J. Highw. Transp. Res. Dev. 2025, 19(4): 31−36

Journal of Highway and Transportation Research and Development

Home page: www.sciopen.com/journal/2095-6215

3D vehicle detection improving method fusing image and LiDAR in foggy environment

Shaokang Niu*, Zhigang Xu, Zhichao Cui**

School of Information Engineering, Chang'an University, Xi'an, Shaanxi 710018, China

Abstract: 3D vehicle detection still remains challenging due to the occlusion, illumination and special weather. This paper proposes a data-level fusion method to combine the images and LiDAR to achieve 3D vehicle detection in the foggy environment. This method first restores the corrupted multi-modal data though image defogging and point cloud defogging. Then the pseudo-LiDAR is generated by back-projecting the image pixels with the estimated image depth. After revisitation and transformation, the pseudo-LiDAR and LiDAR data can be fused in the 3D space. Finally, the fused point clouds are fed into the BtcDet detection network to obtain detection results. Furthermore, in order to evaluate 3D detection methods, this paper establishes a multi-modal dataset (images and LiDAR) of the foggy scenarios by simulation algorithms to compensate for the lack of foggy dataset. Based on the synthetic dataset, the experiments are conducted to evaluate the proposed methods and other comparative methods. The accuracy of the 3D target detection algorithm is higher than comparative methods.

Keywords: road engineering; 3D vehicle detection; image defogging; point cloud defogging; foggy 3D datasets

1 Introduction

Accurate 3D target perception is a prerequisite for safe driving of autonomous vehicles. 3D target detection aims to detect the 3D position, attitude and type of the object from the traffic multi-modal data [1]. However, limited by special weather environments (e.g., fog) and complex traffic environments, target detection is difficult to be accurately carried out with impaired data. Therefore, the study of multi-modal data target detection in foggy environment has not only fully theoretical significance but also application prospects. In fog weather, the abilities of 3D target detection methods are extremely limited, since the data (such as image, LiDAR) are severely corrupted by the fog, which aggravate the the difficulty of 3D detection through reducing the visibility and increasing a large number of noises. Thanks to the complementary information, fusion of multi-modal traffic data is an effective fashion to improve the vehicle detection in the occlusion situation. The prevalent fusion methods consist of feature-level and data-level [2].

In order to improve the detection accuracy in the foggy and occlusion traffic environment, this paper proposes 3D vehicle detection methods combining the data-level fusion and multi-modal (Image and LiDAR) defogging methods simultaneously. Distinguished from the existing methods, our fusion methods convert image to LiDAR rather than LiDAR-to-Image [3], which loss one dimension information from 3D to 2D space. The fused dense points can compensate for sparse LiDAR to improve the detection ability. Besides, the defogging method for multi-modal data can effectively alleviate the influence of fog and improve the detection accuracy of 3D vehicles.

The major contributions of this paper can be summarized as follows:

First, this paper proposes a data-level fusion method to integrate the images and 3D LiDAR for detection. After converting the images into 3D pseudo-LiDAR, it is fused with the radar data for correction. The advantages of this approach remain the whole information of data as much as possible, enhancing 3D LiDAR point cloud.

Second, this paper puts forward a dynamic point cloud defogging method based on light intensity. Unlike DROR [4] and FCSOR [5] which are based on the distance between points, this method is based on the intensity value of the “fog point cloud” to remove noises.

Third, a novel foggy multi-modal dataset is built for 3D detection. Although foggy images now exist, there is less 3D LiDAR data in foggy environment. Foggy environment also has a very serious impact on 3D LiDAR data. Based on the working principle of LiDAR, this paper utilizes the method that simulating physically accurate fog into clear-weather scenes [6] to implement foggy 3D LiDAR data generation.

