

Review

Object Detection Based on Roadside LiDAR for Cooperative Driving Automation: A Review

Pengpeng Sun *, Chenghao Sun, Runmin Wang and Xiangmo Zhao

School of Information Engineering, Chang'an University, Xi'an 710064, China

* Correspondence: pengpeng.sun@chd.edu.cn

Abstract: Light Detection and Ranging (LiDAR) technology has the advantages of high detection accuracy, a wide range of perception, and not being affected by light. The 3D LiDAR is placed at the commanding height of the traffic scene, the overall situation can be grasped from the perspective of top view, and the trajectory of each object in the traffic scene can be accurately perceived in real time, and then the object information can be distributed to the surrounding vehicles or other roadside LiDAR through advanced wireless communication equipment, which can significantly improve the local perception ability of an autonomous vehicle. This paper first describes the characteristics of roadside LiDAR and the challenges of object detection and then reviews in detail the current methods of object detection based on a single roadside LiDAR and multi-LiDAR cooperatives. Then, some studies for roadside LiDAR perception in adverse weather and datasets released in recent years are introduced. Finally, some current open challenges and future works for roadside LiDAR perception are discussed. To the best of our knowledge, this is the first work to systematically study roadside LiDAR perception methods and datasets. It has an important guiding role in further promoting the research of roadside LiDAR perception for practical applications.

Keywords: object detection; roadside LiDAR; cooperative perception; a review



Citation: Sun, P.; Sun, C.; Wang, R.; Zhao, X. Object Detection Based on Roadside LiDAR for Cooperative Driving Automation: A Review. *Sensors* **2022**, *22*, 9316. <https://doi.org/10.3390/s22239316>

Academic Editors: Wei Li and Honglu Jiang

Received: 4 November 2022

Accepted: 25 November 2022

Published: 30 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the emerging technologies represented by autonomous driving have developed rapidly and become the key technologies to support the development of a new generation of intelligent transportation in the future. In China's Energy Saving and New Energy Vehicle Technology Roadmap, it is expected that the market share of fully autonomous vehicles in China will reach 10% by 2030 [1]. In 2021, ERTRAC released the EU roadmap for autonomous driving, aiming to enter society by 2030 with autopilot as a standard [2]. Autonomous driving has become the inevitable trend of future automobile development. Object detection and motion prediction are the core of the perception system of autonomous vehicles, and real-time, accurate, and blind-spot-free object detection is the key to ensuring the safe driving of the autonomous vehicle. Currently, the maximum sensing distance of the onboard sensors is less than 150 m, which is easily affected by obstacles, adverse weather, and light, as well as the surface reflection intensity and motion state of the perceived object. Therefore, it is urgently necessary to develop vehicle-road collaboration technology to break through the technical bottleneck of insufficient awareness of the autonomous driving environment and improve the safety and reliability of autonomous vehicles [3]. The cooperative mode of intelligent cars and smart roads is the most effective means to ensure the safe operation of autonomous vehicles. The regional cooperative perception of intelligent roadside infrastructure makes up for the lack of local perception ability of a single vehicle and provides a new strategy to solve the over-the-horizon perception problem faced by autonomous vehicles [4].

At present, the intelligent roadside perception system mainly relies on traditional traffic information collectors, such as radar [5,6], cameras [7–9], and their fusion [10–13].

Radar can only capture moving objects in a certain area with low accuracy. The camera can obtain the color, brightness, texture, and other information of the real world, but it is easily affected by light and does not have good stability in extreme situations such as darkness and illumination mutation. Therefore, it has a natural disadvantage in the all-weather and high-precision perception of traffic information. LiDAR has the advantages of high detection accuracy, a wide sensing range, and no influence of light [14]. It has been widely used in object detection [15,16] and vehicle localization in autonomous vehicles [17,18], etc. When the 3D LiDAR is placed at the commanding point of the traffic scene, the overall situation can be viewed from the perspective of top view, and the high-precision motion trajectory of each object in the traffic scene can be sensed in real time. The information can be distributed to the surrounding vehicles or another roadside LiDAR through the advanced “5G + Cellular Vehicle-to-Everything (C-V2X)” wireless communication technology, to realize the vehicle–road cooperative regional perception. This can significantly improve the local sensing ability of autonomous driving. In addition, the roadside sensors are in a static state when working, so the perception accuracy and reliability of scene objects are higher [19,20].

Roadside LiDAR has unique characteristics compared with onboard or airborne LiDAR, such as the sparser distribution of the output point cloud and the lack of diversity of the background point cloud. The goal of deploying LiDAR sensors on the roadside for cooperative autonomous driving to achieve object detection and tracking is still in the exploratory stage, and there are still many difficulties for practical application. At present, there are some research reviews on object detection based on onboard LiDAR, as shown in [21–23], but there are few reviews on roadside LiDAR perception. Recently, Bait et al. [24] reviewed object detection and tracking based on roadside sensors. However, the detection methods and datasets related to roadside LiDAR are less involved. Therefore, this paper reviews the challenges, methods, and datasets involved in object detection using roadside LiDAR in detail, aiming to establish an overall landscape for object detection based on roadside LiDAR and accelerating the commercial application of roadside LiDAR-based perception technology. To the best of our knowledge, this is the first work to provide an in-depth survey of roadside LiDAR perception. The contribution of this paper is summarized as follows:

- ✧ The characteristics of roadside LiDAR, the challenges in object detection tasks, and methods of object detection based on roadside LiDAR in recent years are reviewed in depth, including the methods based on a single roadside LiDAR and cooperative detection of multiple LiDAR.
- ✧ The influence of adverse weather on LiDAR and methods of LiDAR perception in adverse weather are reviewed. Moreover, this paper collects and analyzes the currently published datasets related to roadside LiDAR perception.
- ✧ The existing problems, open challenges, and possible research directions of object detection based on roadside LiDAR are discussed in depth to serve as a reference and stimulate future works.

The remainder of this paper is organized as follows. Section 2 summarizes the characteristics of roadside LiDAR and the challenges of object detection based on roadside LiDAR. Section 3 reviews the object detection methods and datasets based on roadside LiDAR. Section 4 presents some open challenges and possible research directions for future works, and the last section concludes this paper.

2. Characteristics of Roadside LiDAR and Challenges in Object Detection Task

As a perception sensor with high precision, high sensitivity, and a short delay, roadside LiDAR is generally deployed at urban intersections with a complex traffic flow, ramp entrances and exits of expressways, and accident-prone areas on expressways. Through the appropriate layout and networking of multiple roadside LiDARs, we can obtain the traffic data of the road section or the whole intersection from the perspective of top view without blind areas, and then reconstruct the scene of all objects and environments in

the surrounding area, so as to accurately identify the vehicle status and vulnerable traffic groups such as pedestrians and non-motor vehicles in real time, as shown in Figure 1c. On the one hand, this can make up for the deficiency of perception of a single autonomous vehicle, and, at the same time, it can improve the level of road intelligence, effectively alleviate traffic congestion, and improve the capacity and efficiency of the road network. To meet the requirements of roadside perception, roadside LiDAR needs to work for a long time, even without interruption. In addition, due to the dynamic complexity of the traffic environment, the roadside LiDAR needs to have a larger sensing range and higher resolution. To cover a larger area, the roadside LiDAR is generally installed on the roadside infrastructure with a height of more than meters. Figure 1a shows a roadside 3D LiDAR with high-stability 32L-LiDAR-R, which was developed by China's Wanji Technology for roadside perception applications in an intelligent vehicle–road cooperation system. It has 32 scanning beams, with a vertical field of view of 37° and a horizontal field of view of 360° . Its maximum coverage is 150 m. In addition, the difference between the roadside LiDAR and the onboard LiDAR is that the vertical scanning field of the roadside LiDAR is downward ($-37\sim 0^\circ$), and the laser transmitter adopts a multi-level distribution mode with different vertical angle resolutions (0.6° , 1° , 2° , 3°), which not only increases the density of traffic object points but also improves the spatial distribution of point clouds. The input point cloud is shown in Figure 1b.

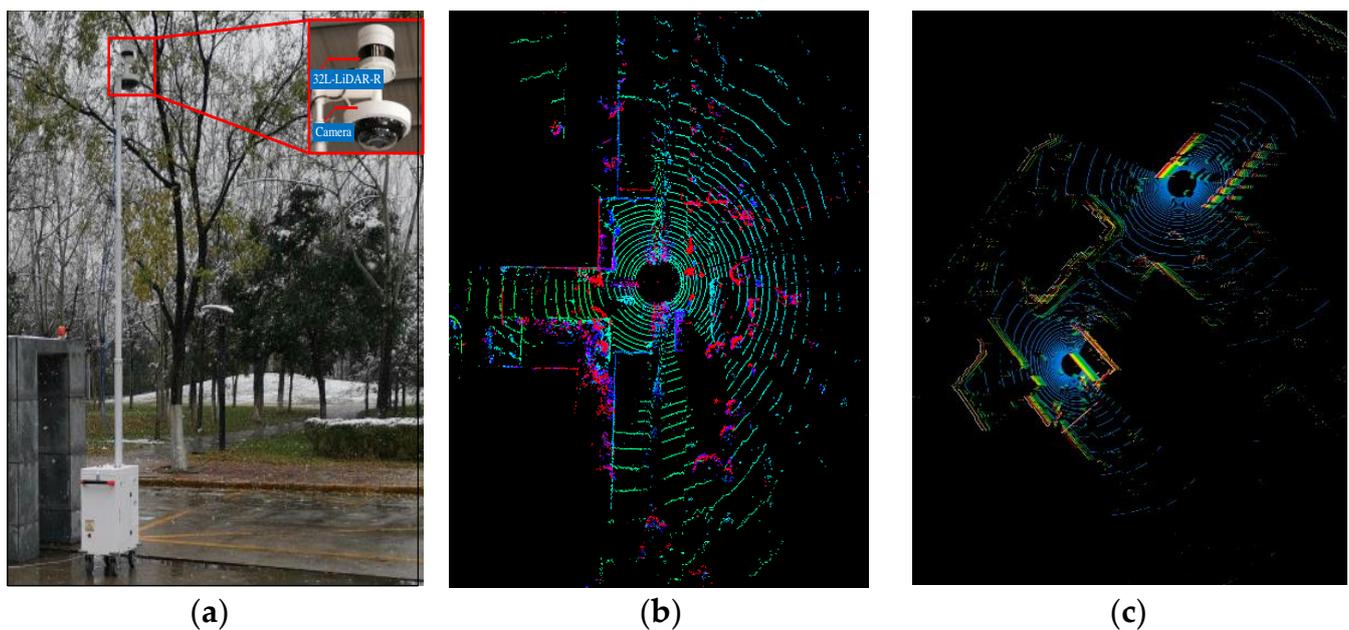


Figure 1. Roadside LiDAR and output point cloud. (a) 32L-LiDAR-R; (b) point clouds from a single LiDAR output; (c) point clouds from multiple LiDAR outputs.

At present, roadside LiDAR is mainly deployed on road sections of some pilot demonstration areas, and there are still many challenges in its deployment on a large scale for cooperative autonomous vehicles to achieve the goal of over-the-horizon perception in traffic scenes. This is manifested in the following:

- (1) The best LiDAR sensor type for roadside perception remains an open question; except for a few LiDAR models especially developed for roadside sensing (e.g., 32L-LiDAR-R), most of the roadside collaborative projects directly deploy the onboard LiDAR at the roadside. LiDAR can be divided into mechanical rotary and solid-state LiDAR according to its structure. Mechanical rotary LiDAR changes faster and more accurately from “line” to “surface” by continuously rotating the transmitting head and arranging multiple beams of laser in the vertical direction to form multiple surfaces, to achieve the purpose of dynamic scanning and dynamic information reception. Commercially

available rotary LiDARs include 16, 32, 40, 64, 80, and 128-beam LiDARs, which can perform 360-degree rotary scanning to achieve the three-dimensional reconstruction of full-scene traffic objects; the detailed parameters of the typical LiDAR are shown in Table 1. Hybrid solid-state LiDAR uses semiconductor “micro-motion” devices (such as a MEMS scanning mirror) to replace the macro-mechanical scanner and changes the emission angle of a single emitter through a micro-galvanometer, to achieve the effect of scanning without the external rotating structure. Therefore, this type of LiDAR can achieve a higher resolution and achieve image-level point cloud output. For example, the Falcon-K LiDAR has a higher resolution equivalent to 300 beams, a horizontal angle of view of 120 degrees, an angular resolution of 0.08 degrees, and an output of 3 million points per second in the double echo return mode.

Table 1. Overview of typical LiDAR sensors.

