测方法提供了可行的思路，最大限度的保留了点云的几何信息，有助于精确捕捉小目标。但这种方法计算量较大、成本高，且容易受到点云密度不均匀的影响，导致检测性能不稳定。基于体素的 3D 检测器一般会先对非结构化点云做预处理，将其转变为紧凑的规则柱体或体素网格，随后通过 3D 或 2D 卷积神经网络提取特征，最终生成 3D 目标候选框[8]，典型模型包括 VoxelNet [9]、SECOND [10]、PointPillars [11]等。这种方法提高了点云数据的处理效率，降低了计算复杂度，但体素化操作可能导致细节信息丢失，影响对小目标的检测效果。基于多模态融合的三维目标检测将图像和点云两种数据相结合，充分利用了图像的语义信息和点云的深度信息，典型模型包括 PointFusion [12]、MLOD、AVOD-Net [13]等。这种方法提高算法的鲁棒性和检测精度，适合复杂场景下的检测任务，但硬件成本高、数据量大，对系统的标定和同步提出了更高的要求。

经过综合考虑，本文使用路端激光雷达进行三维目标检测。依据当前复杂的交通情况以及点云的稀疏性和无序性，现有的基于点云的目标检测算法难以胜任十字路口场景下远距离车辆和行人的检测任务，这对自动驾驶的大规模商业化落地产生了一定的影响。为了进一步提高在交通路口场景下对远距离车辆和行人等目标的检测精度，本文提出了一种基于点的单阶段、可实现端到端训练的路端三维目标检测算法。为了能够在交通路口场景下的稀疏点云中提取更多有效的特征信息，本文采用了一种面相任务的、针对实例的感知下采样方法，能够在减少内存占用和计算开销的同时分层地提取更重要的前景点信息。另外，为了能够进一步精确地预测实例目标的中心，本文还引入了一种联合上下文信息的中心感知模块。为了能够提高远距离目标的检测精度，提升检测效率，选择采用了稀疏检测头。所提算法在公开路端数据集 DAIR-V2X-I [14]上的实验结果表明了算法的优越性。

2. 理论分析

2.1. 网络结构

如图 1 所示为本文所提算法的网络结构图，主要针对路端检测任务进行设计优化。算法主要由点云特征提取网络、中心感知模块和候选框检测头等模块组成。该方法首先将路端原始点云进行下采样，然后将点云送入到几个集合抽象层提取点的特征，通过进行实例感知下采样，保留前景点的同时逐步降低内存成本，提高计算效率。将提取到的特征输入到中心感知模块进行中心预测和聚合，用于确定目标中心位置。最后，使用三维目标检测头来实现目标分类以及三维检测框的位置、尺寸大小和目标朝向角等参数的回归。

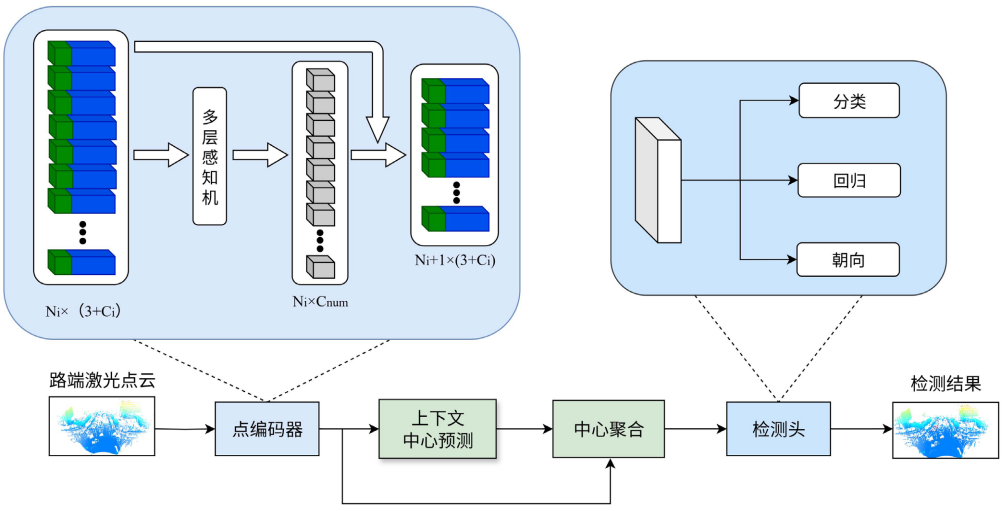

Figure 1. Network structure diagram
图 1. 网络结构图

2.2. 实例感知下采样方法

交通路口场景下的目标检测任务点云数据规模庞大，为了实现高效的目标检测任务，需要通过逐步下采样来减少内存和计算开销。然而，较为激进的下采样方法可能会导致大量前景点信息的丢失，降低检测精度。通过相关研究发现，经过多次随机下采样之后，实例召回率显著下降，表明大量前景点丢弃。基于欧几里得距离的下采样(D-FPS)和基于特征距离的下采样(Feat-FPS)在早期阶段都能获得较好的实例召回率，但在最后一层编码时仍会丢失大量的前景点。在路端远距离检测任务中，对于行人和非机动车