2 Experimental

2.1 Introduction to 3D target detection

3D target detection aims to obtain semantic (category) and geometric information of the target in the 3D space. 3D target detection algorithms can be broadly classified into three categories based on types of input data (i.e. image, laser point cloud and mul-

Received: 28 October 2024; **Revised:** 20 December 2024; **Accepted:** 5 January 2025

* *E-mail address*: niushaokang@chd.edu.cn; ** *Corresponding author*: cui.zhichao@chd.edu.cn

DOI: 10.26599/HTRD.2025.9480083

RESEARCH INSTITUTE OF HIGHWAY MINISTRY OF TRANSPORT | SciOpen

timodal fusion) [7]. However, images do not provide scene depth directly. The laser point cloud methods cannot solve the problem of target detection under long-range sparse point clouds. To compensate for the shortcomings of single data, more and more studies are fusing point clouds and images. More accurate detection results are obtained by complementing each other with information. 3D-CVF [8] uses an auto-calibrated projection to convert 2D camera features into a smooth spatial feature map with the highest correspondence to LiDAR features in the bird's eye view (BEV) domain. MSMDFusion [9] can better utilize depth information and fine-grained cross-modal interactions between LiDAR and cameras, and the framework consists of two important components, MDU and GMA-Conv. So these characteristic differences lead to different feature extraction methods. Convolutional neural networks [10] are currently the dominant image-based deep learning method, which can accomplish the extraction of image features and learning of higher order semantics with a level of abstraction and learning capability. This paper utilizes a data-level fusion approach to solve the problem of low accuracy of 3D target detection in foggy environment. We first defog the image and laser point cloud. Then we generate pseudo-LiDAR from images and fuse them with laser point cloud. Finally, the BtcDet [11] network is used for supervised training to achieve foggy 3D target detection.

2.2 Framework

For foggy scenes, this paper proposes a foggy 3D target detection method by fusing laser point cloud and image, and its framework is shown in Fig. 1. First, the image and point cloud need to be defogged separately. For foggy image, FFA-Net [12] is used to obtain a clearer fog-free image. For foggy LiDAR data, this paper proposes a dynamic point cloud denoising method based on the point cloud light intensity to denoise the laser point cloud. Second, we generate pseudo-LiDAR from the defogged image data, and denoise the pseudo-LiDAR before fusing it with the laser point cloud. Finally, the fused point cloud data is fed into the BtcDet [11] network for 3D target detection in foggy environment.

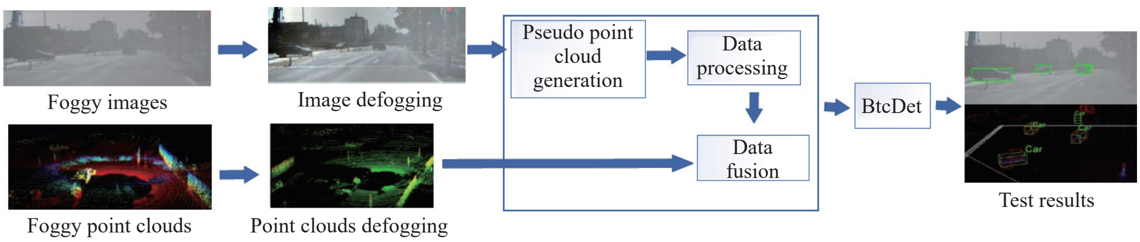

Fig. 1 3D target detection framework in foggy environment.

2.2.1 Image defogging

Due to the degradation of the foggy image and the increase in gray scale, the foggy images need to be defogged in advance for generating high-quality pseudo-point cloud. Thanks to the simple structure and good image defogging effect, the FFA-Net [12] is utilized in this paper for image defogging. The end-to-end structure of FFA-Net [12] is shown in Fig. 2.

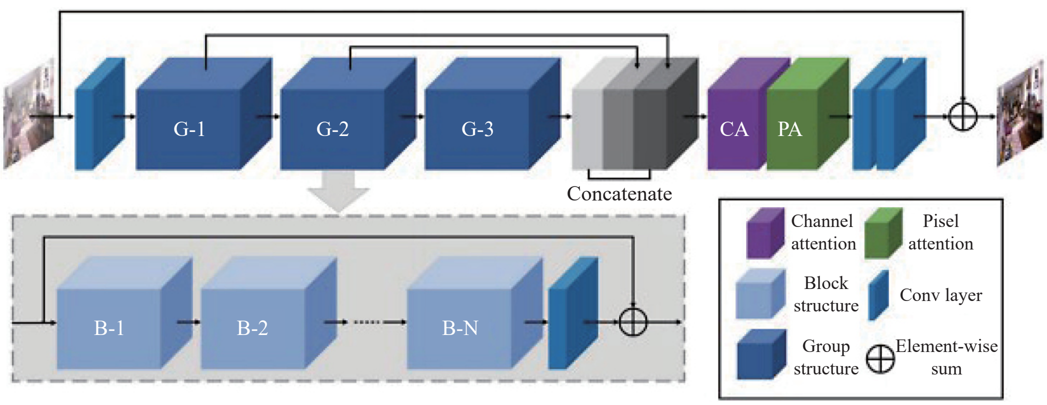

Fig. 2 FFA-Net framework for image defogging.

