Sunshine to Rainstorm: Cross-Weather Knowledge Distillation for Robust 3D Object Detection

Xun Huang[1*], **Hai Wu**[1*], **Xin Li**[2], **Xiaoliang Fan**[1], **Chenglu Wen**[1†], **Cheng Wang**[1]

[1] Xiamen University,
[2] Texas A & M University
{huangxun,wuhai}@stu.xmu.edu.cn, {clwen, fanxiaoliang, cwang}@xmu.edu.cn, xinli@tamu.edu

Abstract

LiDAR-based 3D object detection models inevitably struggle under rainy conditions due to the degraded and noisy scanning signals. Previous research has attempted to address this by simulating the noise from rain to improve the robustness of detection models. However, significant disparities exist between simulated and actual rain-impacted data points. In this work, we propose a novel rain simulation method, termed DRET, that unifies **D**ynamics and **R**ainy **E**nvironment **T**heory to provide a cost-effective means of expanding the available realistic rain data for 3D detection training. Furthermore, we present a **S**unny-to-**R**ainy **K**nowledge **D**istillation (**SRKD**) approach to enhance 3D detection under rainy conditions. Extensive experiments on the Waymo-Open-Dataset show that, when combined with the state-of-the-art DSVT model and other classical 3D detectors, our proposed framework demonstrates significant detection accuracy improvements, without losing efficiency. Remarkably, our framework also improves detection capabilities under sunny conditions, therefore offering a robust solution for 3D detection regardless of whether the weather is rainy or sunny.

Introduction

Recent years have witnessed growing research interest in 3D object detection utilizing point cloud data. For widespread application in autonomous driving, such models must demonstrate robust performance under diverse conditions. Despite progress in 3D object detection on benchmark datasets, achieving consistent and reliable performance, particularly under adverse weather conditions (e.g., rain) remains an open challenge.

Unfortunately, 3D object detection research under rainy weather presents significant challenges at both the data and methods levels. At the data level, a pressing issue is data scarcity. Extant datasets exhibit highly limited rainy samples owing to the high annotation and collection costs. For instance, merely 0.6% of samples in the Waymo Open Dataset (WOD) (Sun et al. 2020) perception subset in rain. This shortage seriously hinders the research on 3D object detection under rainy weather. The low quality of rainy data

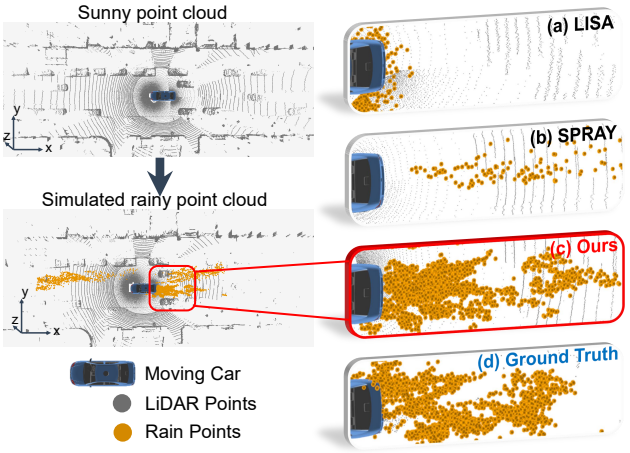

Figure 1: Visualization of rain noise points in groundtruth and different rain simulations.

obstructs 3D detectors. Analysis of rainy point clouds in the WOD reveals two critical phenomena: 1) *Dense rain noise.* As LiDAR light pulses cannot penetrate water particles (Hahner et al. 2021), resulting in noise from water droplets generated by moving vehicles. 2) *Missing points.* Atmospheric parameters like attenuation substantially differ in rain. Thus, many points fall below the LiDAR intensity threshold and are missed. These phenomena create a considerable domain gap between sunny and rainy data. Robust rainy 3D object detection needs to address both data and model challenges.

Prior work, including LISA (Kilic et al. 2021) and SPRAY (Shih et al. 2022), has sought to address the challenge of insufficient training data by simulating adverse weather conditions. While these approaches are valuable, they overlook the rainy environmental theory and object-scene dynamics, leading to limitations in the realism of simulated rain. For example, as depicted in Fig. 1, a significant gap exists between simulated and real rain points, especially in the case of LISA. SPRAY also fails to simulate the phenomenon of missing points that usually occur in rainy scenes.

In addition to creating realistic simulations, the effective utilization of simulated data is also crucial. Many existing methods (Teufel et al. 2022; Hahner et al. 2021, 2022) simply augment training data with simulations, but this ap-

*These authors contributed equally.
†Corresponding author.

