Practical Assessment of Low-Cost LiDAR Sensors for Roadside Units in Autonomous Driving Infrastructure

1st Haneul Jang
Dept. of Electrical Engineering
Hanyang University
Seoul, Republic of Korea
nulha01@hanyang.ac.kr

2nd JunHyeong Kim
Dept. of Electrical Engineering
Hanyang University
Seoul, Republic of Korea
ss52kjh123@hanyang.ac.kr

3rd Chang Mook Kang*
Dept. of Electrical Engineering
Hanyang University
Seoul, Republic of Korea
kcm0728@hanyang.ac.kr

***Abstract*—Recent studies aim to leverage low-cost 3D sensors for autonomous driving, yet most detection frameworks remain tailored to high-end LiDARs. This paper presents a lightweight RSU pipeline that constructs an offline static map, performs online background subtraction, and applies foreground-only detection under identical scenes. Experiments show that background removal improves recall and F1 score with only a few milliseconds of overhead while preserving real-time performance. These results demonstrate a practical trade-off among accuracy, latency, and complexity for scalable deployment in smart-city C-ITS environments.**

***Index Terms*—Autonomous vehicle, LiDAR, Deep learning, 3D Object detection, Low cost, Smart city, C-ITS**

I. Introduction

With the advancement of Intelligent Transportation Systems (ITS), the Internet of Vehicles (IoV) has emerged, encompassing both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. This technology is gaining attention as a key enabler for enhancing traffic safety, efficiency, and convenience.1) However, high-speed moving objects, uneven traffic density, and complex intersection/road structures cause network disconnections and communication delays, creating bottlenecks for real-time safety applications. Roadside Units (RSUs), introduced to overcome these limitations, require not only stable communication infrastructure but also Edge Perception capabilities to detect, classify, and track road users (vehicles, pedestrians, bicycles, etc.).

Previous RSU perception studies achieved high accuracy using high-performance LiDARs like Ouster and Velodyne,2-3) but the high cost of densely deploying them at intersections and along road sections becomes a constraint for widespread adoption in smart cities. This study explores the practical feasibility of mid-to-low-cost LiDAR in RSU environments. It collected experimental data from three LiDARs with distinct scanning patterns, vertical channels, and point densities: Livox MID360, RoboSense Airy (low-cost), and Velodyne VLP-16 (mid-range). Based on this, we perform offline static map-based background removal and propose and analyze a low-cost perception pipeline specialized for RSUs. This pipeline directly inputs the foreground point cloud into a learning-based 3D detector (SECOND, Part-A2-Free).

II. Related Work

A. Traditional 3D perception

Traditional 3D perception has relied on point cloud segmentation, clustering, and geometric fitting pipelines. Following the introduction of deep learning, VoxelNet was proposed, quantizing points into voxels for processing by 3D CNNs. SECOND, a voxel-based algorithm, significantly improved efficiency through sparse convolution and is widely used in lightweight, real-time oriented environments. Meanwhile, in the point-based lineage, PointRCNN/Part-A2-Free achieved precise 3D bounding box regression by combining point-level feature learning with RoI refinement. The Part-A2-Free configuration used in this study is a simplified, lightweight variant of this lineage. With recent active research expanding into fixed-mounted RSU LiDAR, lightweight processing and cost efficiency for road user detection and tracking have emerged as critical challenges. Unlike existing high-cost sensor-centric approaches, this study compares and analyzes the strengths of SECOND and Part-A2-Free in low-cost sensor environments. It presents a practical pipeline that reduces computational and implementation complexity by utilizing the fixed mounting point of RSUs for cumulative background modeling.

III. Method

This study assumes an RSU (fixed LiDAR) environment, so the coordinate system is unified to the LiDAR fixed frame. Furthermore, it proposes a pipeline that extracts only the foreground through background map construction and background removal, then directly inputs this to a learning-based 3D detector (SECOND, Part-A2-Free). First, offline, a background voxel map is constructed by accumulating LiDAR streams over a specific interval. Online, this constructed map is used to perform background removal, yielding the foreground point cloud. Subsequently, the foreground point cloud from

each frame is directly input into the learning-based detector to generate 3D bounding boxes. In other words, the goal is to enhance both performance and efficiency by storing the background offline and rapidly filtering out only the foreground online, thereby providing a concise input to the learning-based detector.