The network is fed with a foggy blurred image, which is passed to a shallow feature extraction section. It is then fed into N cluster structures with multi-hop connections, and the output features of the N cluster structures are fused together by the feature attention module. Finally, these features are passed to the reconstruction section and the global residual learning structure, resulting in a fog-free output. The network can retain information from the lower layers and pass it to the deeper layers. Also, the presence of the weighting mechanism makes the FFA [12] network more focused on valid information such as thick fog areas, high frequency textures and color fidelity.

2.2.2 Laser point cloud defogging

This paper proposes a light intensity-based point cloud defogging algorithm for processing point cloud noise in foggy scenes. The point clouds collected by LiDAR in foggy environment have a large number of noises which cannot be effectively removed using traditional methods such as adiusoutlier removal (ROR), statistical outlier removal (SOR) and voxel grid filters (VGF). Unlike ROR, SOR and other algorithms that remove noises based on the distance between points, our proposed method is based on the fact that the intensity value of the “fog point cloud” is lower than those of other points of other points at the same distance. Thus, the noise points are removed when the intensity values are greater than presetting threshold.

The flow of the point cloud defogging algorithm is shown in Fig. 3, which consists of two steps. Step 1 identifies and removes all points whose intensities are lower than the specified threshold. First, we need to select a suitable intensity threshold I_{thr} to distinguish the "fog point cloud" from the normal target point cloud. The

```
Input:
P(x, y, x, i), Coordinates and distances of 3D point clouds
Function PointCloudDefogging(x, y, z, i):
  1: Calculate the distance D_c between point p and the LIDAR
     origin
  2: Calculate the intensity I_c of the current distance D_c, and I_thr=I_c
  // judge
  3: If I_p>I_thr then Non-foggingpoint, depositrightpoint
  Else :
  4: Suspect point, deposit errpoint
  5: Store errpoint in KDTree
  6: Calculate the radius search distance R_l of the suspected point
     P_w at the current distance D_c
  7: Radius Proximity Search
  // judge
  8: If P_w numberofadjacentpoints K>K_thr then
     Non-foggingpoint, depositrightpoint
  Else:
  9: Fog point, delete
  return 0
```

Fig. 3 Point cloud defogging framework.

intensity profile of the laser point cloud needs to be calculated by Eq. (1) [13] to estimate the intensity value of the laser point returned at each distance. Then the I_{thr} for defogging can then be determined based on the intensity value at each distance. For each point p, if its intensity value I_{p} is higher than the I_{thr}, the point is regarded as a normal point and vice versa as an anomaly p_{w}.

$$I_{\text{c}} = I_{\text{ref}} \times \frac{D_{\text{ref}}^2}{D_{\text{c}}^2} \times \cos(\alpha) \tag{1}$$

where, a is the angle of incidence of the laser in the “fog point cloud”. In order to obtain the intensity profile corresponding to the LiDAR, the point at 5.5 m is chosen as the reference point. And its intensity value is the average of the intensity values of all points at that distance. To determine the variation of the actual intensity with distance, the incidence angle degree $w = 90°$ is assumed here.

However, some correct target points below the threshold will be wrongly deleted in the step 1. For example, if a car is located directly in front of the LiDAR sensor (within 3 m), some important points will be lost when applying the laser point cloud defogging in step 1. Because of the low value of vehicle material strength, it makes the point cloud that truly represents the car incomplete. To solve this problem, “non-fog points” carrying important environmental information are selected from the outliers identified in step 1. Then they are re-classified as inner points.

