

AdWeatherNet: Adverse Weather Denoising with Point Cloud Spatiotemporal Attention

Haozheng Han, Dongyu Du, Jie Luo, Xin Jin*

Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China

Abstract—Adverse weather introduces disruptive noise into LiDAR data within autonomous driving systems, compromising the accuracy and range of 3D perception. Mitigating this challenge for high-precision noise removal becomes intricate due to the varying noise distributions at different distances. A novel spatiotemporal denoising network, AdWeatherNet, is proposed to address this problem. The Spatial Encoder module dynamically encodes spatial features using a designed density evaluation model. Additionally, the Temporal Differential Attention module effectively leverages temporal variation in adjacent point clouds to identify and accurately remove noise. To drive the research, we also introduce an adverse weather dataset, named the *AdScenes* dataset, which features point-wise annotations and a wide variety of weather conditions, making it one of the largest comprehensive datasets in this domain. The experimental results demonstrate the effectiveness of our method, with a remarkable improvement of +9.5% of IoU in rainy scenes, +5.9% of IoU in snowy scenes, and +1.9% of IoU in foggy scenes. Compared to the SOTA, AdWeatherNet enhances the mAP of object detection by an average of +1.8% across all weather conditions. Our method contributes to the development of reliable LiDAR perception systems, fostering the development of autonomous vehicles.

Index Terms—point cloud denoising, adverse weather, all-weather perception, autonomous driving

I. INTRODUCTION

In adverse weather, the quality of light detection and ranging (LiDAR) sensor data degrades because of interactions with airborne particles, including rain droplets [1]–[3], snowflakes [3]–[7] and fog [3], [8], [9]. Denoising to enhance the integrity of the raw data has been a perennial pursuit, despite significant efforts, efficient denoising across all weather conditions remains highly challenging. Because various physical propagation result in distinct noise distribution. For instance, the scattering and refraction of airborne particles disturb the signal phase, leading to the convergence of the point clouds and the loss of the target information. It is particularly evident at close distances, as shown in the yellow region in Fig. 1(a). Simultaneously, the reflection and absorption of the airborne particles attenuate the signal intensity or hinder the signal from being received, resulting in sparse point clouds, especially in the distant area, as shown in the red region in Fig. 1(a). This characteristic is referred to as the intrinsic property of point clouds, which is “Near is dense, and far is sparse”. Therefore, establishing a unified denoising framework to enhance the data quality is crucial for autonomous driving and robot systems [10], as demonstrated by our results in Fig. 1(b).

* is corresponding author.

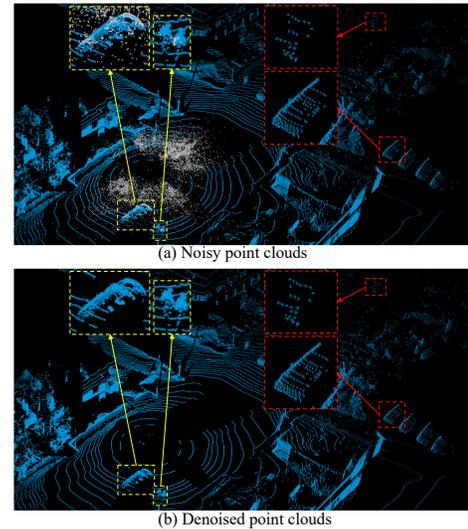


Fig. 1. The comparison between noisy and denoised point clouds in adverse weather. Sparse points are in the red boxes, and dense points are in the yellow boxes. (a) The view of raw point clouds in adverse weather. (b) The view of denoised point clouds by AdWeatherNet.

Both filtering algorithms [11]–[16] and deep-learning algorithms [17]–[19] face challenges in effectively removing noise across all weather conditions due to the difficulty in discerning noise within the dynamic point cloud distribution, characterized by the “Near is dense, and far is sparse” principle. As shown in Table I, the existing adverse weather denoising datasets [3], [13], [18], [20]–[26] are limited in scale and scene diversity, which hinders the model’s learning capacity and can lead to overfitting. Moreover, these datasets typically lack temporal continuity, even though distribution features within a time window tend to be more robust.