这样的小目标来说检测精度会有所下降。为了能够尽可能的保留前景点，本文选择合理的利用每个点的潜在语义信息，因为学习到的点特征可能包含更丰富的语义信息，分层聚合在每个层中运行。根据这一思想，本文采用以下两种面向任务的采样方法，即将前景语义先验知识引入到网络训练管道中。

2.2.1. 实例感知采样

实例感知采样的流程如下：首先，在逐点语义学习的基础上执行选择性下采样。为增强此过程的有效性，网络额外引入一个辅助分支，该分支的作用是充分挖掘潜在特征中的语义信息。随后，编码层叠加两个 MLP 层，以估计每个点的语义类别。在监督方面，模型使用从原始边界框标注转换得到的逐点语义标签作为真值，并采用常规交叉熵损失进行优化。

$$L_{cls-aware} = -\sum_{c=1}^{C}\left(s_i \log\left(\hat{s}_i\right) + \left(1-s_i\right)\log\left(1-\hat{s}_i\right)\right) \tag{1}$$

式中，C 代表类别的总数量，s_i 则对应关键标签。在推理阶段，算法会筛选并保留前景得分排名前 k 的点，将这些点作为代表点传入下一个编码层，最终实现较高的实例召回率。

2.2.2. 中心感知采样

考虑到实例中心位置估计是目标检测任务的关键，本文选用中心感知下采样策略，在该策略中，距离实例中心越近的点，所获权重越高。相应地，实例 i 的软掩码可定义如下：

$$Mask_i = \sqrt[3]{\frac{\min\left(f^*, b^*\right)}{\max\left(f^*, b^*\right)} \times \frac{\min\left(l^*, r^*\right)}{\max\left(l^*, r^*\right)} \times \frac{\min\left(u^*, d^*\right)}{\max\left(u^*, d^*\right)}} \tag{2}$$

式中，f^*，b^*，l^*，r^*，u^*，d^* 六个变量，分别编码了点与边界框前、后、左、右、上、下各表面的距离。据此，我们得到一个归一化的遮罩值：该值在点位于边界框质心处时取得最大值 1，并随着点向边界框表面移动而逐渐减小，直至在边界框表面处为 0。在训练过程中使用掩码根据空间位置对检测框内的点赋予不同的权重，从而隐式地将几何先验融合到网络训练中。其中加权交叉熵损失计算公式如下：

$$L_{ctr-aware} = -\sum_{c=1}^{C}\left(Mask_i \cdot s_i \log\left(\hat{s}_i\right) + \left(1-s_i\right)\log\left(1-\hat{s}_i\right)\right) \tag{3}$$

通过将掩码与前景点损失项相乘，能够提高靠近中心区域的点的概率。值得注意的是，在推理阶段，边界框并非必要条件。当模型训练达到预期标准后，仅需从下采样后的结果中筛选并保留得分最高的前 k 个点。

2.3. 上下文实例中心感知

2.3.1. 上下文中心预测

受 2D 图像检测领域中上下文预测技术的成功案例启发，进而尝试借助边界框周边的上下文线索开展中心预测任务。具体而言，实例中心偏移的预测遵循[15]的方法来确定。与此同时，为最小化中心预测的不确定性，算法中额外引入了正则化项。具体来说，每个实例的所有投票都是根据周围的干扰进行聚合的，其中 c_i 是第 i 个实例的平均目标。中心预测损失函数如下：

$$L_{cent} = \frac{1}{\left|F_+\right|}\frac{1}{\left|S_+\right|}\sum_i\sum_j\left(\left|\Delta c_{\hat{ij}} - 1\Delta c_{ij}\right| + \left|c_{\hat{ij}} - \overline{c_i}\right|\right)\cdot I_S\left(p_{ij}\right) \tag{4}$$

$$where\ \ \overline{c_i} = \frac{1}{\left|S_+\right|}\sum_j c_{\hat{ij}}, \mathfrak{I}_S : P \to \{0,1\} \tag{5}$$