LiDAR	LiDAR Beams	FOV (Vertical)	FOV (Horizontal)	Resolution (Vertical)	Resolution (Horizontal)	Range	Range Accuracy	Points Per Second
VLP-16 [25]	16	30° (−15~+15)	360°	2°	0.1~0.4°	100 m	±3 cm	300,000
VLP-32C [26]	32	40° (−25~15)	360°	≥0.03°	0.1~0.4°	200 m	±3 cm	Single echo mode: 600,000
Pandar40M [26]	40	40° (−25~15)	360°	0.33~6°	0.2°	120 m	±2 cm	Single echo mode: 720,000
Pandar64 [27]	64	40° (−25~15°)	360°	0.167~6°	0.2°	200 m	±2/5 cm	Single echo mode: 1,152,000
HDL-64E [27]	64	26.8° (24.8~2°)	360°	0.33~6°	0.08~0.35°	120 m	±2 cm	Single echo mode: 1,300,000
RS-Ruby Lite [28]	80	40° (−25~15°)	360°	Up to 0.1°	0.1~0.4°	230 m	±3 cm	Single echo mode: 1,440,000
RS-Ruby [29]	128	40° (−25~15°)	360°	Up to 0.1°	0.2~0.4°	200 m	±3 cm	Single echo mode: 2,304,000
Falcon-K [30]	Eq. to 300	25°	120°	0.16°	0.24°	250 m	±2 cm	Double echo mode: 3,000,000

The number of LiDAR beams determines the resolution, detection range, and cost of the LiDAR. Figure 2 shows the output point cloud of a typical LiDAR. There is a large difference in the resolution and detection range of the object for LiDAR with different laser beams. The more laser beams, the higher the resolution, the greater the density of the laser point cloud, and the higher the cost. Because there is no uniform specification to specify which sensor will dominate future roadside sensing applications, the LiDAR sensing algorithms studied at present are all based on a certain beam of LiDAR, and the algorithms developed based on a certain type of LiDAR cannot be directly applied to another, different type of LiDAR. Therefore, it is still a great challenge to explore and develop a sensing method that can be applied to different types of LiDAR sensors.

- (2) In order to allow the roadside LiDAR to obtain a sufficiently large coverage area, the deployment height of the roadside LiDAR is generally above 5 m. Therefore, compared with the point cloud output by onboard LiDAR, the point cloud is more widely distributed and sparser, which increases the difficulty of object detection. In addition, the roadside LiDAR is installed at a fixed location, resulting in a high degree of similarity and lack of diversity in the background point cloud. Moreover, the lack of large-scale roadside LiDAR point cloud datasets limits the use of deep learning methods, which causes the robustness and scene generalization ability to existing detection algorithms to still face enormous challenges in practical applications.

- (3) Roadside LiDAR is fixed and deployed on the roadside, and it needs to work for a long time and in all weather. It will inevitably encounter adverse weather conditions such as rain, snow, and fog. Studies have shown that adverse weather has a greater impact on the performance of LiDAR. Rain, snow, fog, and other adverse weather will reduce the reflection intensity of the object point cloud and the number of object points, while increasing noise and reducing the resolution of the object in the point cloud, as shown in [31–34]. Therefore, the developed algorithm needs to maintain high reliability and accuracy in adverse weather.
- (4) Most of the roadside LiDAR detection methods proposed at present are based on a single LiDAR, but the field of view of a single LiDAR is limited, and the point cloud data obtained by a single LiDAR have certain defects. The accuracy of perception can be significantly improved by fusing multiple LiDAR point clouds with the diversity of the surrounding space to achieve collaborative perception.

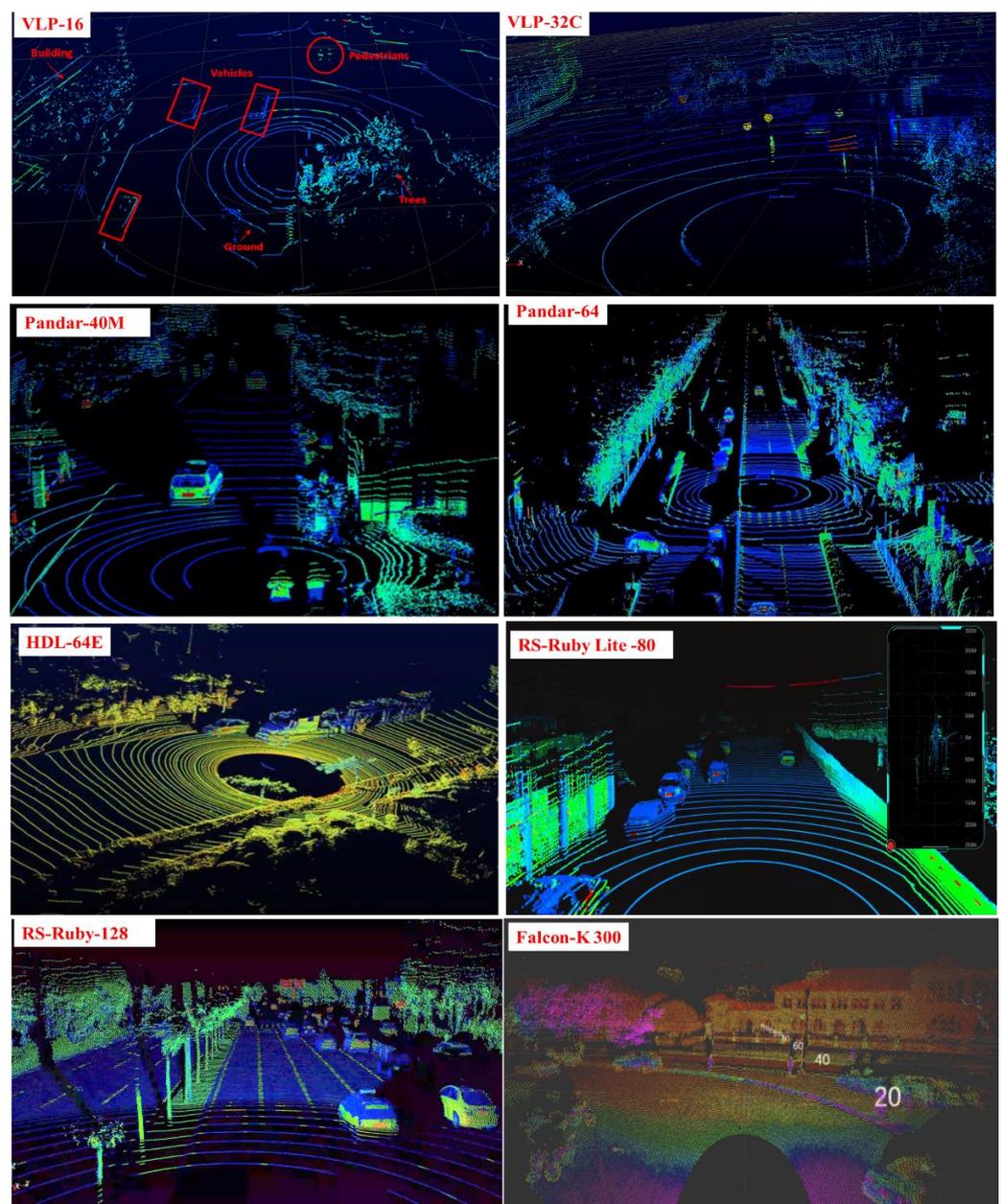


Figure 2. Point cloud of the typical LiDAR.

3. Object Detection Based on Roadside LiDAR

Three-dimensional object detection and motion prediction is a typical task of roadside LiDAR, which is used to obtain the precise location, three-dimensional size, category, and velocity of the object in the traffic scene. Object detection using roadside LiDAR can be divided into a single LiDAR-based method and multi-LiDAR cooperation-based methods. This section reviews the related challenges involved in the above tasks and the methods and datasets proposed in recent years in detail. In addition, this section also analyzes the impact of adverse weather on LiDAR and reviews the current work on LiDAR object detection in adverse weather.

3.1. Object Detection Methods Based on a Single Roadside LiDAR

Compared with the regional sensing method based on multiple LiDAR cooperatives, the perception system using a single roadside LiDAR does not require the performance of complex spatio-temporal registration between multiple sensors and the requirements for deployment scenarios are much lower. Currently, most of the proposed object detection methods are based on a single roadside LiDAR, which can be divided into methods based on traditional machine learning and deep learning, and the technical route is shown in Figure 3.

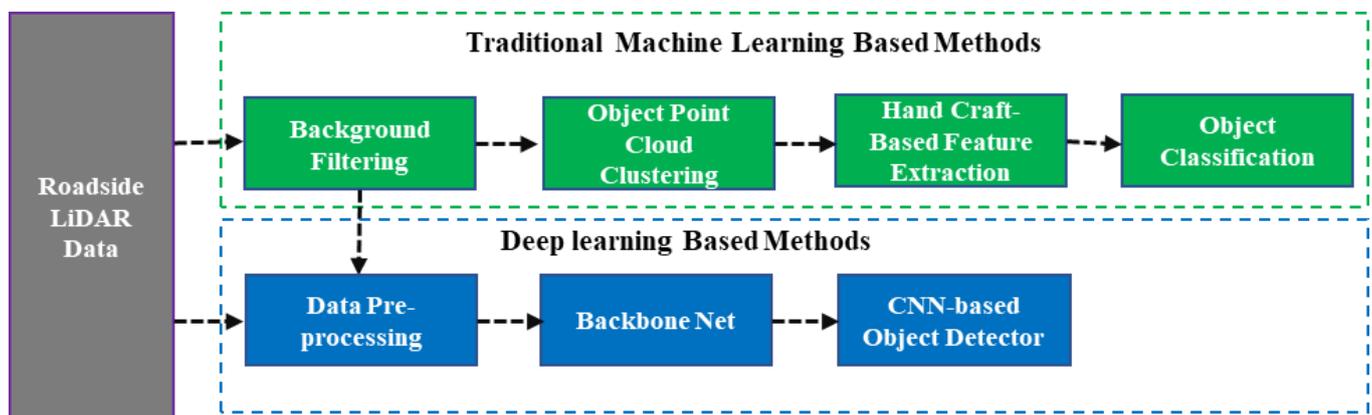


Figure 3. Technical route based on a single roadside LiDAR.

A. Detection Method Based on Traditional Machine

In recent years, the research results of roadside LiDAR object detection methods based on traditional machine learning have been the most numerous. Usually, the object detection task is divided into four main steps: background filtering, object point clustering, feature extraction, and object classification [35–38]. When processing the point clouds captured by roadside LiDAR, some reasonable clusters are generated by background filtering and density-based spatial clustering methods, and then object detection is achieved by feature extraction and the classification of each cluster.

(1) Background filtering: The purpose of background filtering is to retain the interesting object points in the point cloud as much as possible while excluding other irrelevant points (buildings, trees, ground points), to improve the efficiency and accuracy of subsequent object segmentation. The principle of background filtering is mainly to model the background by using the characteristic that the spatial position of the background point in the time sequence point cloud changes little, and then judge whether the current point is the background point or the foreground point according to the difference in the depth, height, point density, and other characteristics between the current frame point cloud and the background model. According to the expression of the point cloud, background filtering methods can be divided into point cloud mapping-based methods and voxel-based methods.

The background filtering method based on point cloud mapping aims to encode the temporal point cloud according to the azimuth angle and the ID of the laser beam, as

shown in Figure 4a. The background modeling and filtering are performed on the point cloud at the same beam and azimuth [39–44]. Zhao et al. [40] proposed an azimuth–height background filtering method. The main idea is to manually select the frame without a foreground point cloud as the background, and distinguish background points and foreground points by comparing the height of each point in the current point cloud with the background point in the background frame at the same beam and the same azimuth angle. Lee et al. [41] encoded the output point cloud according to the vertical angle and horizontal azimuth angle of the LiDAR, assumed that the background depth value at a given vertical angle and horizontal azimuth angle was constant, and took the median value of consecutive multi-frame sampling points at the same azimuth angle as the background depth value. Zhang et al. [43] used the same method to encode the multi-frame point cloud and construct the background model using the maximum distance according to the principle that the static environment is impenetrable. The above methods have high real-time performance, but, in urban scenes, it is difficult to ensure that only background points are obtained in the point cloud when the background frame is acquired. When the roadside LiDAR is deployed in crowded traffic scenes, object point clouds such as vehicles will be misjudged as background points, thus affecting the accuracy of the background model. Liu et al. [44] optimized the construction method of the background point cloud based on the maximum distance. By assuming that the vehicles and another object point only appear in the peer area, they introduced the filtering of the passing region to eliminate the object points introduced in the background point cloud. When the roadside LiDAR swings in the wind, the background model established by coding the LiDAR beams and azimuth angles may also be offset.