Step 2 is based on the above idea. The density of neighboring points around the “fog point cloud” is sparse, while the density of neighboring points around the “non-fog points” of the outliers is high. Therefore when the number of neighboring points within the specified search radius R_1 of an outlier k exceeds the threshold k_{thr} , the point can be saved as an inner point. First the distance between the anomaly P_{w} and the LiDAR sensor R_{P} is calculated by $R_p = \sqrt{{x_p}^2 + {y_p}^2}$. If R_{P} is lower than the minimum search radius SR_{min}, then R_1 uses the minimum search radius value SR_{min}. Otherwise, R_1 is defined using Eq. (2) in dynamic radius outlier removal [4].

$$SR_p = \beta \times \left(R_p \times \alpha\right) \tag{2}$$

where, β is a parameter (β>1) that compensates for the increased interpoint spacing due to object surfaces that are not perpendicular to the LiDAR beam. In this case, the outlier P_{w} is classified as an inner point if the number of adjacent points k within the radius R_1 is greater than the threshold k_{thr}. Otherwise, it is still an outlier and is rejected. Finally, the laser point cloud de-fogging process is completed.

2.2.3 Image and point cloud data fusion

The fusion of the image and point cloud can be mainly divided into three steps: pseudo-LiDAR generation, pseudo-LiDAR data processing, and point cloud fusion. The first step requires to use an image-based pseudo-point cloud generation method. The stereo parallax method is used to create a depth map corresponding to image pixels. Next, the depth map is back-projected into the 3D space to generate the pseudo-LiDAR. However, the directly generated pseudo-LiDAR is too dense and the number of point clouds far exceeds the number of laser point clouds in the same scene. So a voxel filter is used to downsample the pseudo-LiDAR. At the same time, the pseudo-LiDAR has some noises, so it is also necessary to use the DROR [4] to denoise the pseudo-LiDAR. Finally, the processed pseudo-LiDAR P_{f} is fused with the laser point cloud P_{t} to obtain a dense fused point cloud P.

$$P = P_{\text{t}} + P_{\text{f}} \tag{3}$$

2.2.4 BtcDET

BtcDet [11] is used to improve low target detection accuracy under the situation of missing target point shape by occlusion and signal loss. The framework of BtcDet [11] consists with three main parts. The occluded and loss regions (R_{OC} and R_{SM}) can be distinguished firstly. Then shape occupation probabilities $P(\text{O}_{\text{S}})$ is generated. Next, a 3D backbone network is used to extract features, which are fed into the RPN network to generate 3D region proposals. For each region proposal, a local grid covering the proposal box is constructed and the sparse tensor $P(\text{O}_{\text{S}})$ is stitched onto the feature map. Finally, the local geometric features are pooled onto a local grid, the grid features are aggregated, and the final bounding box prediction is generated. The optimized target detection process can be expressed as follows:

$$\arg\max_{\theta} P(O_s|\{p_1 p_2, \cdots, p_n\}, R_{\text{SM}}, R_{\text{OC}}, \theta) \tag{4}$$

$$\arg\max_{\theta} P(X, D|\{p_1 p_2, \cdots, p_n\}, P(O_s), \theta) \tag{5}$$

2.3 Foggy dataset construction

The LiDAR foggy datasets available for 3D detection task are still relatively lacking. In this paper, a physics-based fog simulation method is used to synthetize the foggy laser point cloud. It uses a standard linear system [14] to simulate the LiDAR impulse response, converting a real clear-weather LiDAR point cloud to a foggy laser point cloud. The method has been successfully applied to semantic segmentation tasks [6], proving its effectiveness in the field of foggy LiDAR data simulation and 3D target detection. KITTI [15] is selected as the original sunny dataset. And the images and laser point cloud data are fogged to create a simulated foggy dataset.

2.3.1 Foggy image dataset generation and construction

We use the approach in the literature [16], which uses an atmospheric scattering model to establish the generation process of foggy images.

$$I(x,y) = t(x,y) \cdot J(x,y) + [1 - t(x,y)] \cdot A \tag{6}$$

where, (x, y) are the pixel coordinates; I is the synthetic fog-sky image; J is the original clear image; A is the atmospheric light value; $t(x, y)$ is the transmittance, which indicates the amount of transmittance of the scene reaching the camera. Under uniform medium conditions, $t(x, y)$ depends on the distance from the scene to the camera and is defined as:

$$t(x,y) = \mathrm{e}^{-\beta d(x_1 y)} \tag{7}$$

where, β is the atmospheric scattering coefficient; $d(x, y)$ is the normalized distance of the scene at pixel (x, y).