In this paper, we initially introduce a denoising dataset, named *AdScenes* dataset, rich in weather scenes for adverse weather denoising, which also stands as one of the most extensive datasets in this domain. Furthermore, we propose a novel denoising network, named AdWeatherNet, designed to identify the dynamic disparity between noise and point clouds across both spatial and temporal channels. Experimental results show that our proposed method outperforms the existing methods, achieving improvements of +2.8%, +6.1%, and +3.6% in Precision, as well as +9.5%, +5.9%, and +1.9% in IoU for rainy, snowy, and foggy scenes, respectively. Furthermore, our method enhances the mAP in object detection by +1.8% across all adverse weather, underscoring its significance for downstream perception tasks.

TABLE I
COMPARISON AMONG MAJOR POINT CLOUD DATASETS.

Dataset	Type	Task	Annotation	Weather			
				Rain	Snow	Fog	Multi-intensities
nuScenes	Real	detection	bbox	✓	×	×	×
Waymo	Real	detection	bbox	✓	×	×	×
Oxford RobotCar	Real	detection	bbox	✓	✓	×	×
CADC	Real	detection	bbox	×	✓	×	×
STF	Real	detection	bbox	✓	✓	✓	×
Panoptic nuScenes	Real	segmentation	point-wise	✓	×	×	×
SemanticKITTI	Real	segmentation	point-wise	✓	×	✓	×
Semantic STF	Real	segmentation	point-wise	✓	✓	✓	×
WADS	Real	denoising	point-wise	×	✓	×	×
SnowyKITTI	Synthetic	denoising	point-wise	×	✓	×	×
<i>AdScenes</i>	Real/Synthetic	denoising/detection	point-wise/bbox	✓	✓	✓	✓

II. THE *AdScenes* DATASET

We have constructed the *AdScenes* dataset, a point-wise annotated point cloud dataset, including all adverse weather scenarios. The dataset is divided into the training and testing parts. The testing part includes real-world rainy and synthesized foggy weather scenes. The rainy data is captured by our self-developed LiDAR and camera array synchronous perception system. The training part is enriched using the SemanticKITTI dataset [23] through the LiDAR Light Scattering Enhancement Model (LISA) [27], a physics-based simulation model of particles in the air. This part includes 3 weather scenes: rain, snow, and fog, each with 3 intensity levels: moderate, heavy, and extreme. It covers 10 traffic scenarios, including roads, intersections, sidewalks, parking lots, highways, vegetation, terrain, buildings, poles, and others, and involves 8 types of traffic participants, including cars, buses, bicycles, motorcycles, pedestrians, cyclists, motorcyclists, and other vehicle. The whole dataset is approximately 23 GB with 12.1 billion point-wise annotations. The *Adscenes* dataset categorizes noise as a positive sample (label: 1) and the clean background, including vehicles and pedestrians, etc, as a negative sample (label: 0).

As can be seen from Table I, the proposed *Adscenes* dataset stands out as one of the most extensive datasets of all-weather scenes in the field of point cloud denoising. In terms of the diversity of weather scenes, this dataset also surpasses the existing dataset focused on point cloud semantic denoising.

III. METHOD

Our proposed framework, named AdWeatherNet, takes 3D U-Net as the backbone as shown in Fig. 2(a). Except for the U-Net, it mainly consists of 3 parts: (1) SE utilizes the density prediction model introduced in this paper to dynamically encode features, thereby enhancing the efficiency of encoding the point cloud characterized by “Near is dense, and far is sparse”, as shown in Fig. 2(b). (2) TDA utilizes the variation of adjacent point clouds in the temporal dimension to leverage the consistent features, as shown in Fig. 2 (c). (3) Semantic Outlier Removal is a filter to remove noisy points caused by airborne particles in adverse weather by semantic information extracted by the segmentation head, in which there is a 3D convolution layer with $3 \times 3 \times 3$ kernel.

A. Spatial Encoder

The existing encoding [17]–[19] often results in a notable loss of features from distant point clouds, which are equally important. To tackle this problem, we propose SE, a dynamic encoder designed to assist the model in learning distinct features at various distances. Assuming a uniform distribution of multi-class targets in the outdoor scene, the density of point clouds is inversely proportional to the square of the distance. We thereby formulate a non-linear regression to predict the density of point clouds on the distance, which is given by:

$$\sigma = \gamma \frac{Num}{d^2}, \quad (1)$$

where density σ , density-distance coefficient γ , the total number of point clouds Num , and the distance d . The laser experiences exponential attenuation due to the scattering of airborne particles like raindrops. Therefore, the coefficient of kernel is calculated based on the density prediction in Eq. (1).

$$\delta = \frac{\sigma}{\sigma_0} \frac{d}{d_{\max}} \sqrt{e^{2\alpha d}}, \quad (2)$$

where δ , σ_0 , and d_{\max} are kernel size, the expected density, and the max measured range, respectively. α is the scattering coefficient of airborne particles.