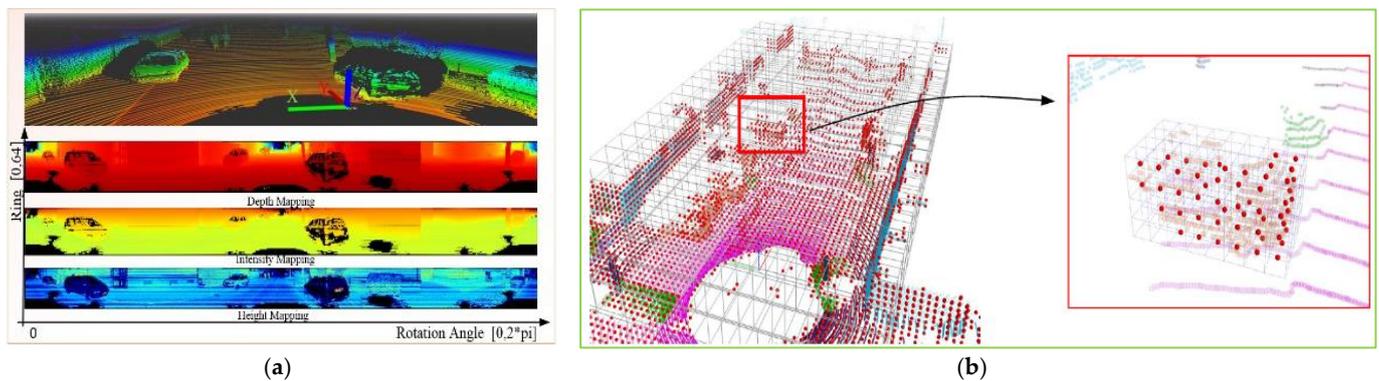


Figure 4. Representation of LiDAR point clouds. (a) Point cloud mapping. (b) Voxel-based [45].

The background modeling method based on a three-dimensional voxel aims to divide the laser point cloud into three-dimensional voxels according to the spatial coordinates of the points, as shown in Figure 4b, and establish a background model based on density statistics using multiple point cloud frames, and then discriminate the background voxels and the foreground voxels according to the point density changes in each voxel [46–49]. In addition, Wu et al. [50] proposed a variable-dimension background filtering method, which uses a dynamic matrix to store the locations of background points, and identifies background points according to the number of neighboring points and the distance between the points in the current frame and the aggregation frames. This method uses a frame-by-frame update instead of multi-frame spatial aggregation to construct and update the background, which greatly reduces the amount of calculation for background point cloud extraction. Zheng et al. [37] used the background difference method in image processing to filter the background of the roadside LiDAR point cloud data. The above methods generally use multi-frame aggregation to filter background points by summarizing the farthest distance value, average value, point density, and other features. Because the threshold parameters mainly depend on the experience of engineers, most methods lack transferability. In addition, none of the above methods can handle scenes in which the

foreground object is stationary for a period of time. Wang et al. [51] established a Gaussian background model with the average height and number of points as parameters for each three-dimensional voxel by considering the density and height distribution characteristics of the laser point cloud in three-dimensional space. These background modeling methods are based on a cubic grid. However, in the case of traffic congestion, it is difficult to choose the grid with the appropriate size to create the background model. Therefore, an adaptive polar grid Gaussian mixture model (APG-GMM) was proposed, which divides the three-dimensional grid based on the vertical angle and horizontal angle resolution of the roadside LiDAR, and a Gaussian mixture background model with the maximum distance value as the hyperparameter for the point cloud in each three-dimensional grid was established [52]. This method improves the accuracy of foreground and background segmentation. To further improve the accuracy of dynamic background modeling, Xia et al. [53] divided the points in the same scan line into multiple sectors and developed a new density background representation model (DBRM) to detect static and dynamic backgrounds (such as leaves) based on the assumption that the background objects are static in space and time. The static background is represented by density statistics, while the dynamic background is modeled by Gaussian mixture probability distribution.

Since LiDAR intensity is not as discriminative as the camera, most detection methods based on roadside LiDAR mainly use range information, while intensity values are often ignored. The intensity of laser radar is mainly affected by the surface reflectivity of obstacles, which is usually used for the detection of road signs, traffic signs, buildings, and other infrastructure, and is suitable for the background detection of roadside LiDAR. Zhang et al. [54] first encoded the 3D point cloud according to the LiDAR beam and azimuth angle where the point is located and used the hash function to store the depth value, intensity, and azimuth angle of the point in a beam–azimuth tensor with three channels. They decomposed the intensity channels based on the dynamic mode decomposition method to decompose the LiDAR data into the low-rank background and sparse foreground. Considering that the static infrastructure is the farthest object illuminated by LiDAR, the dynamic clustering method based on distance histogram analysis is used to separate the moving object from the static background. More details on the background modeling methods for roadside LiDAR are shown in Table 2. The accuracy of these methods varies from one scenario to another; thus, the accuracy is not listed in Table 2.

In summary, the above methods focus on the pre-modeling of the background point clouds, and it is difficult to achieve real-time background updates in changeable scenes.

(2) Object point cloud segmentation and recognition: After filtering the background point cloud, it is necessary to further identify vehicles, pedestrians, and other objects from the filtered foreground point cloud. Firstly, a three-dimensional object point cloud clustering algorithm, such as the point cloud clustering method based on Euler distance [55–58], point density, and its variants [36,43,44,59–62], is used to accurately segment the foreground object point cloud into independent objects. Then, according to the prior knowledge of the object, several handcrafted features, such as the standard deviation and clustering dimension of the cluster point cloud, are extracted from the cluster. Finally, some traditional classifiers, such as SVM, decision trees, and artificial neural networks, are used to realize object recognition. Zhang et al. [55,56] used the Euclidean distance method to cluster the object point cloud after filtering the roadside background point cloud based on the farthest point; extracted 28-dimensional features, such as the vertical distribution histogram, 3D size, and 2D minimum bounding box of the cluster points; and used the SVM classifier for vehicle detection. Upon finishing clustering based on Euclidean distance, Zhang et al. [54] estimated the bounding box to each cluster in which the total number of points is greater than the given threshold and used it as a candidate vehicle target for trajectory tracking. Table 3 shows a more detailed description of the proposed method.

Since the shape of point cloud objects is mostly irregular, the clustering method based on point cloud density can detect objects with arbitrary shapes, so it is also applied to the segmentation of roadside laser radar object point clouds. Among them, the Density-

Based Spatial Clustering of Applications with Noise (DBSCAN) method is the most widely used [43,44,59–62].

Table 2. Summary of background modeling for roadside LiDAR.

References	Data Description	Initial Frames	Criteria of Background/Foreground	Merit	Limitation
Zhao et al. [40]	Azimuth–height table	Manually select the frame without foreground points	Height information of the background frame	High real-time performance.	Manually select the initial frame and poor environmental adaptability.
Lee et al. [41]	Range image	Multi-frame point cloud	The median value and the average height of the sampling points in the accumulated frame.	High real-time performance, low complexity.	Ineffective background modeling for scenes with more dynamic targets in consecutive frames.
Zhang et al. [43]	Spherical range image	954-point cloud frames with few vehicle points	The maximum distance and mean distance of the sampling points in the accumulated frames.	High real-time performance, low complexity.	Not applicable to congested traffic scenarios.
Liu et al. [44]	Spherical range image	315-point cloud frames	The maximum range value of consecutive frames after filtering the object points.	Background modeling uses fewer point cloud frames and can adapt to different levels of traffic scenarios.	Decrease in background modeling accuracy when LiDAR swings due to wind.
Wu et al. [46–49]	Voxel/3D cube	2500-point cloud frames	The density of each cube is learned from accumulated frames.	High background modeling accuracy in areas with high point density at close range.	The size of the cube largely influences accuracy and computational cost.
Wu et al. [50]	Dynamic matrix	Randomly select a frame	The number of neighbors, and the distance between the points in the current frame and the aggregated frames.	Effectively filter the background points under different scenarios.	The value parameter mainly depends on experience and lacks portability.
Wang et al. [51]	Voxel/cubes	Multiple point cloud frames from different periods	Established a Gaussian background model with average height and number of points as parameters.	The robustness of the algorithm is good, and it still achieves good performance in the case of LiDAR shaking.	Different voxel sizes need to be set for different scenes, and it is difficult to select the appropriate size.
Wang et al. [52]	An adaptive grid	One frame of the LiDAR data	Established a Gaussian mixture background model with the maximum distance as the hyper-parameter.	No manual selection of background frames and high background extraction accuracy for sparse point clouds.	The algorithm has high time complexity.
Zhang et al. [54]	Elevation azimuth matrix	Successive frames	Established a background model by regressing the intensity features of continuous frames based on the DMD algorithm.	The algorithm process is applied directly on the scattered and discrete point clouds and has strong robustness.	Reflection intensity is attenuated in adverse weather, resulting in reduced algorithm performance.

Wu et al. [59,60] used the DBSCAN method to cluster and segment the roadside foreground point cloud and realized vehicle recognition based on six features: target length, height, distance from LiDAR, number of points, length–height ratio, and height profile of each cluster. Zhang et al. [43] used DBSCAN to cluster the roadside target point cloud; they adopted a different search radius according to the number of pedestrians and vehicles and filtered out the clusters belonging to noise points according to the total number of points after clustering. After DBSCAN clustering, Chen et al. [61] extracted features such as the center point of the cluster, the farthest distance between the cluster and the LiDAR center, the length and width of the minimum bounding box of the cluster, and the height difference, and classified the target based on SVM. Zhang et al. [62] also used DBSCAN to cluster the target point clouds. Five features were extracted and a probabilistic neural network (PNN) was used to achieve target classification. The performance of the DBSCAN algorithm depends on two parameters, the minimum number of points and the search radius, but the density of points decreases with the increase in the distance from the roadside LiDAR, which may affect the accuracy of density statistics, and the detection of distant vehicles and pedestrians becomes a major challenge. Therefore, Liu et al. [44] used

the DBSCAN method to identify near-range traffic objects such as vehicles and road users. For far-range traffic object detection, they extracted the trajectory of the object according to the distance and moving direction, filtered out the noise by a fast Fourier transform (FFT), and identified the points of the object. Zhao et al. [36] improved the DBSCAN method for the roadside object point cloud segmentation task, divided the detection range into sub-regions based on the sensor distance, generated an ellipsoid search space with different radii in the vertical and horizontal directions, and estimated the point density according to the maximum number of points collected by the search ellipsoid, which improved the clustering accuracy of the target point cloud. In the target classification, the 2D distance, the total number of clusters, and the direction distribution characteristics of clusters are collected, and the backpropagation artificial neural network is used to realize the classification of pedestrians and vehicles. Both the Euclidean cluster and DBSCAN work directly in the 3D world, thus suffering from the large time costs of querying every point. To improve the clustering efficiency, researchers worked on the spherical range image representation of the LiDAR point cloud. This makes it convenient to borrow image processing techniques (e.g., region growing [63], connected-component labeling [64–67]) for faster clustering. Zhang et al. [63] coded the laser points filtered out of the background into a spherical range image according to the number of LiDAR beams and horizon angle resolution of the LiDAR and then used a counted region-growing method based on the distance threshold criterion to achieve object clustering. In [64,65], the angle formed by two adjacent laser beams in the range image is used as a criterion to separate the adjacent points belonging to different clusters. Both the distance threshold and the angle threshold are empirically based, and this empirical condition does not guarantee that it always works for all pairs of points. Therefore, Zhao et al. [67] proposed a divide-and-merge point cloud clustering algorithm, which first clusters the point cloud into many local components, and then merges the locally clustered components by voting on edge point pairs. The results show that the method outperforms all published methods on the SemanticKITTI [68] panoptic leaderboard. In addition, in view of the over-segmentation phenomenon in point cloud clustering, Li et al. [69] utilized connected-component labeling to quickly group points into initial clusters based on the distance difference and merge overlapping regions or neighboring parts based on the minimum mutual point Cartesian distance, thereby reducing the over-segmentation problem in the initial clusters. Shin et al. [70] combined a two-dimensional grid and undirected graph structure to cluster non-ground points first, and then used the Gaussian process regression method to merge the over-segmented parts to improve the segmentation accuracy of the object. The experimental results show that the method achieves real-time processing speed and high segmentation accuracy in most evaluation indicators.

The above research methods all use the traditional perception pipeline, which can produce stable results for the roadside LiDAR used in the research but suffer from generality. Some challenges and problems due to the handcrafted features may arise in practical applications when the LiDAR beam is changed. In addition, the threshold selection of background filtering and clustering is also subjective, feature selection depends on experience and professional skills, and the performance of the algorithm will decline in changing scenarios. These problems and challenges may limit the applicability of most of the above research methods in practical situations.