2.3.2 Foggy laser point cloud dataset generation

Based on the operating principle of LiDAR, the method [17] models the laser transmission and received signal power. The various environment such as sunny, foggy weather is determined by the impulse response of the atmospheric optical channel.

In clear weather, the targets detected from LiDAR are solid objects, referring to this type of targets as hard targets. For both sunny and foggy data in the same 3D scene, the same hard target still exists. However, the scattering of light in foggy days produces excess target data, and this type of targets is referred to as soft targets. The total impulse response function HH(RR) in the fog can be expressed as:

$$H(R)=-\frac{\exp(-2\alpha R)\xi(R)}{R^2}\times[\beta U(R_0-R)+\beta_0\delta(R-R_0)] \tag{8}$$

where, $H_T^{\text{hard}}(R)=\beta_0\delta(R-R_0)$ is the impulse response of the hard target and β_0 denotes the differential reflectance of the target; $H_T^{\text{soft}}(R)=\beta U(R_0-R)$ is the impulse response of the soft fog target and β is the scattering coefficient. The response received by the LiDAR in the fog can then be decomposed into:

$$P_{R,\,\text{fog}}(R)=P_{R,\,\text{fog}}^{\text{hard}}(R)+P_{R,\,\text{fog}}^{\text{soft}}(R) \tag{9}$$

where,

$$R_{R,\text{fog}}^{\text{hard}}(R)=C_A\frac{\exp(-2\alpha R_0)}{R_0^2}\int_0^{2\tau H}P_0\sin^2\left(\frac{\pi}{2\tau H}t\right)\beta_0\delta\left(R-\frac{ct}{2}-R_0\right)\mathrm{d}t=\exp(-2\alpha R_0)P_{R,\text{clear}}(R);$$

$$R_{R,\text{fog}}^{\text{soft}}(R)=C_AP_0\beta\int_0^{2\tau H}\sin^2\left(\frac{\pi}{2\tau H}t\right)\frac{\exp\left[-2\alpha\left(R-\frac{ct}{2}\right)\right]}{\left(R-\frac{ct}{2}\right)^2}\xi\left(R-\frac{ct}{2}\right)U\left(R_0-R+\frac{ct}{2}\right)\mathrm{d}t$$

According to the above equations, the sunny laser point cloud is converted into a foggy laser point cloud by setting the parameters of attenuation coefficient α and scattering coefficient β and hard target reflectivity γ. And corresponding to the conversion parameter β for image foggy simulation, the simulated foggy dataset is created.

3 Results and discussion

The experiment is the comparison experiment, including the comparison of the proposed foggy 3D target detection method and existing detection methods in three scenarios: light fog, medium fog and dense fog.

3.1 Experimental settings

The experiment was conducted in python with a 64-bit operating system, ubuntu 16.04. The deep learning framework is Pytorch 1.10; hardware configuration is CPU Intel Xeon(R) E5-2 678 v3 @ 2.50GHz octa-core and Nvidia GeForce GTX 3 060 Ti; calling GPU acceleration with CUDA 10.2 and cuDNN 8.0.5.

The algorithm is used to build a foggy dataset using the KITTI [15] open source dataset as a benchmark. The dataset is divided into three classes: light fog, medium fog, and dense fog, with the same data scenarios for different concentration classes. And each dense fog class dataset contains 3,712 training sets and 3,769 test sets. Then three levels of easy, medium, and difficult were determined based on target size, occlusion status, and truncation level. Table 1 shows the parameters in the three fog levels. β is atmospheric scattering coefficient, α is laser point cloud attenuation coefficient, γ is hard target reflectance and r is actual detection range.

Table 1 Parameters of the three fog levels.

Parameters	Light fog	Medium fog	Dense fog
β	0.04	0.06	0.08
α	0.075	0.1	0.125
γ	$3e^{-0.6}$	$2e^{-0.6}$	$e^{-0.6}$
r	40m	30m	20m

3.2 Comparison experiments with existing testing methods

In this paper, we follow the official KITTI [15] evaluation protocol and evaluate the vehicle detection results at three concentration levels: light fog, medium fog, and dense fog at LoU = 0.7. AVOD [18], PointRCNN [19], F-Pointnet [20], and SECOND [22] were also trained at three fog concentrations. These are used as the comparison objects to verify the effectiveness of the 3D target detection method in fog in this paper. Table 2 shows the comparison results of the detection performances of different 3D target detection methods trained with light fog, medium fog, and dense fog data.