Based on Eq. (2), we propose a density-adaptive 3D convolution that employs a larger kernel size to capture more features in sparser point clouds. This allows us to extract additional features with a larger receptive field. The final output is obtained using max pooling to exploit the maximum value of features.

B. Temporal Differential Attention

The TDA module is designed to harness larger discrepancies in noisy features relative to clean features. This understanding allows us to forecast a differential attention map a_d using the sigmoid confidence as specified in Eq. (3). The function $h(\cdot)$ signifies a submanifold sparse convolution that aims to extract features while maintaining the sparsity of point clouds. The features f_{t-1} and f_t are the extracted features at time $t-1$ and t respectively, each having dimensions of $N \times C$.

$$a_d = \text{sigmoid}(h(f_{t-1}) \cdot (f_t - f_{t-1})), \quad (3)$$

$$f_d = a_d f_t, \quad (4)$$

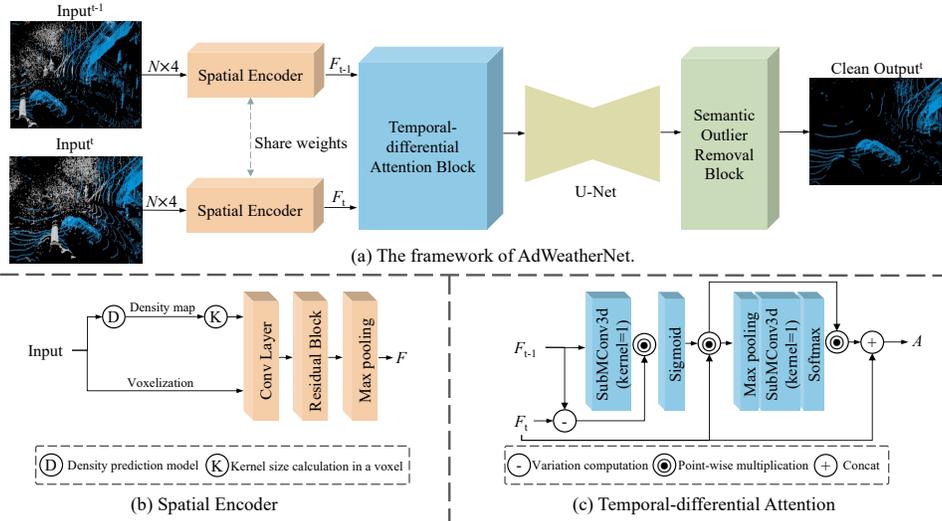


Fig. 2. The architecture of the proposed AdWeatherNet. Here white points are noise and blue points are clean environments. The inputs are 2 scans of adjacent point clouds, and the output is a clean point cloud. (a) The framework of AdWeatherNet. (b) The illustration of SE. (c) The illustration of TDA.

$$a_c = \text{softmax}(h(p(f_d))), \quad (5)$$

$$f_c = a_c f_d, \quad (6)$$

$$O = f_c + f_t, \quad (7)$$

where f_c and f_c are differential attentive features and channel attentive features, respectively. O is the final output feature of the TDA module. The channel attention map a_c employs the max value of differential attention from each voxel of point clouds, using softmax confidences as described in Eq. (5). The function $p(\cdot)$ represents a max pooling operation. The variation in clean information is subtle, causing it to be overshadowed by noisy features in the differential attention map. To address this issue, a residual block is incorporated. Features from the current frame of point clouds are point-wise added to the channel attention map, resulting in the final output.

C. Network and Training Details

The proposed denoising network includes 3 key steps. Firstly, SE and TDA are applied to extract spatiotemporal features from the input raw point cloud data. Secondly, these features are introduced into a U-Net architecture with a segmentation head, which contains a 3D convolution layer with $3 \times 3 \times 3$ kernel size to perform further processing on the data. To remove noisy points caused by airborne particles such as snowflakes and rain droplets, Softmax confidences are utilized with a mask map that is normalized in the range of $[0, 1]$. Finally, the semantic outlier removal is applied to remove the predicted noisy points and yield cleaner point clouds.