B. Detection method based on deep learning

Compared with the method based on traditional machine learning, the method based on deep learning can learn object features autonomously from a large number of data samples to achieve end-to-end object detection.

Deep learning methods have achieved remarkable results in the field of object detection based on vehicle-mounted LiDAR point clouds and images. On the one hand, the object detection method based on vehicle LiDAR is inspired by the image object detection method. Point clouds are mapped into a bird's-eye view (such as BirdNet [71], BirdNet + [72], PIXOR [73], and YOLO3D [74]) or projected to a front view based on the horizontal and

vertical angles of the points (such as LaserNet [75], FVNet [76], RangeDet [77]) to obtain a structured data representation, which is then fed into a feedforward convolutional neural network for 3D object detection. Although the projection-based method benefits from the mature 2D detector, it inevitably loses 3D spatial information due to the spatial quantization coding. To solve this problem, object detection methods based on voxels and LiDAR point clouds have been proposed. Representative networks include VoxelNet [78], PointPillars [79], Voxel-FPN [80], IPV-RCNN [81], etc. Since the resolution of the grid is much lower than the accuracy of the three-dimensional LiDAR point cloud, the accuracy of the object detection results is slightly reduced. The method based on the point cloud aims to directly process the LiDAR point cloud without any transformation and apply the multi-layer perceptron to directly extract the features of the points from the point cloud. Along this direction, many frameworks have been proposed for object detection [82] and semantic segmentation [83].

Although the deep learning method has achieved remarkable results in the processing of vehicle LiDAR point clouds, few studies have applied deep learning-based perception algorithms to roadside LiDAR systems. Because of the similarity and lack of diversity of the background point cloud output by the roadside LiDAR, it easily leads to the overfitting of the deep learning model. When the deployment location of the roadside LiDAR changes, the performance of the algorithm will be significantly reduced, which seriously limits the application of powerful deep learning methods in the roadside environment. To solve this issue, [51] first proposed to use the background filtering method to filter the background points, so as to increase the diversity of the roadside LiDAR point cloud. Then, the non-background point cloud is used as the input of the object detection network (e.g., PointPillars [79], SECOND [84], TANNet [85]) for the training and testing of the model. Zhang et al. [57] adopted a similar approach for roadside LiDAR object detection and utilized PointVoxel-RCNN [86] (PV-RCNN) to detect vehicles and pedestrians from the extracted moving points. The experimental results show that the generalization ability of the above detection methods for target detection in different scenes has been improved. However, due to the lack of publicly available roadside LiDAR datasets for network training and testing, the detection accuracy of the algorithm is still lower than that of the vehicle LiDAR.

Recently, Zhou et al. [87] explored and studied the training of a convolutional neural network (CNN) by utilizing a large-scale autonomous driving dataset and reusing it for vehicle detection from roadside LiDAR data. The proposed CNN model was modified from the PointPillars network [79], as shown in Figure 5, by adding dense connections to achieve more comprehensive feature extraction. It is worth noting that there are great differences between roadside LiDAR and onboard LiDAR in terms of installation height, output point cloud density distribution, data occlusion, and background point cloud. To empower the model with the capability of training on onboard datasets while inferring on the roadside, Bai et al. [88] propose a Roadside Point-Cloud Encoder and Decoder (RPEaD) to transform roadside point clouds into a space in which the model trained on the onboard datasets can work. In order to reduce the shifting in Z-axis coordinates between the onboard LiDAR point cloud and the roadside LiDAR point cloud, the feature extraction network also adopts a structure similar to [79,87]; that is, the point cloud feature extraction network is based on pillars. Moreover, Zimmer et al. [89] designed a model, DASE-ProPillars, for 3D object detection using roadside LiDAR. They created a semi-synthetic infrastructure dataset with 6000 frames of the point cloud using the CARLA simulator, to make up for the deficiency of existing publicly available roadside LiDAR datasets. The comparison of the above methods is shown in Table 4.

Table 3. Summary of the features used for LiDAR data classification.

References	Object Clustering	Selected Features	Classifier	Applicability
Zhang et al. [55,56]	Euclidean cluster	<ul style="list-style-type: none"> ✚ 20D vertical point distribution histogram of the cluster. ✚ The standard deviation of points in X, Y, Z. ✚ Volume size of the cluster: length, width, maximum height, and minimum height. ✚ Area of the 2D minimum bounding box of the cluster. 	SVM classifier with RBF kernel	Vehicle detection
Zhang et al. [54]	Distance-based	<ul style="list-style-type: none"> ✚ 3D bounding box to each cluster. ✚ The number of points. 	-	Vehicle detection
Wu et al. [59,60]	DBSCAN	<ul style="list-style-type: none"> ✚ Object length, height. ✚ Difference between height and length. ✚ 10D object height profile. ✚ Distance between object and the LiDAR. ✚ The number of points. 	Naïve Bayes, K-nearest neighbor classification, decision tree, and SVM	Vehicle classification
Zhang et al. [43]	DBSCAN	<ul style="list-style-type: none"> ✚ Compare the distances between objects and LiDAR and the distances between ground and LiDAR. 	-	Vehicle and pedestrian detection
Chen et al. [61]	DBSCAN	<ul style="list-style-type: none"> ✚ Average X, Y, Z of the points. ✚ The nearest distance between the points and the LiDAR. ✚ Object length and width. ✚ Difference between the height of the object and length. 	SVM classifier	Vehicle detection
Zhang et al. [62]	DBSCAN	<ul style="list-style-type: none"> ✚ The number of points. ✚ Maximum intensity change. ✚ Distance between tracking point and LiDAR. ✚ Maximum distance in the XY plane and maximum distance in Z-axis 	Probabilistic neural network (PNN)	Pedestrian, bicycle, passenger car, and track classification
Zhao et al. [30,36]	Improved DBSCAN	<ul style="list-style-type: none"> ✚ The number of points. ✚ 2D distance. ✚ The direction of the clustered points' distribution. 	Backpropagation artificial neural network (BP-ANN)	Pedestrian and vehicle classification

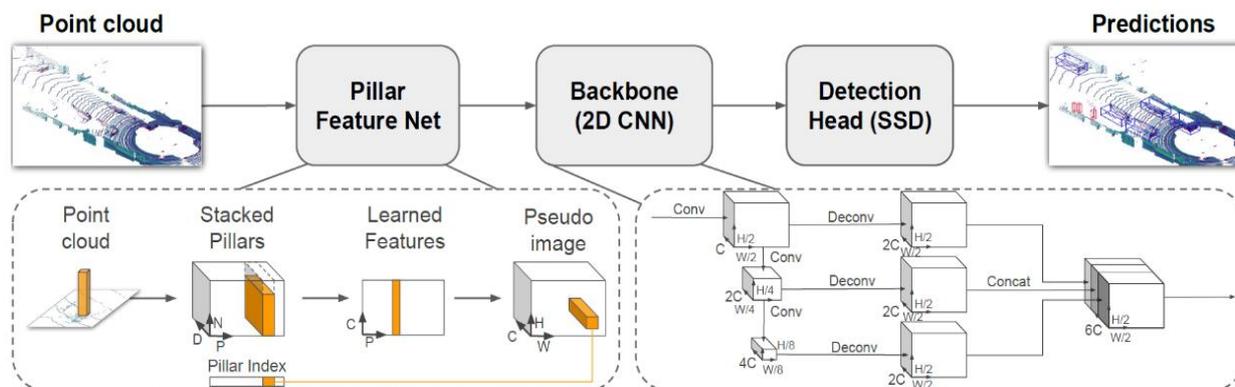


Figure 5. PointPillars 3D object detection model architecture [79].

Table 4. Comparison of object detection methods for roadside LiDAR based on deep learning.

References	Year	Architecture	Dataset	Strategy
Wang et al. [51]	2021	Background filtering module + 3D CNN detectors	Training: collected using roadside LiDAR (800 frames). Testing: collected using roadside LiDAR (100 frames).	<ul style="list-style-type: none"> Filtering out the background points before training model to improve the generalization ability and performance of 3D detectors.
Zhang et al. [57]	2022	Background filtering, clustering, tracker module + PointVoxel-RCNN detector	Training: collected using a RS-LiDAR-32 roadside LiDAR (700 frames). Testing: collected using a RS-LiDAR-32 roadside LiDAR (63 frames).	<ul style="list-style-type: none"> Candidate objects are generated by background filtering, moving point clustering, and UKF tracker before detection.
Zhou et al. [87]	2022	Modified PointPillars	Training: a large-scale autonomous driving LiDAR dataset, PandaSet, captured through a Panda 64 LiDAR (11200 frames). Testing: Roadside LiDAR data collected using a Velodyne VLP-32C. (1000 frames).	<ul style="list-style-type: none"> Reusing CNNs pre-trained on autonomous driving data to detect vehicles from roadside LiDAR data. Dense connections between convolutional layers are introduced on the basis of PointPillar to achieve more effective features.
Bai et al. [88]	2022	RPEaD + PointPillars + FPN	Training: large-scale autonomous driving LiDAR dataset, Nuscenec. Testing: 130 frames manually labelled based on the drone's view.	<ul style="list-style-type: none"> Transform roadside point clouds into coordinates suitable for training on the onboard dataset by RPEaD. To further reduce the sensitivity of the model to the difference point clouds in height, they voxelized the point cloud to generate point cloud pillars. Designed feature pyramid network to generate predicted bounding boxes.
Zimmer et al. [89]	2022	DASE-ProPillars, an improved version of the PointPillars model	Training: a semi-synthetic dataset with 6000 frames generated by a OSI-64 simulated LiDAR sensor. A9 dataset, Regensburg. Testing: A9 dataset, Regensburg Next Project.	<ul style="list-style-type: none"> Introduce five extensions to improve PointPillars. Create a semi-synthetic roadside LiDAR dataset to train the proposed model. 50 LiDAR frames collected from roadside LiDARs were manually labeled to fine-tune the detector. Perform transfer learning from the A9 dataset to the dataset from the Regensburg Next project.

3.2. Multi-LiDAR Cooperative Detection Method

Due to the limited field of view of a single LiDAR, the acquired point cloud data have certain defects. The fusion of multiple LiDAR point clouds with spatial diversity characteristics to achieve regional cooperative perception can significantly improve the accuracy of scene perception. Cooperative perception can fuse data from multiple LiDAR sensors through V2I [90–92] or I2I [19,93,94], to realize information sharing and reduce sensor deployment costs. Based on the general procedure of 3D object detection, cooperative perception can be classified into three categories depending on the sensing information to be shared between LiDAR: raw, feature, and object-level cooperative perception [95]. Raw point cloud cooperative perception seeks to fuse LiDAR point clouds with different spatial distributions into point clouds with more comprehensive object information after alignment, and then apply the fused point clouds to object detection, as shown in Figure 6. Since the most abundant information of the same object is maintained after point cloud fusion, the object detection has the highest accuracy. However, a large amount of data needs to be exchanged between different sensors. The issue of how to realize real-time and distortion-free point cloud sharing under limited communication bandwidth is a huge challenge [96].

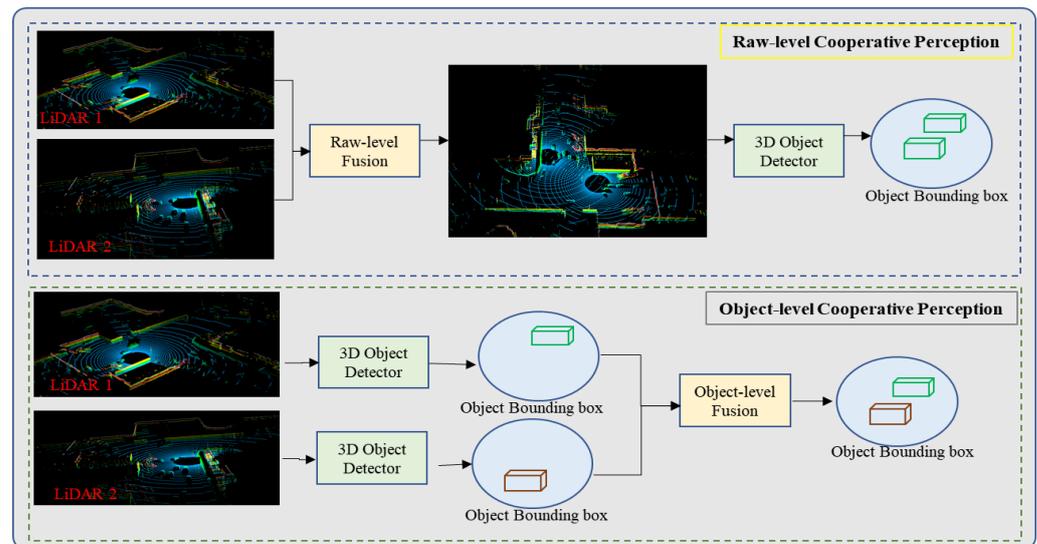


Figure 6. Logical illustration of raw-level and object-level schemes for cooperative object detection.