其中，$\Delta c_{\hat{ij}}$ 表示从点 p_{ij} 到中心点的地面真值偏移，I_S 是一个指标函数，用于确定该点是否用于估计实例中心。$\left|S_+\right|$ 用于预测实例中心的点数。

2.3.2. 基于中心的实例聚合

首先，通过 PointNet++模块从偏移后的中心点学习实例的潜在表示。具体地，将相邻点转换至局部规范坐标系后，利用共享 MLP 结合对称函数完成特征聚合；最终，将聚合后的中心点特征输入提案生成头，完成分类边界框的预测。提案需被编码为涵盖位置、大小及方向的多维表示形式，最后借助设定了特定 IOU 阈值的 3D NMS 后处理操作，实现结果过滤。

2.4. 损失函数

在训练过程中，由于框架中使用了多任务损失来进行联合优化，所以总损失 L_{total} 由下采样策略损失 L_{sample}、中心预测损失 L_{cent}、分类损失 L_{cls} 和包围框生成损失 L_{box} 四部分组成：

$$L_{total} = L_{sample} + L_{cent} + L_{cls} + L_{box} \tag{6}$$

此外，包围框生成损失可以进一步分解为位置损失、尺寸损失、角度箱损失、角度分辨率损失、角点损失等组成：

$$L_{box} = L_{loc} + L_{size} + L_{angle-bin} + L_{angle-res} + L_{corner} \tag{7}$$

3. 实验

3.1. 实验数据集

本文选择使用大规模开源路端 3D 检测数据集 DAIR-V2X-I 进行算法的评估和验证，数据集由多个十字路口真实场景下的图像数据和激光雷达点云数据组成。该数据集包含 10084 个图像和点云样本，其中公开训练集及验证集共 7481 个样本[16]，主要类别有小汽车、自行车、行人等共计 15 类。该数据集依据目标的截断程度与遮挡程度，将所有目标划分为简单、中等、困难三个不同的检测等级。在本次实验中，模型训练阶段使用训练集输入数据，后续的实验验证环节则依赖验证集完成结果校验。

现有的目标检测算法及框架大部分基于 KITTI 数据集数据格式进行开发测试，这里用到的 DAIR-V2X-I 路端感知数据集与 KITTI 数据集在数据结构及格式上存在较大差异，为了方便开展实验研究，将 DAIR-V2X-I 路端感知数据集转化为应用更加普遍的 KITTI 数据集格式。并且，对数据集 DAIR-V2X-I 原有的 10 分类进行合并操作，将手推车(Barrowlist)、三轮车(Tricyclist)和摩托车(Motorcyclist)与骑行者(Cyclisy)进行合并同时将标注形式转换为 KITTI 数据集标注形式，方便后续的数据读取和评价指标计算。

3.2. 实验设置

本实验依托 3D 目标检测框架 OpenPCDet 开展，所采用的软硬件环境配置如下：硬件方面，搭载 AMD EPYC 7542 32 核处理器(主频 2.89 GHz)、NVIDIA GeForce RTX 3090 显卡(24G 显存)及 54 GB 内存；软件方面，系统为 Ubuntu 20.04.2 LTS，配套 Python 3.7、Cuda 11.2 与 Pytorch 1.10.0。采用 Adam 优化器，权重衰减值为 0.01，动量为 0.9，学习率衰减值为 0.1，批次大小为 4，最大迭代次数为 80。根据数据集的特点设置点云范围、包围框大小，以及中心高度，对汽车(Car)、行人(Pedestrian)和骑行者(Cyclist)这 3 类目标进行检测。

3.3. 实验结果及分析

所提算法在路端感知数据集 DAIR-V2X-I 上进行实验，然后使用平均精度值(AP)和每秒帧数(FPS)对

3D 目标检测性能进行评估。实验中 IOU 阈值按目标类别差异化设置：汽车类别的 IOU 阈值设定为 0.7，行人与骑行者两类目标的 IOU 阈值则统一设置为 0.5。为了评估所提算法的性能，将此文所提算法与另外三种经典检测算法在 DAIR-V2X-I 验证集上进行对比试验，实验结果如表 1 所示。其中加粗字体代表最佳检测结果。

Table 1. Comparison results between the proposed algorithm and classical algorithms on the DAIR-V2X-I validation set
表 1. 所提算法与经典算法在 DAIR-V2X-I 验证集上的对比结果