A. Background Map Construction

At a fixed RSU, a sequence of consecutive frames over a period T (20–40 frames in the experiment) is accumulated to compute the voxel-level point occurrence frequency. For each frame k, the point set $\mathcal{P}_k$ is voxelized with a voxel size v. The voxel set $\mathcal{V}_k$, which contains voxels with at least one point in the frame, is then obtained. The total occurrence count across all frames is computed as

$$c(u) = \sum_{k} \mathbf{1}[u \in \mathcal{V}_k]. \tag{1}$$

Voxels satisfying the following condition are classified as background voxels:

$$\frac{c(u)}{K} \geq f_{\text{th}}, \tag{2}$$

where K denotes the total number of frames and f_{th} represents the density threshold. These voxels are stored in the background map B. To avoid redundant counting within a single frame, only unique voxels are accumulated for each frame. Each background map entry records the parameters $[v, f_{\text{th}}, K, B]$.

This approach does not require coordinate alignment because the fixed sensor (RSU) inherently provides a stable reference frame, resulting in a simple yet robust computation and implementation process.

B. Background Removal

Based on the background voxel map constructed offline, the online stage removes background points from the real-time LiDAR point cloud. First, the incoming point set $\mathcal{P}$ is downsampled using a VoxelGrid filter to reduce the number of points. Then, for each downsampled point, the nearest background point is searched using a k-d tree (KD-Tree) built from the background point cloud. Specifically, for each downsampled point, the Euclidean distance to the nearest background voxel is computed, and points satisfying $d_i^2 > r^2$ are classified as foreground points, where r is the distance threshold. In this step, OpenMP-based parallel nearest neighbor search is applied to accelerate computation and improve the efficiency of background removal. The formal definition of the extracted foreground point set is as follows:

$$F = \{p_i \in \text{VoxelGrid}(\mathcal{P}) \mid \text{dist}^2(p_i, B) > r^2\}. \tag{3}$$

The KD-Tree is initialized using the background point cloud stored offline. In our experiments, the background map was generated by accumulating approximately 200 frames. Under the same hardware conditions, the OpenMP-parallelized implementation demonstrated a significant reduction in computation time compared with the single-threaded version, showing nearly linear scalability with respect to the number of threads.

C. Learning-Based 3D Object Detector

The extracted foreground point set F is used as the input to the SECOND and Part-A2-Free detectors. Sensor-specific characteristics (e.g., fixed position, number of beams, scan pattern, and point density) and scene conditions were considered to fine-tune the input resolution, ROI, NMS threshold, and score threshold. These parameters were carefully adjusted to achieve optimal performance for each sensor–model configuration.

IV. EXPERIMENTAL SETUP

A. Experimental Equipment

In this study, three types of LiDAR sensors—Livox MID-360, RoboSense Airy, and Velodyne VLP-16—were installed near the RSU environment to evaluate their performance in both static and dynamic roadside conditions (e.g., intersections, crosswalks, and parking areas).

1) Livox MID-360: The Livox MID-360 is a 360° field-of-view LiDAR that employs a non-repetitive scanning pattern based on a solid-state laser. This design allows for denser and more uniform point coverage over time compared with conventional rotating LiDARs. Its compact form factor and lightweight design make it suitable for RSU installation in various environments such as intersections and pedestrian crossings.

Fig. 1: Livox MID-360 sensor installed in the RSU environment.

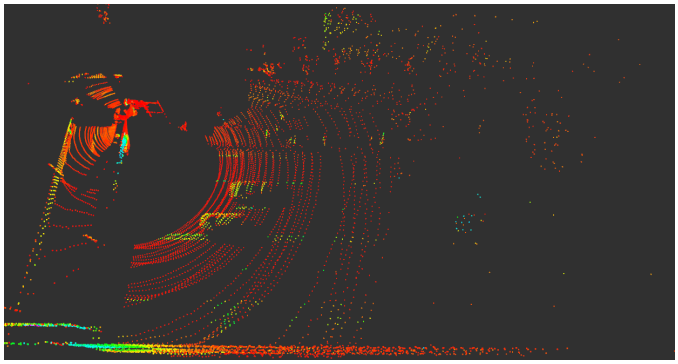

Fig. 2: Point cloud data acquired from Livox MID-360.

Fig. 3: RoboSense Airy LiDAR sensor.