The object-level cooperative perception method first performs 3D object detection based on the point cloud output by each LiDAR to obtain information such as the position, bounding box, and confidence score of the object from different perspectives, and shares them to obtain more accurate object information [19,88,97], as shown in Figure 6. Since the scheme only needs to exchange the information of the object detected by LiDARs, the real-time cooperative perception among multiple LiDARs can be realized with a small communication bandwidth. However, when an object is missed in a laser radar due to occlusion of the perspective or low point density, it will be missed based on the object-level cooperative perception system. Therefore, the effect of cooperative perception based on the object level is worse than that based on point clouds and point cloud features [19].

To overcome the shortcomings of raw and object-level cooperative perception, the feature-level cooperative perception method needs to be further explored to achieve a trade-off between detection precision and bandwidth requirements. The feature-level cooperative perception first uses the feature coding layer based on voxel [98], pillar [90], or BEV [99] to generate feature maps from the point clouds output by each LiDAR, and then fuses them for object detection. Bai et al. [90] proposed a novel feature-level cooperative perception method from multiple 3D LiDARs. To avoid the extremely time-consuming 3D convolution, they first extracted the pillar features from the point cloud voxelized into pillars, and used a grid-wise feature fusion strategy to fuse multi-LiDAR data with high expandability; the architecture is shown in Figure 7. Chen et al. [98] extended their previous work [96] by fusing voxel features and using deep neural networks to learn deep features to achieve cooperative perception. Two sequence frames in the KITTI Vision Benchmark Suit (KITTI) dataset [100] were used to simulate the collaborative dataset to test the proposed method. However, KITTI only contains a limited number of traffic scenarios and is not specifically designed for cooperative perception. It is expensive to obtain a real dataset dedicated to multi-LiDAR cooperative perception, and it takes a great deal of time to manually label. Therefore, many methods proposed in recent years, such as [90,99,101], are based on synthetic data for more comprehensive cooperative perception experiments. Data generators and simulation tools such as CARLA [102] and SUMO [103] can not only generate a large amount of high-fidelity data for cooperative perception research under various traffic conditions but also provide accurate truth data. Marvasti et al. [99] used CARLA to generate simulated point cloud datasets to compare different data fusion strategies. The results show that both raw data and deep feature fusion are significantly better than object-oriented fusion, especially when vehicle positioning errors are introduced. In addition, Wang et al. [104] also confirmed that, on the simulated dataset LiDARsim [101], the shared compressed depth feature map can achieve high-precision object detection while

meeting the communication bandwidth requirements. However, due to the down-sampling of the raw point cloud data, the resolution of the shared feature map is low, so the strategy is not robust in high-precision object bounding box prediction.

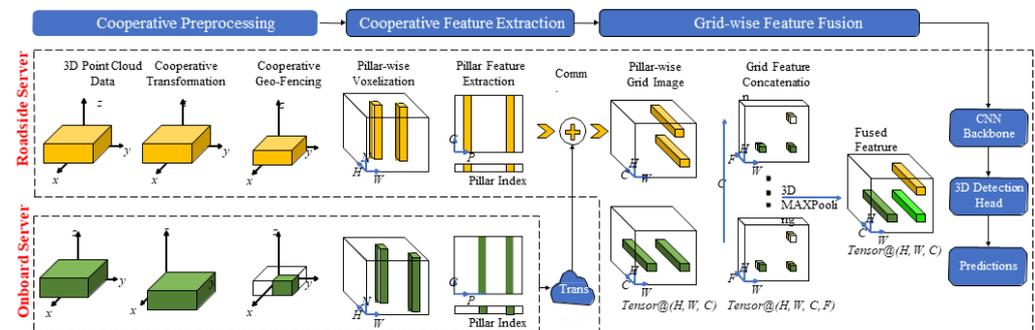


Figure 7. The system architecture of a feature-based cooperative perception framework [90].

In summary, deep feature fusion has achieved initial success, but there is still too much redundancy due to the sparsity of the shared feature map. These deep features are highly abstract and difficult to select, compress, and fuse through neural networks. For autonomous driving with real-time communication requirements, high-precision data sharing and low communication overhead are still huge challenges for cooperative perception technology.

3.3. Object Detection under Adverse Weather Conditions

At present, the roadside LiDAR perception algorithm is mainly oriented toward normal weather conditions, while roadside LiDAR deployed on the road infrastructure will withstand various adverse weather types (such as rain, snow, fog, etc.). Nevertheless, the performance of LiDAR will be degraded to some extent in adverse weather, as shown in [25,31–34,105]. In foggy or snowy weather, the backscattering of light from water droplets or snowflakes received by LiDAR will produce a number of incorrect detection points, which also reduces the number of object points in the output point cloud and increases the difficulty of object detection, as shown in Figure 8. In addition, the effective detection distance and the reflection intensity of the object surface of the laser radar in adverse weather such as rain, fog, or snowfall will also be reduced. Therefore, to improve the target recognition accuracy of laser radar in adverse weather, the laser point cloud processing algorithm must deal with these influences.

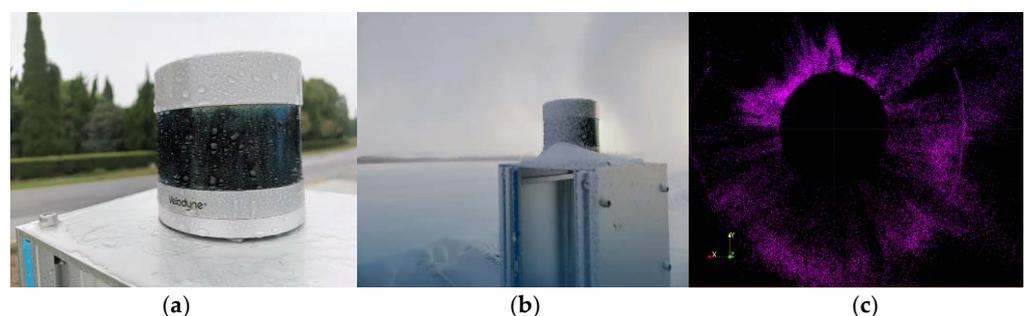


Figure 8. LiDAR operating in rain and snow and its output point cloud. (a) Under rain. (b) Under snow. (c) Point cloud under snow.

In recent years, some work related to LiDAR point cloud enhancement in adverse weather has been carried out, mainly through the filtering algorithm to preprocess the LiDAR point cloud output to improve the accuracy of object detection in adverse weather. Park et al. [106] proposed a filter based on the point cloud reflection intensity, which uses point cloud reflection intensity characteristics to filter snow particles and retain important environmental characteristics based on the analysis of laser and snow particle

characteristics. Robin et al. [107] proposed a CNN-based method to understand and filter noise in point clouds, and the results showed a significant improvement in performance compared with the latest geometry-based filtering. Roriz et al. [108] proposed a method of Dynamic Intensity Outlier Removal (DIOR), which combines the Dynamic Radius Outlier Removal (DROR) and Low-Intensity Outlier Removal (LIOR) algorithms to achieve higher accuracy while ensuring real-time performance. To improve the all-weather ability of roadside LiDAR, based on the analysis of the shortcomings of the existing filters, a combined denoising algorithm is proposed by combining the methods of crop box filter, ray ground filter, voxel filter, and statistical outlier filter [109]. Moreover, Wu et al. [110,111] studied the characteristics of roadside LiDAR data in rain and snow, proposed an improved density clustering method—3D-SDBSCAN—to distinguish vehicle points and snowflakes in LiDAR data, and used adaptive parameters to detect vehicles at different distances from roadside LiDAR sensors. The above methods are also mainly based on prior information and have poor adaptability to changeable scenes.

The rapid construction of an optimized detection network for large-scale LiDAR point cloud datasets in adverse weather provides a new idea for improving the performance of object detection in severe weather. Collecting and labeling sufficient training data in a diverse range of adverse weather conditions is laborious and prohibitively expensive. To address this issue, Yang et al. [112] modeled the performance of LiDAR under various fog conditions based on a 30-m artificial fog chamber established in Europe, and then loaded the pre-trained noise model on the LiDAR data recorded under clear weather conditions to quickly construct a large-scale LiDAR point cloud dataset under fog conditions. Kilic et al. [113] proposed a physics-based approach to simulate the LiDAR point clouds of scenarios in adverse weather conditions on existing datasets collected under normal weather conditions. Through the method, the accuracy and reliability of the 3D object detector can be improved by using the abundant existing real datasets collected in clear weather. Recently, Hahner et al. [114,115] proposed a physical method dedicated to fog and snowfall simulation, applicable to any LiDAR dataset. These partially synthetic data can be used to improve the robustness of the LiDAR point cloud perception algorithm in real fog and snowfall environments.

To summarize, there is a lack of large-scale roadside LiDAR point cloud datasets in poor weather. At present, the main method to improve the detection performance of LiDAR in adverse weather focuses on point cloud denoising. Although the roadside LiDAR point cloud collected in normal weather can be transformed to adverse weather through the physical modeling of LiDAR, there is still a lack of quantitative performance parameters of LiDAR in adverse weather such as rain and snow, which leads to a lack of confidence in the established LiDAR model. It is impossible to rely on the existing model to generate high-fidelity roadside LiDAR point clouds in adverse weather.

3.4. Datasets

High-quality roadside datasets have significant industrial value, which can accelerate the iterative optimization of roadside perception models in vehicle–infrastructure collaboration and play a positive role in promoting innovative research in academia. In recent years, many object detection and tracking datasets collected by LiDAR and cameras have been published, such as KITTI [100], H3D [116], ApolloScape [117], Waymo [118], NuScenes [119], PandaSet [120], and Panoptic Nuscenes [121]. However, these are based on data from onboard sensors; there are relatively few public roadside datasets, especially 3D point cloud datasets, which cannot meet the current needs for the iterative optimization of roadside perception models. Since last year, some datasets collected by roadside LiDAR and cameras have been released one after another. We have summarized the released datasets, and the data samples are shown in Figure 9.



Figure 9. Dataset collected by roadside LiDAR and cameras.

BAAI-VANJEE: In 2021, Deng et al. [122] released a roadside perception dataset, called the BAAI-VANJEE dataset, to support the Connected Automated Vehicle Highway technologies. This dataset provides 2500 frames of point cloud data and 5000 frames of RGB images collected from a complex urban intersection and highway scenes in China, covering different weather conditions (sun, cloud, rain) and times (day, night). It was collected by a 32-beam roadside LiDAR sensor and two cameras placed on the roadside at approximately 4.5 m. More detailed parameters are shown in Table 5.

Table 5. Survey of datasets recorded by infrastructure sensors.

Dataset	Year	LiDAR	Cameras	Annotated LiDAR Frames	3D Boxes	2D Boxes	Classes	Traffic Scenario	Weather and Times	Sensor Height
BAAI-VANJEE [122]	2021	1 32L-LiDAR-R 32-beam LiDAR	2 RGB cameras	2500 frames	74 k	105 k	12	Urban	Sunny/cloudy/rainy, day/night	4.5 m
IPS300+ [28]	2022	1 Robosense Ruby-Lite 80-beam LiDAR	2 color cameras	14,198 frames	454 M	-	7	Urban	Day/night	5.5 m
DAIR-X2X_I [30]	2022	1 300-beam LiDAR	1 RGB camera	10,084 frames	493 k	-	10	Urban highway	Sunny/rainy/foggy, day/night	-
A9-Dataset [123]	2022	1 Ouster-OS1 64-beam LiDAR	1 RGB camera	1098 frames	14 k	-	7	Autobahn highway	daylight	7 m
LUMPI [27]	2022	1 VLP-16, 1 HDL-64, 1 Pandar64, 1 PandarQT	1 PiCam 1 ATOM 1 YiCam	145 min	-	-	6	Urban	Sunny/cloudy/hazy/	-

IPS300+: In order to promote the research on roadside multi-modal perception in cooperative vehicle infrastructure systems, Wang et al. [28] published a bimodal dataset in 2022 with a band of roadside LiDAR and cameras. The collection scene is an urban intersection covering an area of 3000 square meters, covering a radius of 300 m. Two perception units (IPU) are installed on the diagonal of the intersection at a distance of 5.5 m from the ground for data collection. Each perception unit consists of an 80-beam RoboSense Ruby-Lite LiDAR and two Sensing-SG5 color cameras. The proposed dataset includes 14,198 frames of data covering different times. The point clouds registered by the two IPUs are stored as a single PCD file in one frame for annotation, and each frame has an average of 319.84 tags, including seven categories of pedestrians, cyclists, tricycles, cars, buses, trucks, and engineering vehicles.