	汽车			行人			骑行者			mAP	FPS
	简单	中等	困难	简单	中等	困难	简单	中等	困难		
SECOND	87.73	77.91	78.03	60.81	55.68	55.86	**80.62**	61.80	61.63	65.32	25.71
3DSSD	87.15	77.03	77.01	56.42	55.53	55.51	76.62	57.15	57.03	62.13	32.54
PointPillars	87.28	77.86	77.97	59.12	57.53	**57.96**	80.16	60.32	59.73	65.43	30.32
本文	**87.88**	**78.02**	**78.10**	**60.92**	**57.93**	56.54	80.32	**61.93**	**61.91**	**65.51**	**36.74**

由表 1 可知：所提算法在 DAIR-V2X-I 验证集上进行实验，与 3DSSD 相比，在汽车、行人、骑行者的检测精度在简单检测难度下分别提高了 0.73 百分点、4.5 百分点、3.7 百分点，在中等检测难度下分别提高了 0.99 百分点、2.4 百分点、4.9 百分点，在困难检测难度下分别提高了 1.09 百分点、1.03 百分点、4.88 百分点，mAP 提高了 3.38 百分点，FPS 提高了 4.2 fram/s。

所提算法在 DAIR-V2X-I 验证集上对不同类别目标均取得了一定的检测精度提升，尤其是在行人和骑行者两类小目标上提升幅度更为显著。这一结果反映出本文设计的实例感知下采样策略在保持整体检测性能的同时，对小体积目标的特征表达具有更强的鲁棒性。其原因主要体现在以下三个方面。

首先，行人和骑行者目标在点云中往往只包含极少的点数，且这些点受距离、遮挡和扫描角度的影响较大。传统的随机采样(RS)和基于几何距离的最远点采样(D-FPS)策略容易在多层下采样过程中丢失这些稀疏但关键的前景点，导致特征学习阶段缺乏语义显著性。而本文提出的实例感知下采样方法通过引入语义先验，引导网络在采样时优先保留具有前景属性的点，从源头上降低了前景点被误舍弃的概率。

其次，本文在点特征提取过程中引入了多层语义聚合机制，使得模型能够在较深层特征空间中保留来自不同尺度的语义线索。对于小目标而言，局部几何特征的完整性至关重要，多尺度聚合使模型在特征表达上更具判别性，从而提高了对远距离及被遮挡小目标的识别能力。

最后，中心感知模块的引入在特征回归阶段进一步增强了对目标中心位置的建模能力。对于尺寸较小、形态差异大的行人和骑行者类别，准确估计中心点可以有效抑制由点云不均匀分布带来的误检与偏移，提高了整体定位精度和边界一致性。

综上，所提算法通过实例感知下采样与中心感知机制的协同作用，显著改善了传统方法在小目标检测任务中的特征丢失与定位偏差问题，从而在复杂交通场景下实现了更优的检测性能和更强的泛化能力。

3.4. 可视化结果与分析

为直观展示所提算法在复杂交通路口场景下对不同类别目标(尤其是远距离小目标)的检测能力，本文在 DAIR-V2X-I 路端感知数据集上对检测结果进行可视化分析，结果如图 2 和图 3 所示。其中，红色框标注汽车检测结果，蓝色框标注行人检测结果，绿色框标注骑行者检测结果，点云视图与对应真实场景图像的联动展示可清晰反映算法的检测效果。

图 2 上半部分为点云场景的检测结果可视化，下半部分为对应真实路侧场景图像。在该交通路口场景中，无论是路口待行区的近距离车辆，还是道路远端的车辆，红色检测框均能准确包围目标，体现了

算法对不同距离汽车目标的鲁棒检测能力；对于斑马线上的行人(蓝色框)和非机动车道的骑行者(绿色框)，即使目标在场景中像素占比小、点云点数少，算法仍能精准识别并标注，验证了实例感知下采样策略对小目标前景点的有效保留。

Figure 2. Detection results in DAIR-V2X-I roadside scenes 1
图 2. DAIR-V2X-I 路侧场景 1 检测结果

Figure 3. Detection results in DAIR-V2X-I roadside scenes 2
图 3. DAIR-V2X-I 路侧场景 2 检测结果