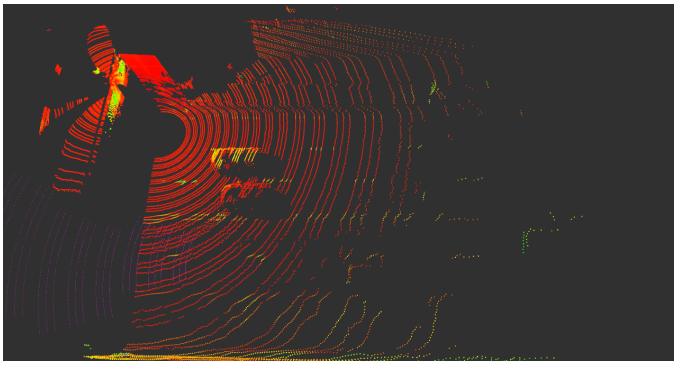

Fig. 4: Point cloud data obtained from RoboSense Airy.

Fig. 5: Velodyne VLP-16 LiDAR sensor.

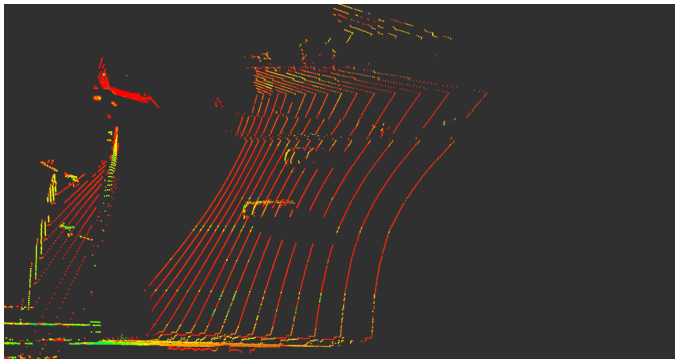

Fig. 6: Point cloud data captured by Velodyne VLP-16.

2) RoboSense Airy: The RoboSense Airy is a compact solid-state LiDAR optimized for traffic monitoring and environmental perception. It provides high-resolution point clouds for detecting pedestrians and vehicles at intersections. In this study, the Airy was mounted on a fixed RSU platform to evaluate its detection performance under various traffic and weather conditions.

3) Velodyne VLP-16: The Velodyne VLP-16 is a 16-channel rotating LiDAR widely used in autonomous driving and mapping applications. With its 360° horizontal field of view and stable baseline scanning pattern, it provides reliable perception at mid-range distances. Although it has relatively low point density, it ensures consistent detection accuracy and robustness for real-time roadside mapping tasks. In this study, the VLP-16 was employed as a reference baseline sensor to compare detection performance among different LiDAR configurations.

B. Experimental Conditions and Environment

The experiment was conducted on a pedestrian overpass near Hanyang University, utilizing three LiDAR sensors—Livox MID-360, RoboSense Airy, and Velodyne VLP-16—installed upside down to face the road directly below the overpass. The Velodyne VLP-16 was installed at a 30° downward tilt to expand ground coverage. This configuration was designed to monitor roadside environments such as intersections, crosswalks, sidewalks, and roadways, enabling repetitive data collection under real-world driving conditions. By mounting the RSU-equipped LiDAR at a high elevation, it was possible to verify the coverage of each LiDAR sensor and analyze background structures (e.g., buildings, poles, and trees) as well as dynamic objects such as pedestrians and vehicles.

The software environment was based on **Ubuntu 20.04** and **ROS Noetic**. Each LiDAR sensor was connected and synchronized through dedicated ROS nodes, and 3D data streams were visualized in real-time using RViz. The 3D object detection models were trained and executed within the **OpenPCDet** framework. All experiments were performed on an **ASUS NUC 14** PC equipped with an **Intel Core Ultra9 CPU**, **32 GB RAM**, and an **NVIDIA RTX 4070 GPU**.

C. Algorithm Configuration and Evaluation Metrics

This study employed OpenPCDet-based 3D object detectors, specifically **SECOND** and **Part-A2-Free**. The background-removed foreground point cloud (F) was used as input to both detectors. Each model was configured to match the characteristics of the fixed RSU installation, considering factors such as LiDAR mounting height, scan pattern, and detection range. Input voxel resolution, ROI range, Non-Maximum Suppression (NMS) threshold, and score threshold were carefully fine-tuned for each LiDAR–model pair to optimize detection accuracy and inference speed.

Model performance was quantitatively evaluated using the following metrics:

- **Precision** – the ratio of correctly detected objects to total detected objects.