DAIR-V2X-I: To accelerate computer vision research and innovation for vehicle–infrastructure cooperative autonomous driving, Yu et al. [30] released the DAIR-V2X dataset in 2022. The dataset was collected from 10 km of urban roads, 10 km of expressways, and 28 intersections in Beijing’s high-level automatic driving demonstration area. Four pairs of 300-beam roadside LiDAR and high-resolution cameras were deployed at each of the 28 intersections.

DAIR-V2X-I is a subset of DAIR-V2X, which is dedicated to roadside cooperative perception, and contains 10,084 frames of jointly annotated images and roadside LiDAR point cloud data, respectively. The annotator exhaustively labels each of the 10 object classes in each image and point cloud frame, including different vehicles, pedestrians, and different cyclists.

A9 Dataset: In 2022, Christian et al. [123] presented the A9 dataset based on roadside sensors from the 3-km-long Providentia++ test field near Munich in Germany. Sensors include a camera, radar, and 64-beam Ouster LiDAR, and are mounted on the gantry bridges and masts and provide vistas of the road. The dataset offers labeled images and LiDAR point clouds of multiple road segments and from different angle recordings of dense traffic on the A9 autobahn during daylight. The release R0 consists of 1098 labeled frames and 14,459 labeled 3D objects, including nine categories of objects, such as car, trailer, truck, van, pedestrian, bus, motorcycle, bicycle, and others.

LUMPI: Recently, Bush et al. [27] published a multi-view dataset; they used three different configurations of cameras and LiDAR to collect data, and all LiDARs were synchronized by GPS, covering the intersection area through different combinations of LiDAR and installation methods. They collected 145 min of data over three different days with varying weather conditions and labeled six categories of objects, including person, car, bicycle, motorcycle, bus, and truck. The detailed information is shown in Table 5.

4. Discussion and Future Works

Although the research on traffic object detection based on roadside LiDAR is still in the exploratory stage, it plays an increasingly important role in the realization of non-blind areas and the over-the-horizon perception ability of cooperative automatic driving vehicles. At present, many solutions have been put forward for this problem, but there are still many difficulties for commercial applications. This section will discuss in detail some very important but rarely studied open problems in object detection based on roadside LiDAR, as well as future work directions.

- (1) **Towards Scene-Adaptive High-Precision Perception:** Due to the homogeneity and lack of diversity of the background point clouds output by the roadside LiDAR, the existing detection methods mainly use traditional background filtering, clustering segmentation, feature extraction, and classification methods. This cascade method is easily affected by any changes and errors in the upstream model. The reliability and accuracy of the algorithm are greatly affected by the LiDAR deployment environment. These problems will limit the application of traditional perception methods in real scenes. Therefore, it is necessary to use the powerful learning ability of deep learning to develop an integrated object detection network with scene adaptability to improve the accuracy and reliability of object detection based on roadside LiDAR. Therefore, on the one hand, we can try to collect and annotate a large number of roadside LiDAR data from the simulation environment or the real scene to promote the research of roadside LiDAR perception methods based on data-driven considerations. On the other hand, based on the existing small number of labeled roadside point cloud datasets, we can attempt to carry out research on the deep learning method of roadside LiDAR based on few-shot learning [124].
- (2) **Adaptation in Different Roadside LiDARs:** According to the research review in this paper, the roadside LiDAR used in the existing methods or published datasets has great differences in the scanning mode and the number of scanning lines; specifically, the number of scanning lines of LiDAR ranges from 16 to 300. The observation angle of the object, sparsity of the output point cloud, and regional occlusion are different for the roadside LiDAR with a different number of scanning lines and different installation height, which makes it difficult for the existing perception methods to be applied to different roadside LiDARs. Thus far, there is no uniform standard specification for which type of LiDAR will dominate the roadside perception application in the future. Therefore, it is very important to promote research on the common perception

algorithm of different roadside LiDARs, which can mine the characteristics of LiDAR point cloud data and try to build a domain-invariant data representation, so that the detection model trained based on existing LiDAR data can be reused for new LiDARs [125].

- (3) **Adaptation in Different Weather Scenarios:** The existing perception algorithms based on roadside LiDAR are usually developed for scenes under normal weather, and do not perform well in adverse weather conditions such as rain and snow. However, the construction of roadside LiDAR point cloud datasets under adverse weather is time-consuming and costly. Therefore, we can try to establish a physical model of LiDAR with high confidence under adverse weather based on the study of the impact of adverse weather on LiDAR parameters, and then use the existing roadside LiDAR point cloud collected under normal weather to quickly construct a large-scale LiDAR point cloud dataset under adverse weather. This will promote the research of roadside LiDAR perception algorithms for weather domain adaptation.
- (4) **Towards Multiple Roadside LiDAR Cooperation:** Through a review of the roadside LiDAR perception approach in recent years, most of the studies above are based on a single LiDAR sensor, while few of them use multiple roadside LiDARs to enhance perception. However, the field of view of a single laser radar is limited, and the point cloud data obtained have some defects. The accuracy of scene perception can be significantly improved by integrating multiple laser radar point clouds from different perspectives in the surrounding space to achieve cooperative perception. At present, high-precision data sharing and low communication overhead are major challenges for multi-side LiDAR cooperative perception. If we can explore the fusion strategy of object information and key point features based on the output of multiple LiDARs, and construct the object detection model of multiple roadside LiDARs based on the fusion of key point depth features, it will provide a new idea for the realization of the enhanced perception of traffic objects under low communication bandwidth.
- (5) **Towards Multi-Model Cooperation:** The perception system based on multi-modal sensor information fusion can significantly improve the perceived performance of a single modal sensor through the complementarity of different types of modal information (such as LiDAR point clouds and images) and appropriate fusion technology [126]. However, the performance of multi-modal fusion has been limited due to the spatiotemporal asynchrony between sensors, domain bias, and noise in different modal data. Future work can explore more effective spatiotemporal registration and data fusion strategies for different modal sensors, and thus achieve better perception performance.

5. Conclusions

This paper reviewed several current issues and trends of object detection based on roadside LiDAR. This review mainly includes the following parts: the characteristics of roadside LiDAR and the challenges of object detection, object detection based on a single roadside LiDAR, object detection based on multiple roadside LiDARs, the challenges of roadside LiDAR in adverse weather and related work, the roadside LiDAR dataset, and some open problems and future work directions. The major findings in the above studies are as follows.

- (1) Due to the particularity of the deployment location of roadside LiDAR, most of the current object detection methods based on roadside LiDAR mainly adopt the traditional point cloud processing method, which has low accuracy compared with the existing object detection method based on onboard LiDAR and has poor adaptability to changing scenes and LiDAR with different beams.
- (2) The roadside LiDAR is deployed on the roadside infrastructure for a long time, and the developed algorithm must consider the impact of adverse weather on the LiDAR. At present, there is a lack of roadside LiDAR point cloud datasets in adverse weather, and the work of roadside LiDAR detection in adverse weather is mainly focused on point cloud denoising.

- (3) Most of the algorithms are based on the data of a single roadside LiDAR, and the cooperativity of the LiDAR can better exploit the advantages of the roadside LiDAR, so it is necessary to continue to promote research on the cooperative perception of multiple roadside LiDAR.
- (4) Since last year, some roadside LiDAR and image datasets have been released, but the coverage of the scene is limited, the data collected in adverse weather are sparse, and the number of LiDAR lines used in each dataset is also very different.

Roadside LiDAR will play an important role in the future cooperative perception system. There are still many challenges in the practical application of the existing perception methods based on roadside LiDAR. It is hoped that the review in this paper will play a guiding role in further promoting the research of roadside LiDAR perception methods.

Author Contributions: Conceptualization, P.S. and X.Z.; writing—original draft preparation, P.S.; writing—review and editing, P.S., C.S. and R.W.; supervision, P.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the 111 Project of China, grant number B14043; in part by the Key Project of the National Natural Science Foundation of China, grant number 52232015; and in part by the Key R&D Program of Shaanxi Province, grant number 2021LLRH-04-01-03.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank the Autonomous Driving Test Laboratory of Chang'an University for its support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. China SAR. *Energy-Saving and New Energy Vehicle Technology Roadmap 2.0*; China SAR: Beijing, China, 2020.
2. Rana, M.; Hossain, K. Connected and autonomous vehicles and infrastructures: A literature review. *Int. J. Pavement Res. Technol.* **2021**. [[CrossRef](#)]
3. Van Brummelen, J.; O'Brien, M.; Gruyer, D.; Najjaran, H. Autonomous vehicle perception: The technology of today and tomorrow. *Transp. Res. Part C Emerg. Technol.* **2018**, *89*, 384–406. [[CrossRef](#)]
4. Liu, S.; Yu, B.; Tang, J.; Zhu, Q. Towards fully intelligent transportation through infrastructure-vehicle cooperative autonomous driving: Challenges and opportunities. In Proceedings of the 2021 58th ACM/IEEE Design Automation Conference (DAC), San Francisco, CA, USA, 5–9 December 2021; pp. 1323–1326.
5. Sun, Q.; Chen, X.; Liang, F.; Tang, X.; He, S.; Lu, H. Target Recognition of Millimeter-wave Radar based on YOLOX. *J. Phys. Conf. Ser.* **2022**, *2289*, 012012. [[CrossRef](#)]
6. Sheeny, M.; Wallace, A.; Wang, S. 300 GHz radar object recognition based on deep neural networks and transfer learning. *IET Radar Sonar Navig.* **2020**, *14*, 1483–1493. [[CrossRef](#)]
7. Chetouane, A.; Mabrouk, S.; Jemili, I.; Mosbah, M. Vision-based vehicle detection for road traffic congestion classification. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e5983. [[CrossRef](#)]
8. Zou, Z.; Zhang, R.; Shen, S.; Pandey, G.; Chakravarty, P.; Parchami, A.; Liu, H. Real-time full-stack traffic scene perception for autonomous driving with roadside cameras. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–25 May 2022; pp. 890–896.
9. Cho, G.; Shinyama, Y.; Nakazato, J.; Maruta, K.; Sakaguchi, K. Object recognition network using continuous roadside cameras. In Proceedings of the 2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring), Helsinki, Finland, 19–22 June 2022.
10. Roy, A.; Gale, N.; Hong, L. Automated traffic surveillance using fusion of Doppler radar and video information. *Math. Comput. Model.* **2011**, *54*, 531–543. [[CrossRef](#)]
11. Bai, J.; Li, S.; Huang, L.; Chen, H. Robust detection and tracking method for moving object based on radar and camera data fusion. *IEEE Sens. J.* **2021**, *21*, 10761–10774. [[CrossRef](#)]
12. Wang, L.; Zhang, Z.; Di, X.; Tian, J. A roadside camera-radar sensing fusion system for intelligent transportation. In Proceedings of the 2020 17th European Radar Conference (EuRAD), Utrecht, The Netherlands, 10–15 January 2021; pp. 282–285.
13. Liu, P.; Yu, G.; Wang, Z.; Zhou, B.; Chen, P. Object Classification Based on Enhanced Evidence Theory: Radar–Vision Fusion Approach for Roadside Application. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–12. [[CrossRef](#)]