Fig. 4 plots the PR curves of different 3D target detection methods under light fog, medium fog and dense fog, respectively.

First, from Fig. 4, it can be visualized that the detection accuracy of various 3D target detection algorithms decreases as the fog concentration rises. As the fog concentration increases, the image becomes more blurred so that the target feature information from the image is gradually reduced. Meanwhile, the detection distance of LiDAR decreases with the increase of fog concentration, and the useful point cloud information collected also gradually decreases, and the useless noise information increases. All these factors will reduce the target detection accuracy.

Second, Table 2 shows that the F-pointnet [20] and AVOD [18] algorithms based on image and laser point cloud fusion methods have better detection results than the PointRCNN [19] and SECOND [22] algorithms based on pure laser point clouds under

Table 2 Comparison of vehicle detection performance under light, medium and dense fog conditions: average accuracy (%).

Methods	Fog conditions	Easy		Medium		Difficult	
		BEV	3D	BEV	3D	BEV	3D
PointRCNN [19]	Hight	84.34	71.06	66.33	54.21	58.79	48.60
AVOD [18]		88.86	76.46	79.53	64.47	69.05	57.02
F-pointnet [20]		87.31	79.50	75.86	61.40	66.95	53.52
SECOND [22]		83.92	68.35	60.04	50.71	53.47	43.86
Ours Method		**88.96**	**80.01**	**81.35**	**68.63**	**73.51**	**63.89**
PointRCNN [19]	Medium	83.24	67.65	65.88	53.40	58.54	46.77
AVOD [18]		84.71	71.17	75.97	54.53	66.81	52.79
F-pointnet [20]		85.34	73.81	72.49	52.44	64.01	48.60
SECOND [22]		83.15	65.24	58.91	49.62	53.07	42.98
Ours Method		**86.53**	**76.64**	**76.80**	**62.19**	**69.45**	**58.93**
PointRCNN [19]	Dense	81.99	65.74	64.08	51.85	58.08	45.41
AVOD [18]		82.53	65.98	67.83	51.91	59.26	45.65
F-pointnet [20]		81.36	66.09	63.46	45.95	55.09	40.43
SECOND [22]		81.64	63.45	57.02	47.39	52.46	41.75
Ours Method		**84.79**	**71.34**	**71.95**	**57.32**	**62.83**	**50.17**