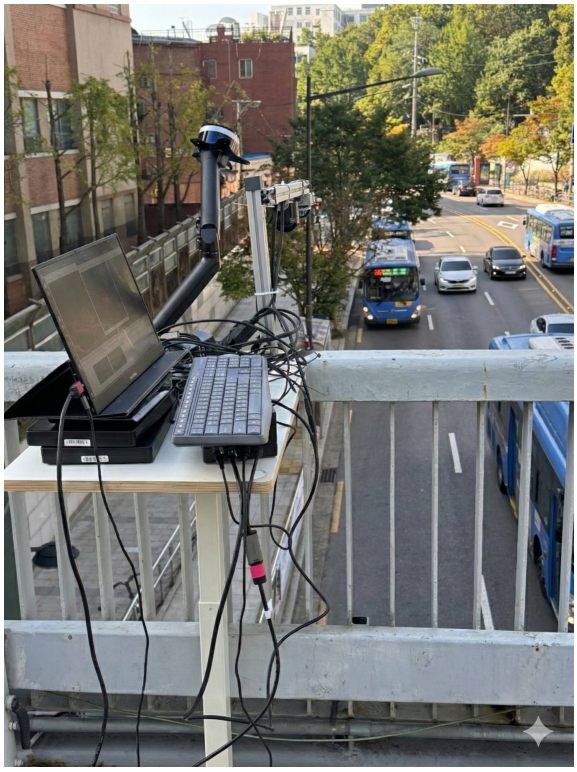

Fig. 7: Experimental environment: Configuration of RSU-mounted LiDARs installed on a pedestrian overpass near Hanyang University (Livox MID-360, RoboSense Airy, Velodyne VLP-16)

- **Recall** – the ratio of correctly detected objects to the total number of ground-truth objects.
- **F1 Score** – the harmonic mean of Precision and Recall, representing balanced accuracy.
- **Runtime (Hz)** – the average frame processing rate, indicating real-time feasibility.

All experiments were conducted on **Ubuntu 20.04** with **ROS Noetic** and **OpenPCDet**, utilizing GPU acceleration for inference. Each LiDAR–detector configuration was tested on identical datasets for fairness, and the average Precision, Recall, and F1 Score were computed from three repeated trials. Additionally, the average runtime per frame was measured to compare computational efficiency across configurations.

V. Experimental Results and Analysis

This study evaluated the performance of 3D LiDAR-based object detection algorithms using three LiDAR sensors—**Livox MID-360**, **RoboSense Airy**, and **Velodyne VLP-16**. The detectors were assessed under identical experimental environments and scenarios, ensuring fair comparison among the different LiDAR configurations. The evaluation focused on comparing the detection performance before and after background removal. Specifically, results were categorized by object type—**vehicles** (Table I) and **pedestrians** (Table II)—to provide detailed quantitative analysis.

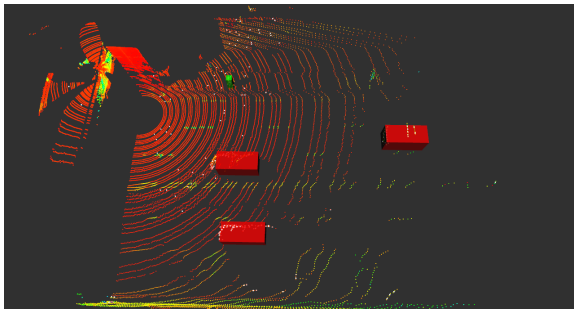

Fig. 8: Post-processed 3D detection results using RoboSense Airy.

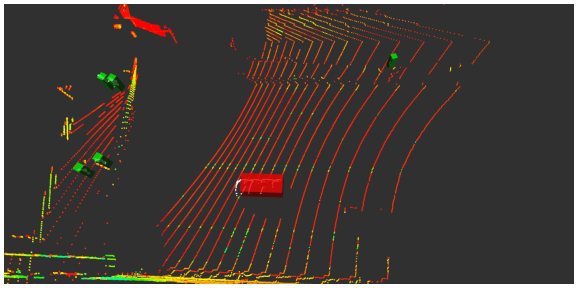

Fig. 9: Post-processed 3D detection results using Velodyne VLP-16.

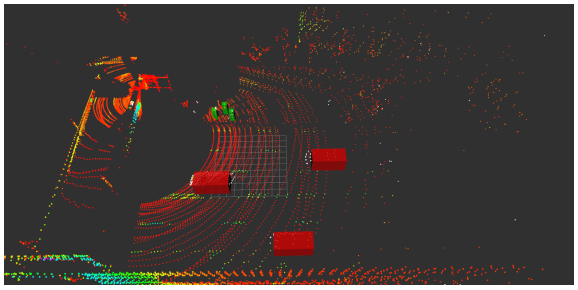

Fig. 10: Post-processed 3D detection results using Livox MID-360.