The training process involves a weight-wise sum of two loss functions to optimize the network performance.

$$\mathcal{L} = \beta \mathcal{L}_{ls} + (1 - \beta) \mathcal{L}_{ce}, \quad (8)$$

where \mathcal{L}_{ls} is a Lovász Softmax [28] loss and \mathcal{L}_{ce} is a standard cross-entropy loss. β is a weighted coefficient.

IV. EXPERIMENTS AND RESULTS

Our AdWeatherNet is compared with 5 state-of-the-art (SOTA) algorithms that are representative of point cloud denoising. DROR* [12] and DSOR* [13] are the state-of-the-art methods among density-based outlier removal methods. LIOR [14] is the representative low-intensity outlier removal method. WeatherNet [17] is the first cnn-based deep learning method. 4DenoiseNet [18] is the first network that uses the temporal information of adjacent point clouds.

A. Quantitative and Qualitative Results

As shown in Table II, our proposed method improves the Precision by +2.8%, the Recall by +6.2% and the IoU by +9.5% when tested in the rainy scenes. It additionally achieves superior performance compared to the SOTA algorithms, exhibiting improvements of +5.9% and +1.9% in IoU, as well as +6.1% and +3.6% in Precision for snowy and foggy scenes, respectively. Among all weather conditions, ours achieves no lower than a high Recall of 96.4%. It proves that AdWeatherNet has the best performance in robust and accurate denoising across all weather conditions. It is noteworthy that the high Precision and Recall signify AdWeatherNet's effectiveness in both denoising and feature extraction.

As shown in Fig. 3, the results of filtering methods, such as DROR, DSOR, and LIOR preserve more noise than the deep learning methods. WeatherNet exhibits poorer performance than the other deep learning methods. AdWeatherNet is slightly better than 4DenoiseNet for denoising. When focusing on pedestrians and vehicles in scenes, AdWeatherNet preserves more complete contours of these objects than 4DenoiseNet. In summary, AdWeatherNet has the best performance of denoising among all the methods.

As shown in Table IV, ablation study results show that the baseline with the Baseline-TDA achieves the second-highest precision of 96.2% and the highest recall of 98.6%. The TDA module slightly outperformed the SE module in enhancing overall performance.

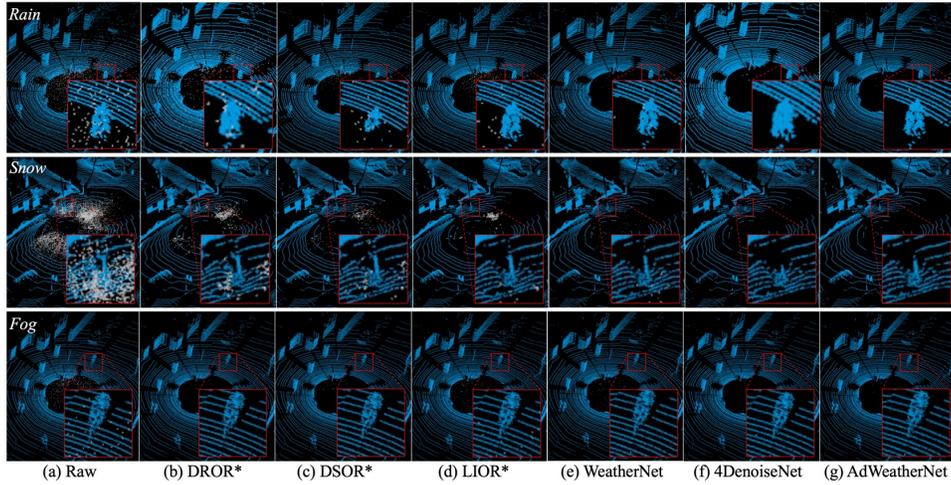


Fig. 3. Visual results of point cloud denoising algorithms. (a) raw point clouds, where rainy scenes are on the top and snowy scenes are on the bottom. (b) Results of DROR filter. (c) Results of DSOR filter. (d) Results of LIOR filter. (e) Results of WeatherNet. (f) Results of 4DenoiseNet. (g) Results of ours. *: no training required.