图 3 同样采用“点云检测视图 + 真实场景图像”的展示形式，聚焦另一典型交通路口场景。在车辆密集、行人与骑行者穿插的路口环境中，算法对汽车、行人、骑行者的检测框无明显误检、漏检情况，且检测框与目标的空间贴合度较高，说明中心感知模块有效提升了实例中心估计的精确性；针对道路远端的小型车辆和行人，算法仍能稳定输出检测结果，进一步佐证了本文下采样策略在保留远距离前景点信息上的优势——相比传统采样易丢弃小目标点云的问题，实例感知下采样可针对性保留行人、骑行者等小目标的关键语义点，从而在可视化结果中体现出对这类目标更突出的检测性能提升。

4. 结论

引用实例感知下采样策略和实例中心估计，提出一种单阶段三维目标检测算法，旨在解决现阶段点云目标检测算法对远距离目标检测精度差、误检、漏检等问题。为了能够提高在十字路口场景下的检测精度，采用了实例感知下采样策略和中心感知模块。本研究在 DAIR-V2X 路端数据集上开展实验，验证了所提算法的有效性。结果显示，该算法对三类目标的检测精度均有提升，且检测速度达到 36.74 FPS，满足实时性需求，实现了检测精度与检测速度的良好平衡。

参考文献

[1] Yu, H.B., Luo, Y.Z., Shu, M., *et al*. (2022) DAIR-V2X: A Largescale Dataset for Vehicle-Infrastructure Cooperative 3D Object Detection. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), New Orleans, 18-24 June 2022, 21329-21338.

[2] 《中国公路学报》编辑部. 中国汽车工程学术研究综述 2023[J]. 中国公路学报, 2023, 36(11): 1-192.

[3] Liu, Z., Wu, Z. and Toth, R. (2020) SMOKE: Single-Stage Monocular 3D Object Detection via Keypoint Estimation. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (*CVPRW*), Seattle, 14-19 June 2020, 4289-4298. https://doi.org/10.1109/cvprw50498.2020.00506

[4] Chen, X.Z., Kundu, K., Zhang, Z.Y., *et al*. (2016) Monocular 3D Object Detection for Autonomous Driving. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 2147-2156. https://doi.org/10.1109/cvpr.2016.236

[5] He, T. and Soatto, S. (2019) MONO3D++: Monocular 3D Vehicle Detection with Two-Scale 3D Hypotheses and Task Priors. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 8409-8416. https://doi.org/10.1609/aaai.v33i01.33018409

[6] Charles, R.Q., Su, H., Kaichun, M. and Guibas, L.J. (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, 21-26 July 2017, 77-85. https://doi.org/10.1109/cvpr.2017.16

[7] Charles, R.Q., Yi, L., Su, H., *et al*. (2017) PointNet++ : Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Proceedings of the* 31*st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 5105-5114.

[8] Li, J., Luo, C. and Yang, X. (2023) PillarNeXt: Rethinking Network Designs for 3D Object Detection in Lidar Point Clouds. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Vancouver, 17-24 June 2023, 17567-17576. https://doi.org/10.1109/cvpr52729.2023.01685

[9] 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

[10] Yan, Y., Mao, Y. and Li, B. (2018) SECOND: Sparsely Embedded Convolutional Detection. *Sensors*, **18**, Article 3337. https://doi.org/10.3390/s18103337

[11] Lang, A.H., Vora, S., Caesar, H., Zhou, L., Yang, J. and Beijbom, O. (2019). PointPillars: Fast Encoders for Object Detection from Point Clouds. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Long Beach, 15-20 June 2019, 12689-12697. https://doi.org/10.1109/cvpr.2019.01298

[12] Xu, D., Anguelov, D. and Jain, A. (2018) PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 244-253. https://doi.org/10.1109/cvpr.2018.00033

[13] Ku, J., Mozifian, M., Lee, J., Harakeh, A. and Waslander, S.L. (2017) Joint 3D Proposal Generation and Object Detection

from View Aggregation. 2018 *IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS*), Madrid, 1-5 October 2018, 1-8. https://doi.org/10.1109/iros.2018.8594049

[14] Yu, H., Luo, Y., Shu, M., Huo, Y., Yang, Z., Shi, Y., *et al*. (2022) DAIR-V2X: A Large-Scale Dataset for Vehicle-Infrastructure Cooperative 3D Object Detection. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), New Orleans, 18-24 June 2022, 21329-21338. https://doi.org/10.1109/cvpr52688.2022.02067

[15] Qi, C.R., Litany, O., He, K.M., *et al*. (2019) Deep Hough Voting for 3D Object Detection in Point Clouds. https://arxiv.org/abs/1904.09664

[16] 张勇, 石志广, 沈奇, 等. 基于特征融合的改进型 PointPillar 点云目标检测[J]. 光学精密工程, 2023, 31(19): 2910-2920.