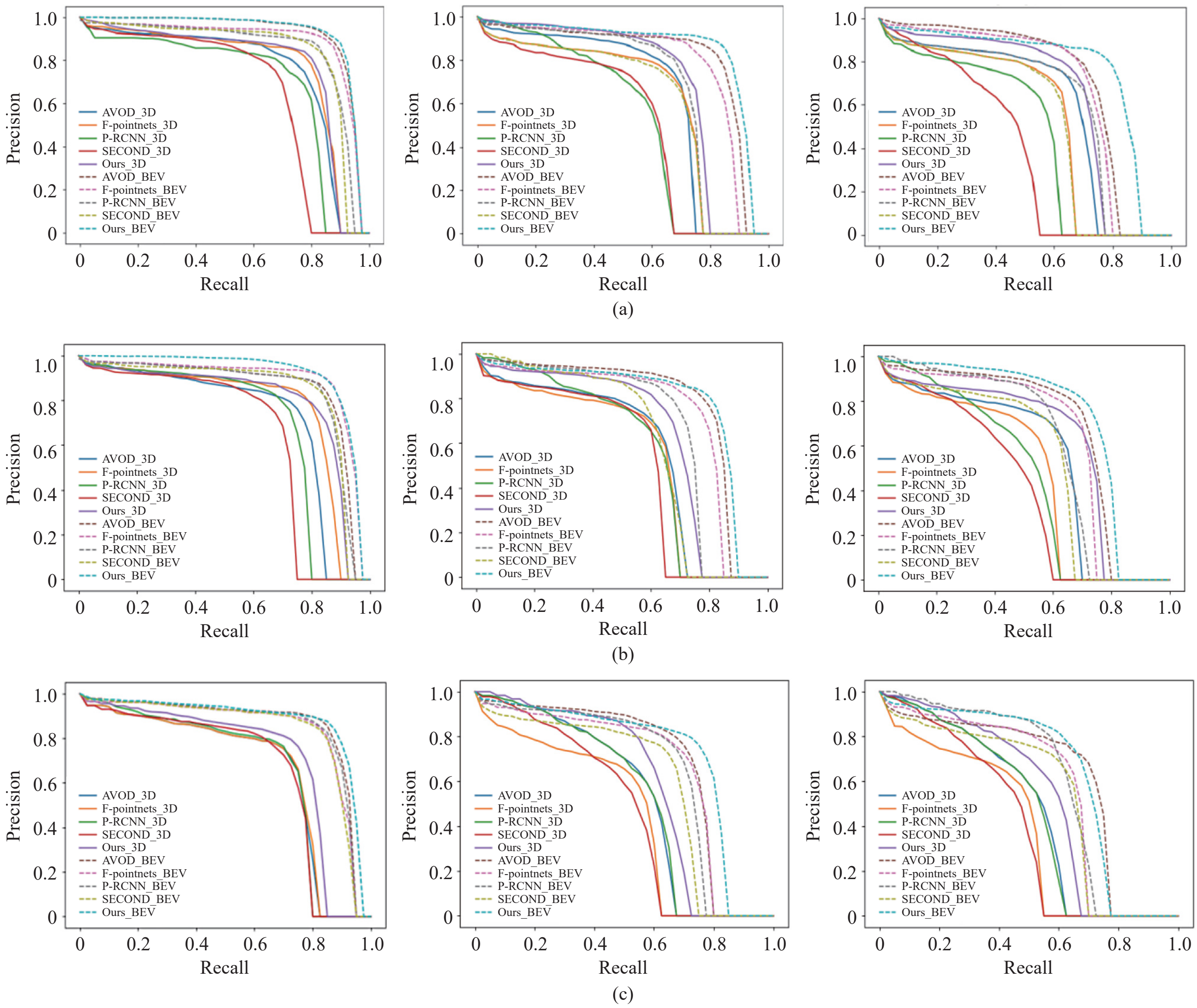

Fig. 4 PR graph for each result under light, medium and dense fog environment. (a) Light fog with easy, medium and difficult modesg. (b) Medium fog with easy, medium and difficult modes. (c) Dense fog with easy, medium and difficult modes.

light fog. However, Table 2 shows that the difference in detection accuracy between the two types of methods is reduced under medium and dense fog. As the fog concentration increases, the image becomes so blurred that cannot provide effective target feature information for the fusion-based detection method, resulting in a significant decrease in detection accuracy.

At last, the following conclusions can be drawn. The performance indexes of the 3D target detection algorithm proposed in this paper are higher than those of the comparison algorithms in all types of fog concentration levels and detection difficulty levels, which also proves the effectiveness of the algorithm. In this paper, the image-based pseudo-point cloud is fused with the original point cloud, and the point cloud density is enhanced so that it can represent the appearance of the target more effectively. Then the laser point cloud is defogged using the defogging method proposed in this paper to further improve the accuracy of target detection under light fog. Compared with the F-pointnet [20] and AVOD [18], this paper defogs the images to make them clearer and more realistic, which is more conducive to data fusion.

4 Conclusions

For the problem of low accuracy of 3D target detection in fog, this paper proposes a 3D target detection method that fuses image and LiDAR data. First, the image and point cloud data are defogged to solve the problem of low data quality due to the presence of a large amount of “noise” in the foggy image and point cloud data. Then a dynamic point cloud defogging method based on light intensity is proposed, which is different from the filtering method using “radius”, but using “intensity + radius”, which has better point cloud defogging effect. After the defogging process is completed, the images are fused with the point cloud data and put into the BtcDet [11] target detector. The experiment proves that this method can effectively remove the influence of fog on the image and point cloud data, and improve the detection accuracy of 3D vehicles in fog.

Acknowledgements

None.

Funding

This work was supported by the Nation Key Research and Development Program of China (No.2019YFB1600100) and was partially supported by the Natural Science Foundation of Shaanxi

Province under Grant 2023-JC-QN-0778 and the Foundation Research Founds for the Central Universities, CHD (Grand No. 300102323101).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author contribution statement

Author 1 was responsible for the main work of the entire study.Authors 2 and 3 provided guidance and revisions throughout the whole process.

Data availability statement

The data supporting the findings of this study are not publicly available because they are currently being used for ongoing follow-up research. The authors will consider reasonable requests for access once the subsequent studies have been completed.