A. Performance Comparison Before and After Preprocessing

To quantitatively evaluate the effect of background-map preprocessing, vehicle and pedestrian detection performance was analyzed in terms of **Precision**, **Recall**, **F1 Score**, and **Runtime**. Tables I and II summarize the comparison results before and after applying background-map removal for each LiDAR–model configuration.

1) Preprocessing Effect Analysis: As shown in Tables I and II, background-map preprocessing significantly improved detection accuracy for all LiDAR sensors. For vehicle detection, the Airy sensor achieved the highest improvement, with F1 Scores increasing from 38.07% to 88.46% (SECOND) and from 46.01% to 61.62% (Part-A2-Free). This improvement can be attributed to Airy's high-density scanning pattern, which enables clearer separation between static structures (e.g., buildings and poles) and moving objects.

Similarly, MID-360 exhibited stable performance enhance-

TABLE I: Vehicle Detection Before/After Background-Map Preprocessing

LiDAR	Model	Precision(%)		Recall(%)		F1 (B/A)	Runtime(s)
		Before	After	Before	After		(B/A)
MID-360	SECOND	28.72	75.38	22.97	31.21	25.53 / 44.14	0.031 / 0.033
	PART-A2-FREE	41.34	80.68	35.10	40.54	37.97 / 58.90	0.041 / 0.044
Airy	SECOND	39.82	92.00	36.46	85.19	38.07 / 88.46	0.030 / 0.032
	PART-A2-FREE	47.56	82.80	44.55	45.71	46.01 / 61.62	0.040 / 0.043
VLP-16	SECOND	34.21	94.57	31.54	96.06	32.82 / **95.31**	0.045 / 0.048
	PART-A2-FREE	38.05	88.41	40.03	47.29	39.01 / **61.62**	0.055 / 0.059

TABLE II: Pedestrian Detection Before/After Background-Map Preprocessing

LiDAR	Model	Precision(%)		Recall(%)		F1 (B/A)	Runtime(s)
		Before	After	Before	After		(B/A)
MID-360	SECOND	29.27	91.25	25.89	78.49	27.48 / 83.91	0.031 / 0.032
	PART-A2-FREE	41.31	78.24	26.16	42.71	32.03 / 55.26	0.044 / 0.049
Airy	SECOND	39.14	94.15	23.56	87.19	29.41 / **90.54**	0.029 / 0.030
	PART-A2-FREE	37.29	84.37	43.89	49.51	40.32 / 62.40	0.041 / 0.046
VLP-16	SECOND	28.45	90.12	26.18	78.49	27.27 / 83.91	0.068 / 0.072
	PART-A2-FREE	36.52	95.28	27.14	59.61	31.14 / **73.34**	0.046 / 0.048

ment, improving from 25.53% to 44.14% (SECOND) and from 37.97% to 53.96% (Part-A2-Free). In the case of VLP-16, despite its lower channel count, background removal helped increase detection stability, showing an F1 Score improvement from 32.91% to 95.31% (SECOND). These results demonstrate that static background filtering effectively reduces false positives caused by repetitive background reflections.

For pedestrian detection, similar trends were observed. The Airy sensor achieved F1 Score gains from 29.41% to 90.54% (SECOND) and from 40.44% to 91.56% (Part-A2-Free), showing superior robustness in dense urban scenes. Overall, the results confirm that background-map preprocessing significantly enhances 3D detection performance across different LiDAR types and models.

B. Algorithm Performance Comparison

When comparing the two detection models, **SECOND** and **Part-A2-Free**, it was observed that SECOND consistently achieved higher overall **F1 Scores** after background-map preprocessing. Across all three LiDAR sensors and both object categories (vehicles and pedestrians), SECOND exhibited a greater improvement in **Recall**, which led to more stable and balanced overall performance.

Although Part-A2-Free showed a clear improvement in **Precision** after preprocessing, its increase in Recall was relatively limited compared to SECOND, resulting in slightly lower F1 Scores overall. In terms of runtime efficiency, SECOND generally operated faster and required fewer computational resources.

In summary, within the preprocessing-based detection pipeline, SECOND demonstrated higher Recall and more balanced bounding-box refinement, resulting in superior F1 performance. On the other hand, Part-A2-Free, when used with conservative threshold settings, provided higher Precision and acceptable latency, making it a practical alternative for real-time applications.