TABLE II
QUANTITATIVE EVALUATION. TRAINING AND INFERENCE ARE BOTH CONDUCTED ON NVIDIA GTX 3090 GPU.
*: NO TRAINING REQUIRED. **BOLD FONT**: THE STATE-OF-THE-ART RESULTS.

Weather	Rain			Snow			Fog		
	Precision	Recall	IoU	Precision	Recall	IoU	Precision	Recall	IoU
DROR*	83.4	91.2	81.6	71.5	91.9	63.6	76.1	89.9	60.9
DSOR*	79.8	91.4	77.5	76.6	94.7	74.6	79.5	93.3	72.4
LIOR*	70.6	80.7	59.8	72.4	83.5	63.3	70.6	80.1	60.7
WeatherNet	88.4	87.2	86.4	87.6	97.8	85.0	83.7	95.2	76.5
4DenoiseNet	96.5	92.3	88.3	91.5	98.7	90.3	93.2	96.2	85.2
AdWeatherNet	99.3	98.5	97.8	97.6	98.5	96.2	96.8	96.4	87.1

TABLE III
THE IMPROVEMENT OF OBJECT DETECTION IN ADVERSE WEATHER.

Weather	mAP			
	Raw	4DenoiseNet	AdWeatherNet	Gain
Rain	73.52	76.33	79.91	+3.58
Snow	69.78	73.33	74.59	+1.26
Fog	75.87	77.52	78.12	+0.60

TABLE IV
ABLATION STUDY.

Metrics	Precision	Recall	IoU
Baseline	93.4	98.1	91.8
Baseline-SE	95.8	98.5	94.4
Baseline-TDA	96.2	98.6	94.9
Full Model	97.6	98.5	96.2

B. Enhanced Object Detection

Noise can significantly affect 3D object detection by distorting the object geometry. Our proposed method focuses on removing noise from point clouds, leading to significant improvements in object detection. We choose Pointpillars [29] as an object detection network and compare the results of 3D object detection using raw data with the denoised data from our method and 4DenoiseNet. As shown in Table III, denoised point clouds improve the mAP of object detection by +6.39%, +4.81%, and +2.25% in rainy, snowy, and foggy scenes. Compared with the SOTA algorithm, it improves by +3.58%, +1.26%, and +0.60% respectively. In foggy scenes, there is a relatively minor improvement in the mAP of object detection. This is attributed to the fact that fog has a stronger impact on the absorption and refraction of laser beams than on reflection. In this particular scene, noise is not the primary influencing factor.

V. CONCLUSION

AdWeatherNet, a spatiotemporally attentive denoising network, is proposed to improve the quality of point clouds in

adverse weather. It establishes a new SOTA benchmark in point cloud denoising in adverse weather, including rain, snow, and fog. It enhances both Precision and Recall, reflecting a decrease in the inaccuracies associated with false positives and missed recognitions of noise. It proves that AdWeatherNet improves the mAP of 3D object detection in adverse weather as well, which shows the potential to become an essential component for LiDAR perception systems. The *AdScenes* dataset is one of the largest and most diverse point cloud semantic denoising datasets. Building upon our work, a promising avenue for future research is real-time 3D perception for autonomous driving systems.

ACKNOWLEDGMENT

The work was supported in part by Shenzhen Science and Technology Project under Grant JSGG20220831095602005, and Natural Science Foundation of China under Grant 61991451, China.