Use of AI statement

No generative AI tools were used in the preparation, writing, or editing of this manuscript. All content was produced entirely by the authors.

References

[1] Afghari AP, Hezaveh AM, Haque MM, et al. A home-based approach to understanding seatbelt use in single-occupant vehicles in Tennessee: Application of a latent class binary logit model. Accident Analysis & Prevention, 2020, 146: 105743.

[2] Zhang JW, Zhang M, Fang ZC, et al. RVDet: Feature-level fusion of radar and camera for object detection. 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Indianapolis: IEEE, 2021: 2822-2828. doi: 10.1109/ITSC48978.2021.9564627.

[3] Li XH, Huo KX, Yan XF, et al. Vehicle detection based on feature-weighted visual enhancement for LiDAR-camera fusion. Journal of Highway and Transportation Research and Development, 2023, 40(2): 182-189.

[4] Charron N, Phillips S, Waslander SL. De-noising of LiDAR point clouds corrupted by snowfall. 2018 15th Conference on Computer and Robot Vision (CRV). New York: IEEE, 2018: 254-261. DOI:10.1109/CRV.2018.00043

[5] Balta H, Velagic J, Bosschaerts W, et al. Fast statistical outlier removal based method for large 3D point clouds of outdoor environments. IFAC-PapersOnLine, 2018, 51(22): 348–353.

[6] Hahner M, Sakaridis C, Dai DX, et al. Fog simulation on real LiDAR point clouds for 3D object detection in adverse weather. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 15263-15272. https://doi.org/10.1109/ICCV48922.2021.01500.

[7] Li H, Liu Y, Xiong SW, et al. Pedestrian detection algorithm based on video sequences and laser point cloud. Frontiers of Computer Science, 2015, 9(3): 402–414.

[8] Yoo JH, Kim Y, Kim J, et al. 3D-CVF: Generating Joint Camera and LiDAR features using cross-view spatial feature fusion for 3D object detection. Computer Vision– ECCV 2020. Cham: Springer International Publishing, 2020: 720-736. https://doi.org/10.1007/978-3-030-58583-9_43.

[9] Jiao Y, Jie ZQ, Chen SX, et al. MSMDFusion: Fusing LiDAR and camera at multiple scales with multi-depth seeds for 3D object detection. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver: IEEE, 2023: 21643-21652. https://doi.org/10.1109/CVPR52729.2023.02073

[10] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2017, 60(6): 84–90.

[11] Xu Q, Zhong Y, Neumann U. Behind the curtain: Learning occluded shapes for 3D object detection. arXiv preprint arXiv:2112.02205, 2021. https://arxiv.org/abs/2112.02205.

[12] Qin X, Wang ZL, Bai YC, et al. FFA-net: Feature fusion attention network for single image dehazing. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 11908–11915.

[13] Kashani AG, Olsen MJ, Parrish CE, et al. Wilson N. A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. Sensors,2015, 15(11): 28099–28128. https://doi.org/10.3390/s151128099

[14] Rasshofer RH, Spies M, Spies H. Influences of weather phenomena on automotive laser radar systems. Advances in Radio Science, 2011, 9: 49–60.

[15] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving?The KITTI vision benchmark suite. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 3354-3361. https://doi.org/10.1109/CVPR.2012.6248074.

[16] Pei YT, Huang YP, Zou Q, et al. Does haze removal help cnn-based image classification?. Proceedings of 15th European Conference. Munich: 2018: 697 - 712. https://doi.org/10.1007/978-3-030-01249-6_42.

[17] Bijelic M, Gruber T, Mannan F, et al. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 11679-11689.

[18] Chabok A, Påhlman L, Hjern F, et al. Randomized clinical trial of antibiotics in acute uncomplicated diverticulitis. British Journal of Surgery, 2012, 99(4): 532–539. https://doi.org/10.1002/bjs.8688.

[19] Shi SS, Wang XG, Li HS. PointRCNN: 3D object proposal generation and detection from point cloud. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 770-779. DOI: 10.1109/CVPR.2019.00086.

[20] Qi CR, Liu W, Wu C, et al. Frustum pointnets for 3D object detection from RGB-D data. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 918–927. https://doi.org/10.1109/CVPR.2018.00102