14. Barad, J. *Roadside Lidar Helping to Build Smart and Safe Transportation Infrastructure*; SAE Technical Paper; Velodyne Lidar: San Jose, CA, USA, 2021.
15. Wang, B.; Lan, J.; Gao, J. LiDAR Filtering in 3D Object Detection Based on Improved RANSAC. *Remote Sens.* **2022**, *14*, 2110. [[CrossRef](#)]
16. Zhao, X.; Sun, P.; Xu, Z.; Min, H.; Yu, H. Fusion of 3D LIDAR and camera data for object detection in autonomous vehicle applications. *IEEE Sens. J.* **2020**, *20*, 4901–4913. [[CrossRef](#)]
17. Lin, X.; Wang, F.; Yang, B.; Zhang, W. Autonomous vehicle localization with prior visual point cloud map constraints in GNSS-challenged environments. *Remote Sens.* **2021**, *13*, 506. [[CrossRef](#)]
18. Liu, H.; Ye, Q.; Wang, H.; Chen, L.; Yang, J. A precise and robust segmentation-based lidar localization system for automated urban driving. *Remote Sens.* **2019**, *11*, 1348. [[CrossRef](#)]
19. Arnold, E.; Dianati, M.; de Temple, R.; Fallah, S. Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 1852–1864. [[CrossRef](#)]
20. Cui, G.; Zhang, W.; Xiao, Y.; Yao, L.; Fang, Z. Cooperative perception technology of autonomous driving in the internet of vehicles environment: A review. *Sensors* **2022**, *22*, 5535. [[CrossRef](#)]
21. Li, Y.; Ma, L.; Zhong, Z.; Liu, F.; Chapman, M.A.; Cao, D.; Li, J. Deep learning for lidar point clouds in autonomous driving: A review. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 3412–3432. [[CrossRef](#)]
22. Roriz, R.; Cabral, J.; Gomes, T. Automotive LiDAR technology: A survey. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 6282–6297. [[CrossRef](#)]
23. Wu, D.; Liang, Z.; Chen, G. Deep learning for LiDAR-only and LiDAR-fusion 3D perception: A survey. *Intell. Robot.* **2022**, *2*, 105–129. [[CrossRef](#)]
24. Bai, Z.; Wu, G.; Qi, X.; Liu, Y.; Oguchi, K.; Barth, M.J. Infrastructure-based object detection and tracking for cooperative driving automation: A survey. In Proceedings of the 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 4–9 June 2022; pp. 1366–1373.
25. Bula, J.; Derron, M.H.; Mariethoz, G. Dense point cloud acquisition with a low-cost Velodyne VLP-16. *Geosci. Instrum. Methods Data Syst.* **2020**, *9*, 385–396. [[CrossRef](#)]
26. Carballo, A.; Lambert, J.; Monrroy, A.; Wong, D.; Narksri, P.; Kitsukawa, Y.; Takeda, K. LIBRE: The multiple 3d lidar dataset. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 1094–1101.
27. Busch, S.; Koetsier, C.; Axmann, J.; Brenner, C. LUMPI: The Leibniz University Mul-ti-Perspective Intersection Dataset. In Proceedings of the 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 4–9 June 2022; pp. 1127–1134.
28. Wang, H.; Zhang, X.; Li, Z.; Li, J.; Wang, K.; Lei, Z.; Haibing, R. IPS300+: A Challenging multi-modal data sets, including point clouds and images for Intersection Perception System. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–25 May 2022; pp. 2539–2545.
29. Robosense Technology Co., Ltd. RS-Ruby 128-Channel Mechanical LiDAR. Available online: <https://www.robosense.cn/rslidar/RS-Ruby> (accessed on 10 November 2022).
30. Yu, H.; Luo, Y.; Shu, M.; Huo, Y.; Yang, Z.; Shi, Y.; Nie, Z. DAIR-V2X: A Large-Scale Dataset for Vehicle-Infrastructure Cooperative 3D Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 21361–21370.
31. Rasshofer, R.H.; Spies, M.; Spies, H. Influences of weather phenomena on automotive laser radar systems. *Adv. Radio Sci.* **2011**, *9*, 49–60. [[CrossRef](#)]
32. Filgueira, A.; González-Jorge, H.; Lagüela, S.; Díaz-Vilariño, L.; Arias, P. Quantifying the influence of rain in LiDAR performance. *Measurement* **2017**, *95*, 143–148. [[CrossRef](#)]
33. Li, Y.; Duthon, P.; Colomb, M.; Ibanez-Guzman, J. What happens for a ToF LiDAR in fog? *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 6670–6681. [[CrossRef](#)]
34. Michaud, S.; Lalonde, J.F.; Giguere, P. Towards characterizing the behavior of LiDARs in snowy conditions. In Proceedings of the International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015.
35. Sun, Y.; Xu, H.; Wu, J.; Zheng, J.; Dietrich, K.M. 3-D data processing to extract vehicle trajectories from roadside LiDAR data. *Transp. Res. Rec.* **2018**, *2672*, 14–22. [[CrossRef](#)]
36. Zhao, J.; Xu, H.; Liu, H.; Wu, J.; Zheng, Y.; Wu, D. Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors. *Transp. Res. Part C Emerg. Technol.* **2019**, *100*, 68–87. [[CrossRef](#)]
37. Zheng, J.; Yang, S.; Wang, X.; Xiao, Y.; Li, T. Background Noise Filtering and Clustering With 3D LiDAR Deployed in Roadside of Urban Environments. *IEEE Sens. J.* **2021**, *21*, 20629–20639. [[CrossRef](#)]
38. Sahin, O.; Nezafat, R.V.; Cetin, M. Methods for classification of truck trailers using side-fire light detection and ranging (LiDAR) Data. *J. Intell. Transp. Syst.* **2021**, *26*, 1–13. [[CrossRef](#)]
39. Wu, J.; Xu, H.; Sun, Y.; Zheng, J.; Yue, R. Automatic Background Filtering Method for Roadside LiDAR Data. *Transp. Res. Rec.* **2018**, *2672*, 106–114. [[CrossRef](#)]
40. Zhao, J.; Xu, H.; Xia, X.; Liu, H. Azimuth-Height background filtering method for roadside LiDAR data. In Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 23–30 October 2019; pp. 2421–2426.
41. Lee, H.; Coifman, B. Side-fire lidar-based vehicle classification. *Transp. Res. Rec.* **2012**, *2308*, 173–183. [[CrossRef](#)]

42. Song, Y.; Zhang, H.; Liu, Y.; Liu, J.; Zhang, H.; Song, X. Background filtering and object detection with a stationary LiDAR using a layer-based method. *IEEE Access* **2020**, *8*, 184426–184436. [[CrossRef](#)]
43. Zhang, Z.; Zheng, J.; Xu, H.; Wang, X.; Fan, X.; Chen, R. Automatic background construction and object detection based on roadside LiDAR. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 4086–4097. [[CrossRef](#)]
44. Liu, H.; Lin, C.; Gong, B.; Wu, D. Extending the Detection Range for Low-Channel Roadside LiDAR by Static Background Construction. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [[CrossRef](#)]
45. Xu, J.; Zhang, R.; Dou, J.; Zhu, Y.; Sun, J.; Pu, S. Rpvnet: A deep and efficient range-point-voxel fusion network for lidar point cloud segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 16024–16033.
46. Zhang, Z.Y.; Zheng, J.Y.; Wang, X.; Fan, X. Background filtering and vehicle detection with roadside lidar based on point association. In Proceedings of the 2018 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018; pp. 7938–7943.
47. Wu, J.; Tian, Y.; Xu, H.; Yue, R.; Wang, A.; Song, X. Automatic ground points filtering of roadside LiDAR data using a channel-based filtering algorithm. *Opt. Laser Technol.* **2019**, *115*, 374–383. [[CrossRef](#)]
48. Lv, B.; Xu, H.; Wu, J.; Tian, Y.; Yuan, C. Raster-based background filtering for roadside LiDAR data. *IEEE Access* **2019**, *7*, 76779–76788. [[CrossRef](#)]
49. Cui, Y.; Wu, J.; Xu, H.; Wang, A. Lane change identification and prediction with roadside LiDAR data. *Opt. Laser Technol.* **2020**, *123*, 105934. [[CrossRef](#)]
50. Wu, J.; Lv, C.; Pi, R.; Ma, Z.; Zhang, H.; Sun, R.; Wang, K. A Variable Dimension-Based Method for Roadside LiDAR Background Filtering. *IEEE Sens. J.* **2021**, *22*, 832–841. [[CrossRef](#)]
51. Wang, G.; Wu, J.; Xu, T.; Tian, B. 3D vehicle detection with RSU LiDAR for autonomous mine. *IEEE Trans. Veh. Technol.* **2021**, *70*, 344–355. [[CrossRef](#)]
52. Wang, L.; Lan, J. Adaptive Polar-Grid Gaussian-Mixture Model for Foreground Segmentation Using Roadside LiDAR. *Remote Sens.* **2022**, *14*, 2522. [[CrossRef](#)]
53. Xia, Y.; Sun, Z.; Tok, A.; Ritchie, S. A dense background representation method for traffic surveillance based on roadside LiDAR. *Opt. Lasers Eng.* **2022**, *152*, 106982. [[CrossRef](#)]
54. Zhang, T.; Jin, P.J. Roadside lidar vehicle detection and tracking using range and intensity background subtraction. *J. Adv. Transp.* **2022**, *2022*, 2771085. [[CrossRef](#)]
55. Zhang, J.; Xiao, W.; Coifman, B.; Mills, J. Image-based vehicle tracking from roadside LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2019**, *2019*, 1177–1183. [[CrossRef](#)]
56. Zhang, J.; Xiao, W.; Coifman, B.; Mills, J.P. Vehicle tracking and speed estimation from roadside lidar. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5597–5608. [[CrossRef](#)]
57. Zhang, J.; Xiao, W.; Mills, J.P. Optimizing Moving Object Trajectories from Roadside Lidar Data by Joint Detection and Tracking. *Remote Sens.* **2022**, *14*, 2124. [[CrossRef](#)]
58. Zhang, L.; Zheng, J.; Sun, R.; Tao, Y. Gc-net: Gridding and clustering for traffic object detection with roadside lidar. *IEEE Intell. Syst.* **2020**, *36*, 104–113. [[CrossRef](#)]
59. Wu, J. An automatic procedure for vehicle tracking with a roadside LiDAR sensor. *ITE J.* **2018**, *88*, 32–37.
60. Cui, Y.; Xu, H.; Wu, J.; Sun, Y.; Zhao, J. Automatic vehicle tracking with roadside LiDAR data for the connected-vehicles system. *IEEE Intell. Syst.* **2019**, *34*, 44–51. [[CrossRef](#)]
61. Chen, J.; Tian, S.; Xu, H.; Yue, R.; Sun, Y.; Cui, Y. Architecture of vehicle trajectories extraction with roadside LiDAR serving connected vehicles. *IEEE Access* **2019**, *7*, 100406–100415. [[CrossRef](#)]
62. Zhang, J.; Pi, R.; Ma, X.; Wu, J.; Li, H.; Yang, Z. Object classification with roadside lidar data using a probabilistic neural network. *Electronics* **2021**, *10*, 803. [[CrossRef](#)]
63. Zhang, Y.; Bhattarai, N.; Zhao, J.; Liu, H.; Xu, H. An Unsupervised Clustering Method for Processing Roadside LiDAR Data with Improved Computational Efficiency. *IEEE Sens. J.* **2022**, *22*, 10684–10691. [[CrossRef](#)]
64. Bogoslavskiy, I.; Stachniss, C. Fast range image-based segmentation of sparse 3D laser scans for online operation. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Republic of Korea, 9–14 October 2016; pp. 163–169.
65. Yuan, X.; Mao, Y.; Zhao, C. Unsupervised segmentation of urban 3d point cloud based on lidar-image. In Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), Dali, China, 6–8 December 2019; pp. 2565–2570.
66. Hasecke, F.; Hahn, L.; Kummert, A. Flic: Fast lidar image clustering. *arXiv* **2021**, arXiv:2003.00575.
67. Zhao, Y.; Zhang, X.; Huang, X. A divide-and-merge point cloud clustering algorithm for LiDAR panoptic segmentation. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–25 May 2022; pp. 7029–7035.
68. Behley, J.; Garbade, M.; Milioto, A.; Quenzel, J.; Behnke, S.; Stachniss, C.; Gall, J. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9297–9307.
69. Li, Y.; Le Bihan, C.; Pourtau, T.; Ristorcelli, T. Insclustering: Instantly clustering lidar range measures for autonomous vehicle. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 20–23 September 2020; pp. 1–6.