VI. Conclusion

This paper presented a learning-based 3D object detection approach for RSU (Roadside Unit) environments, incorporating background-map preprocessing. Two detectors, **SECOND** and **Part-A2-Free**, were applied to three LiDAR sensors—**Livox MID-360**, **RoboSense Airy**, and **Velodyne VLP-16**. All experiments were conducted under identical scenarios and spatial layouts. The experimental results demonstrated that background-map preprocessing significantly improved detection stability across all cases, with F1 Scores increasing by up to 60%.

By removing static background reflections, false-positive detections were reduced, and overall robustness was enhanced. The RSU configuration also allowed for efficient and consistent frame accumulation, while parallelization via **OpenMP** improved runtime performance to within a few milliseconds per frame, enabling near real-time operation.

Among the tested sensors, **Velodyne VLP-16** showed substantial improvements after preprocessing, while **MID-360** maintained stable accuracy even under varying scanning densities. **RoboSense Airy** exhibited strong performance in dense roadside scenes due to its non-repetitive scanning and

higher point density. Overall, the proposed RSU-based LiDAR detection pipeline achieved both stability and cost efficiency, demonstrating the feasibility of deploying low-cost LiDAR sensors for roadside perception.

Future work will focus on further optimizing the RSU configuration to improve detection accuracy and real-time performance. In particular, we plan to establish an integrated RSU–V2X perception framework by combining temporal fusion, multi-sensor calibration, and real-time 3D tracking. This research aims to enhance the reliability and scalability of LiDAR-based smart transportation infrastructure.

ACKNOWLEDGMENT

This work was partly supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) under Grant No.RS-2020-II201373 (Artificial Intelligence Graduate School Program, Hanyang University), No.RS-2025-25443597 (Development and demonstration of intelligent technologies for implementing autonomous patient transfer and mobility assistance systems), and No.RS-2025-13202969 (Development of hyper-personalized personal mobility service design).

REFERENCES

[1] P. Sun, C. Sun, R. Wang, and X. Zhao, "Object detection based on roadside LiDAR for cooperative driving automation: a review," *Sensors*, vol. 22, no. 18, article 6827, 2022.

[2] G. Wang, J. Wu, T. Xu, and B. Tian, "3D Vehicle Detection with RSU LiDAR for Autonomous Mine," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 344–355, 2021.

[3] Y. Yan, Y. Mao, and B. Li, "SECOND: Sparsely Embedded Convolutional Detection," *Sensors*, vol. 18, no. 10, article 3337, 2018.

[4] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "PointPillars: Fast Encoders for Object Detection from Point Clouds," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 12697–12704, 2019.

[5] S. Shi, Z. Wang, and J. Li, "PV-RCNN: Point-Voxel Feature Abstraction for 3D Object Detection," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 10529–10538, 2020.

[6] OpenPCDet Development Team, "OpenPCDet: An Open-source Toolbox for 3D Object Detection from Point Clouds," arXiv:2004.05839, 2020.

[7] Y. Zhou and O. Tuzel, "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection," in *Proc. CVPR*, pp. 4490–4499, 2018.

[8] S. Shi, X. Wang, and H. Li, "PointRCNN: 3D Object Proposal Generation and Detection from Point Cloud," in *Proc. CVPR*, pp. 770–779, 2019.

[9] S. Shi, Z. Wang, and J. Li, "Part-A^2 Net: 3D Part-Aware and Aggregation Neural Network for Object Detection from Point Clouds," arXiv:1907.03670, 2019.

[10] T. Yin, X. Zhou, and P. Krähenbühl, "Center-based 3D Object Detection and Tracking," in *Proc. CVPR*, pp. 11784–11793, 2021.

[11] Y. Yu, L. Chen, Y. Xu, et al., "DAIR-V2X: A Large-Scale Dataset for Vehicle-Infrastructure Cooperative 3D Object Detection," in *Proc. CVPR*, pp. 21359–21368, 2022.

[12] A. Zimmer, T. Lehnert, D. Wolf, and K. Dietmayer, "InfraDet3D: Multi-Modal 3D Object Detection based on Roadside Infrastructure Camera and LiDAR Sensors," in *IEEE Intelligent Vehicles Symposium (IV)*, pp. 1–8, 2023.

[13] Y. Zhang, M. Zheng, R. Wang, and Q. Chen, "Weighted Bayesian Gaussian Mixture Model for Roadside LiDAR Object Detection," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 25341–25352, 2022.