REFERENCES

- [1] Goodin C, Carruth D, Doude M, et al. Predicting the Influence of Rain on LIDAR in ADAS[J]. *Electronics*, 2019, 8(1): 89.
- [2] Wallace A M, Halimi A, Buller G S. Full waveform LiDAR for adverse weather conditions[J]. *IEEE transactions on vehicular technology*, 2020, 69(7): 7064-7077.
- [3] Bijelic M, Gruber T, Mannan F, et al. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 11682-11692.
- [4] Jokela M, Kutila M, Pyykönen P. Testing and validation of automotive point-cloud sensors in adverse weather conditions[J]. *Applied Sciences*, 2019, 9(11): 2341.
- [5] Kutila M, Pyykönen P, Jokela M, et al. Benchmarking automotive LiDAR performance in arctic conditions[C]//*2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020: 1-8.
- [6] Michaud S, Lalonde J F, Giguere P. Towards characterizing the behavior of LiDARs in snowy conditions[C]//*International Conference on Intelligent Robots and Systems (IROS)*. 2015.
- [7] O'Brien H W. Visibility and light attenuation in falling snow[J]. *Journal of Applied Meteorology and Climatology*, 1970, 9(4): 671-683.
- [8] Bijelic M, Gruber T, Ritter W. A benchmark for lidar sensors in fog: Is detection breaking down?[C]//*2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2018: 760-767.
- [9] Heinzler R, Schindler P, Seekircher J, et al. Weather influence and classification with automotive lidar sensors[C]//*2019 IEEE intelligent vehicles symposium (IV)*. IEEE, 2019: 1527-1534.
- [10] Wu J, Xu H, Tian Y, et al. Vehicle detection under adverse weather from roadside LiDAR data[J]. *Sensors*, 2020, 20(12): 3433.
- [11] Rusu R B, Cousins S. 3d is here: Point cloud library (pcl)[C]//*2011 IEEE international conference on robotics and automation*. IEEE, 2011: 1-4.
- [12] Charron N, Phillips S, Waslander S L. De-noising of lidar point clouds corrupted by snowfall[C]//*2018 15th Conference on Computer and Robot Vision (CRV)*. IEEE, 2018: 254-261.
- [13] Kurup A, Bos J. Dsor: A scalable statistical filter for removing falling snow from lidar point clouds in severe winter weather[J]. *arXiv preprint arXiv:2109.07078*, 2021.
- [14] Park J I, Park J, Kim K S. Fast and accurate desnowing algorithm for LiDAR point clouds[J]. *IEEE Access*, 2020, 8: 160202-160212.
- [15] Wang W, You X, Chen L, et al. A scalable and accurate de-snowing algorithm for LiDAR point clouds in winter[J]. *Remote Sensing*, 2022, 14(6): 1468.
- [16] Li B, Li J, Chen G, et al. De-snowing LiDAR point clouds with intensity and spatial-temporal features[C]//*2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022: 2359-2365.
- [17] Heinzler R, Piewak F, Schindler P, et al. Cnn-based lidar point cloud denoising in adverse weather[J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 2514-2521.
- [18] Seppänen A, Ojala R, Tammi K. 4denoisenet: Adverse weather denoising from adjacent point clouds[J]. *IEEE Robotics and Automation Letters*, 2022, 8(1): 456-463.
- [19] Yu M Y, Vasudevan R, Johnson-Roberson M. Lisownet: Real-time snow removal for lidar point clouds[C]//*2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022: 6820-6826.
- [20] Caesar H, Bankiti V, Lang A H, et al. nuscenes: A multimodal dataset for autonomous driving[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020: 11621-11631.
- [21] Fong W K, Mohan R, Hurtado J V, et al. Panoptic nuscenes: A large-scale benchmark for lidar panoptic segmentation and tracking[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 3795-3802.
- [22] Barnes D, Gadd M, Murcutt P, et al. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset[C]//*2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020: 6433-6438.
- [23] Behley J, Garbade M, Milioto A, et al. Semantickitti: A dataset for semantic scene understanding of lidar sequences[C]//*Proceedings of the IEEE/CVF international conference on computer vision*. 2019: 9297-9307.
- [24] Sun P, Kretzschmar H, Dotiwalla X, et al. Scalability in perception for autonomous driving: Waymo open dataset[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020: 2446-2454.
- [25] Pitropov M, Garcia D E, Rebello J, et al. Canadian adverse driving conditions dataset[J]. *The International Journal of Robotics Research*, 2021, 40(4-5): 681-690.
- [26] Xiao A, Huang J, Xuan W, et al. 3d semantic segmentation in the wild: Learning generalized models for adverse-condition point clouds[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023: 9382-9392.
- [27] Kilic V, Hegde D, Sindagi V, et al. Lidar light scattering augmentation (lisa): Physics-based simulation of adverse weather conditions for 3d object detection[J]. *arXiv preprint arXiv:2107.07004*, 2021.
- [28] Berman M, Triki A R, Blaschko M B. The lovasz-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018: 4413-4421.
- [29] Lang A H, Vora S, Caesar H, et al. Pointpillars: Fast encoders for object detection from point clouds[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019: 12697-12705.