70. Shin, M.O.; Oh, G.M.; Kim, S.W.; Seo, S.W. Real-time and accurate segmentation of 3-D point clouds based on Gaussian process regression. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 3363–3377. [[CrossRef](#)]
71. Beltran, J.; Guindel, C.; Moreno, F.M.; Cruzado, D.; Garcia, F.; De La Escalera, A. BirdNet: A 3D Object Detection Framework from LiDAR Information. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 3517–3523. [[CrossRef](#)]
72. Barrera, A.; Beltrán, J.; Guindel, C.; Iglesias, J.A.; García, F. BirdNet+: Two-Stage 3D Object Detection in LiDAR Through a Sparsity-Invariant Bird’s Eye View. *IEEE Access* **2021**, *9*, 160299–160316. [[CrossRef](#)]
73. Yang, B.; Luo, W.; Urtasun, R. Pixor: Real-time 3d object detection from point clouds. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7652–7660.
74. Ali, W.; Abdelkarim, S.; Zidan, M.; Zahran, M.; El Sallab, A. Yolo3d: End-to-end real-time 3d oriented object bounding box detection from lidar point cloud. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.
75. Meyer, G.P.; Laddha, A.; Kee, E.; Vallespi-Gonzalez, C.; Wellington, C.K. Lasernet: An efficient probabilistic 3d object detector for autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12677–12686.
76. Zhou, J.; Tan, X.; Shao, Z.; Ma, L. FVNet: 3D front-view proposal generation for real-time object detection from point clouds. In Proceedings of the 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Suzhou, China, 19–21 October 2019; pp. 1–8.
77. Fan, L.; Xiong, X.; Wang, F.; Wang, N.; Zhang, Z. Rangedet: In defense of range view for lidar-based 3d object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 2918–2927.
78. Zhou, Y.; Tuzel, O. Voxelnet: End-to-end learning for point cloud based 3d object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4490–4499.
79. Lang, A.H.; Vora, S.; Caesar, H.; Zhou, L.; Yang, J.; Beijbom, O. Pointpillars: Fast encoders for object detection from point clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12697–12705.
80. Kuang, H.; Wang, B.; An, J.; Zhang, M.; Zhang, Z. Voxel-FPN: Multi-scale voxel feature aggregation for 3D object detection from LIDAR point clouds. *Sensors* **2020**, *20*, 704. [[CrossRef](#)]
81. Li, Y.; Yang, S.; Zheng, Y.; Lu, H. Improved point-voxel region convolutional neural network: 3D object detectors for autonomous driving. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 9311–9317. [[CrossRef](#)]
82. Wang, L.C.; Goldluecke, B. Sparse-Pointnet: See further in autonomous vehicles. *IEEE Robot. Autom. Lett.* **2021**, *6*, 7049–7056. [[CrossRef](#)]
83. Shi, S.; Wang, Z.; Wang, X.; Li, H. Part-a2 net: 3d part-aware and aggregation neural network for object detection from point cloud. *arXiv* **2019**, arXiv:1907.03670.
84. Yan, Y.; Mao, Y.; Li, B. Second: Sparsely embedded convolutional detection. *Sensors* **2018**, *18*, 3337. [[CrossRef](#)]
85. Liu, Z.; Zhao, X.; Huang, T.; Hu, R.; Zhou, Y.; Bai, X. Tanet: Robust 3d object detection from point clouds with triple attention. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 11677–11684.
86. Shi, S.; Guo, C.; Jiang, L.; Wang, Z.; Shi, J.; Wang, X.; Li, H. Pv-rcnn: Point-voxel feature set abstraction for 3d object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10529–10538.
87. Zhou, S.; Xu, H.; Zhang, G.; Ma, T.; Yang, Y. Leveraging Deep Convolutional Neural Networks Pre-Trained on Autonomous Driving Data for Vehicle Detection from Roadside LiDAR Data. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 22367–22377. [[CrossRef](#)]
88. Bai, Z.; Nayak, S.P.; Zhao, X.; Wu, G.; Barth, M.J.; Qi, X.; Oguchi, K. Cyber Mobility Mirror: Deep Learning-based Real-time 3D Object Perception and Reconstruction Using Roadside LiDAR. *arXiv* **2022**, arXiv:2202.13505. [[CrossRef](#)]
89. Zimmer, W.; Grabler, M.; Knoll, A. Real-Time and Robust 3D Object Detection Within Road-Side LiDARs Using Domain Adaptation. *arXiv* **2022**, arXiv:2204.00132.
90. Bai, Z.; Wu, G.; Barth, M.J.; Liu, Y.; Sisbot, E.A.; Oguchi, K. PillarGrid: Deep Learning-based Cooperative Perception for 3D Object Detection from Onboard-Roadside LiDAR. *arXiv* **2022**, arXiv:2203.06319.
91. Mo, Y.; Zhang, P.; Chen, Z.; Ran, B. A method of vehicle-infrastructure cooperative perception based vehicle state information fusion using improved kalman filter. *Multimed. Tools Appl.* **2022**, *81*, 4603–4620. [[CrossRef](#)]
92. Wang, J.; Guo, X.; Wang, H.; Jiang, P.; Chen, T.; Sun, Z. Pillar-Based Cooperative Perception from Point Clouds for 6G-Enabled Cooperative Autonomous Vehicles. *Wirel. Commun. Tions Mob. Comput.* **2022**, *2022*, 3646272. [[CrossRef](#)]
93. Zhang, Z.; Zheng, J.; Tao, Y.; Xiao, Y.; Yu, S.; Asiri, S.; Li, T. Traffic Sign Based Point Cloud Data Registration with Roadside LiDARs in Complex Traffic Environments. *Electronics* **2022**, *11*, 1559. [[CrossRef](#)]
94. Wu, J.; Zhang, Y.; Tian, Y.; Yue, R.; Zhang, H. Automatic vehicle tracking with LiDAR-enhanced roadside infrastructure. *J. Test. Eval.* **2020**, *49*, 121–133. [[CrossRef](#)]
95. Ghamisi, P.; Rasti, B.; Yokoya, N.; Wang, Q.; Hofle, B.; Bruzzone, L.; Benediktsson, J.A. Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 6–39. [[CrossRef](#)]

96. Chen, Q.; Tang, S.; Yang, Q.; Fu, S. Cooper: Cooperative perception for connected autonomous vehicles based on 3d point clouds. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 7–10 July 2019; pp. 514–524.
97. Hurl, B.; Cohen, R.; Czarnecki, K.; Waslander, S. Trupercept: Trust modelling for autonomous vehicle cooperative perception from synthetic data. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 341–347.
98. Chen, Q.; Ma, X.; Tang, S.; Guo, J.; Yang, Q.; Fu, S. F-cooper: Feature based cooperative perception for autonomous vehicle edge computing system using 3D point clouds. In Proceedings of the 4th ACM/IEEE Symposium on Edge Computing, Arlington, VA, USA, 7–9 November 2019; pp. 88–100.
99. Marvasti, E.E.; Raftari, A.; Marvasti, A.E.; Fallah, Y.P.; Guo, R.; Lu, H. Cooperative lidar object detection via feature sharing in deep networks. In Proceedings of the 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall), Victoria, BC, Canada, 18 November–16 December 2020; pp. 1–7.
100. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: The kitti dataset. *Int. J. Robot. Res.* **2013**, *32*, 1231–1237. [[CrossRef](#)]
101. Manivasagam, S.; Wang, S.; Wong, K.; Zeng, W.; Sazanovich, M.; Tan, S.; Urtasun, R. Lidarsim: Realistic lidar simulation by leveraging the real world. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11167–11176.
102. Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A.; Koltun, V. CARLA: An open urban driving simulator. In Proceedings of the 1st Conference on Robot Learning, Mountain View, CA, USA, 13–15 November 2017; pp. 1–16.
103. Lopez, P.A.; Behrisch, M.; Bieker-Walz, L.; Erdmann, J.; Flötteröd, Y.P.; Hilbrich, R.; Wießner, E. Microscopic traffic simulation using sumo. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2575–2582.
104. Wang, T.H.; Manivasagam, S.; Liang, M.; Yang, B.; Zeng, W.; Urtasun, R. V2vnet: Vehicle-to-vehicle communication for joint perception and prediction. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 605–621.
105. Kutila, M.; Pyykönen, P.; Holzhüter, H.; Colomb, M.; Duthon, P. Automotive LiDAR performance verification in fog and rain. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 1695–1701.
106. Park, J.I.; Park, J.; Kim, K.S. Fast and accurate desnowing algorithm for LiDAR point clouds. *IEEE Access* **2020**, *8*, 160202–160212. [[CrossRef](#)]
107. Heinzler, R.; Piewak, F.; Schindler, P.; Stork, W. Cnn-based lidar point cloud de-noising in adverse weather. *IEEE Robot. Autom. Lett.* **2020**, *5*, 2514–2521. [[CrossRef](#)]
108. Roriz, R.; Campos, A.; Pinto, S.; Gomes, T. DIOR: A Hardware-assisted Weather Denoising Solution for LiDAR Point Clouds. *IEEE Sens. J.* **2021**, *22*, 1621–1628. [[CrossRef](#)]
109. Lu, Q.; Lan, X.; Xu, J.; Song, L.; Lv, B.; Wu, J. A combined denoising algorithm for roadside LiDAR point clouds under snowy condition. In Proceedings of the International Conference on Intelligent Traffic Systems and Smart City (ITSSC 2021), Zhengzhou, China, 19–21 November 2021; pp. 351–356.
110. Wu, J.; Xu, H.; Zheng, J.; Zhao, J. Automatic vehicle detection with roadside LiDAR data under rainy and snowy conditions. *IEEE Intell. Transp. Syst. Mag.* **2020**, *13*, 197–209. [[CrossRef](#)]
111. Wu, J.; Xu, H.; Tian, Y.; Pi, R.; Yue, R. Vehicle detection under adverse weather from roadside LiDAR data. *Sensors* **2020**, *20*, 3433. [[CrossRef](#)]
112. Yang, T.; Li, Y.; Ruichek, Y.; Yan, Z. Performance Modeling a Near-Infrared ToF LiDAR Under Fog: A Data-Driven Approach. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 11227–11236. [[CrossRef](#)]
113. Kilic, V.; Hegde, D.; Sindagi, V.; Cooper, A.B.; Foster, M.A.; Patel, V.M. Lidar light scattering augmentation (LISA): Physics-based simulation of adverse weather conditions for 3D object detection. *arXiv* **2021**, arXiv:2107.07004.
114. Hahner, M.; Sakaridis, C.; Dai, D.; Van Gool, L. Fog simulation on real LiDAR point clouds for 3D object detection in adverse weather. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 15283–15292.
115. Hahner, M.; Sakaridis, C.; Bijelic, M.; Heide, F.; Yu, F.; Dai, D.; Van Gool, L. Lidar snowfall simulation for robust 3d object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 16364–16374.
116. Patil, A.; Malla, S.; Gang, H.; Chen, Y.T. The h3d dataset for full-surround 3d multi-object detection and tracking in crowded urban scenes. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 9552–9557.
117. Huang, X.; Wang, P.; Cheng, X.; Zhou, D.; Geng, Q.; Yang, R. The apolloscape open dataset for autonomous driving and its application. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 2702–2719. [[CrossRef](#)]
118. Sun, P.; Kretschmar, H.; Dotiwalla, X.; Chouard, A.; Patnaik, V.; Tsui, P.; Anguelov, D. Scalability in perception for autonomous driving: Waymo open dataset. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2446–2454.

119. Caesar, H.; Bankiti, V.; Lang, A.H.; Vora, S.; Liong, V.E.; Xu, Q.; Beijbom, O. nuscenes: A multimodal dataset for autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11621–11631.
120. Xiao, P.; Shao, Z.; Hao, S.; Zhang, Z.; Chai, X.; Jiao, J.; Yang, D. Pandaset: Advanced sensor suite dataset for autonomous driving. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; pp. 3095–3101.
121. Fong, W.K.; Mohan, R.; Hurtado, J.V.; Zhou, L.; Caesar, H.; Beijbom, O.; Valada, A. Panoptic nuscenes: A large-scale benchmark for lidar panoptic segmentation and tracking. *IEEE Robot. Autom. Lett.* **2022**, *7*, 3795–3802. [[CrossRef](#)]
122. Yongqiang, D.; Dengjiang, W.; Gang, C.; Bing, M.; Xijia, G.; Yajun, W.; Juanjuan, L. BAAI-VANJEE Roadside Dataset: Towards the Connected Automated Vehicle Highway technologies in Challenging Environments of China. *arXiv* **2021**, arXiv:2105.14370.
123. Creß, C.; Zimmer, W.; Strand, L.; Fortkord, M.; Dai, S.; Lakshminarasimhan, V.; Knoll, A. A9-dataset: Multi-sensor infrastructure-based dataset for mobility research. In Proceedings of the 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 4–9 June 2022; pp. 965–970.
124. Wang, Y.; Yao, Q.; Kwok, J.T.; Ni, L.M. Generalizing from a few examples: A survey on few-shot learning. *ACM Comput. Surv.* **2020**, *53*, 1–34. [[CrossRef](#)]
125. Triess, L.T.; Dreissig, M.; Rist, C.B.; Zöllner, J.M. A survey on deep domain adaptation for lidar perception. In Proceedings of the 2021 IEEE Intelligent Vehicles Symposium Workshops (IV Workshops), Nagoya, Japan, 11–15 July 2021; pp. 350–357.
126. Wang, S.; Pi, R.; Li, J.; Guo, X.; Lu, Y.; Li, T.; Tian, Y. Object Tracking Based on the Fusion of Roadside LiDAR and Camera Data. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–14. [[CrossRef](#)]