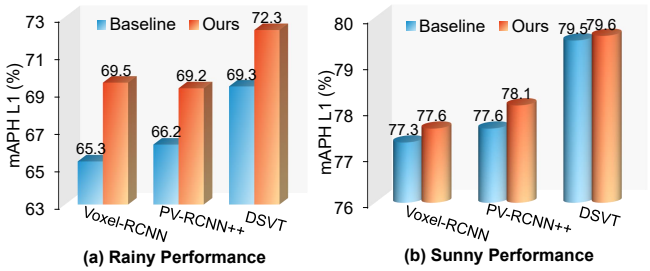

Figure 2: Comparison of the performance under both sunny and rainy weather conditions on the WOD.

proach often falls short in adapting detectors effectively to new rainy environments. LDNet (Do and Yoo 2022) designed an innovative idea of knowledge distillation from sunny to severe weather conditions. Without a realistic rain simulation, this concept cannot be effectively applied. Moreover, LDNet also overlooks the data disparities between different weather conditions. Only by fully addressing the gaps between rainy and sunny weather can models achieve robustness across diverse weather conditions.

To enhance the robustness of models in rain, we propose innovative approaches at both the data and model levels. We introduce DRET, a method that unifies **D**ynamics and **R**ainy **E**nvironment **T**heory to create realistic rain data simulations. We also present SRKD, a **S**unny-to-**R**ainy **K**nowledge **D**istillation framework specifically tailored for 3D detectors in rainy conditions. SRKD trains a student network to learn from a sunny teacher detector, distilling knowledge of the same scene under rainy weather. Additionally, we design a noise-aware prediction correction module to effectively handle the noise associated with rain.

Our framework is highly adaptable and can be easily integrated with various 3D detectors. As shown in Fig.2, extensive experiments on the WOD validated that DRET and SRKD enhance the rain robustness of all 3D detectors without sacrificing efficiency. Remarkably, they also improve detectors' performance under sunny conditions, possibly due to their increased robustness to sparse objects.

Our contributions are : (1) We analyzed the impact of various phenomena on 3D object detection in rainy weather, and proposed a new method, DRET, for more realistic rain simulation in 3D object detection. (2) We designed SRKD, a generalized framework to address the challenges posed by weather disparities in straightforward sunny-to-rainy distillation. It can significantly improve the robustness of the 3D detector in rain without compromising efficiency. (3) We validated the effectiveness and universality of our proposed approach through extensive testing on the state-of-the-art model, DSVT (Wang et al. 2023), as well as other classical 3D detectors, including PV-RCNN++ and Voxel-RCNN.

Related Work

Simulation of rain. Physics-based simulations (Teufel et al. 2022; Hahner et al. 2022, 2021; Kilic et al. 2021) have been explored to reproduce point clouds under adverse weather conditions such as snow and fog. These methods help alleviate the issue of data scarcity to some extent, but they are not fully effective in simulating rainy weather. For instance, LISA (Kilic et al. 2021) employs a rain simulation algorithm that relies on LiDAR light scattering augmentation, but it overlooks the impact of dense rain noise in rainy conditions. SPRAY (Shih et al. 2022) uses dynamics to simulate the water splashing effect caused by vehicles in the rain, but it fails to accurately mimic the actual distribution of splashed water due to the absence of a well-grounded physical theory. Consequently, substantial discrepancies persist between these simulated rainy datasets and actual rainy conditions.

3D object detection. Current 3D object detection methods primarily focus on clear weather conditions and can be categorized into single-stage and two-stage approaches. Single-stage methods, such as SECOND (Yan, Mao, and Li 2018), PointPillars (Deng et al. 2021), SA-SSD (He et al. 2020), and SE-SSD (Zheng et al. 2021), utilize voxel-based sparse convolution or point-based set abstraction for feature extraction. In two-stage algorithms, Voxel-RCNN (Deng et al. 2021), SFD (Wu et al. 2022), and VirConv (Wu et al. 2023) employ voxel-based sparse convolution, while PointRCNN (Shi, Wang, and Li 2019) and STD (Yang et al. 2019) use point-based set abstraction. Notably, PV-RCNN (Shi et al. 2020), CT3D (Sheng et al. 2021), and PV-RCNN++ (Shi et al. 2022) combine voxel-based and point-based operations in the two-stage approach. DSVT (Wang et al. 2023) introduces an efficient and deployable 3D transformer backbone that achieves state-of-the-art performance on WOD. Knowledge distillation has been explored in 3D object detection, including SparseKD (Yang et al. 2022), which examines lightweight models, and other works (Zheng et al. 2022a,b) attempting distillation from multimodal/multi-frame models to single-modal/single-frame models. However, these mainstream methods still lack robustness in rain.

3D object detection in adverse weather. Recent studies, such as (Charron, Phillips, and Waslander 2018; Heinzler et al. 2020), have introduced lightweight approaches that employ techniques like semantic segmentation or filtering algorithms to remove dense noise, thereby minimizing its impact on detection models. However, these methods heavily rely on the denoising model's performance, leading to unstable detection results, and they do not account for the issue of missing points in rainy weather. (Xu et al. 2021) designed a general completion framework that addresses the problem of domain adaptation across different weather and environmental conditions. It achieves promising improvements on the WOD domain adaptation dataset, which contains a significant amount of rainy data. However, it solely focuses on the phenomenon of missing points, disregarding the impact of other phenomena.

Method

The challenges associated with data limitations and model adaptivity create significant obstacles for robust object detection in rainy conditions. Therefore, this work proposes two main approaches: the simulation of realistic rain and

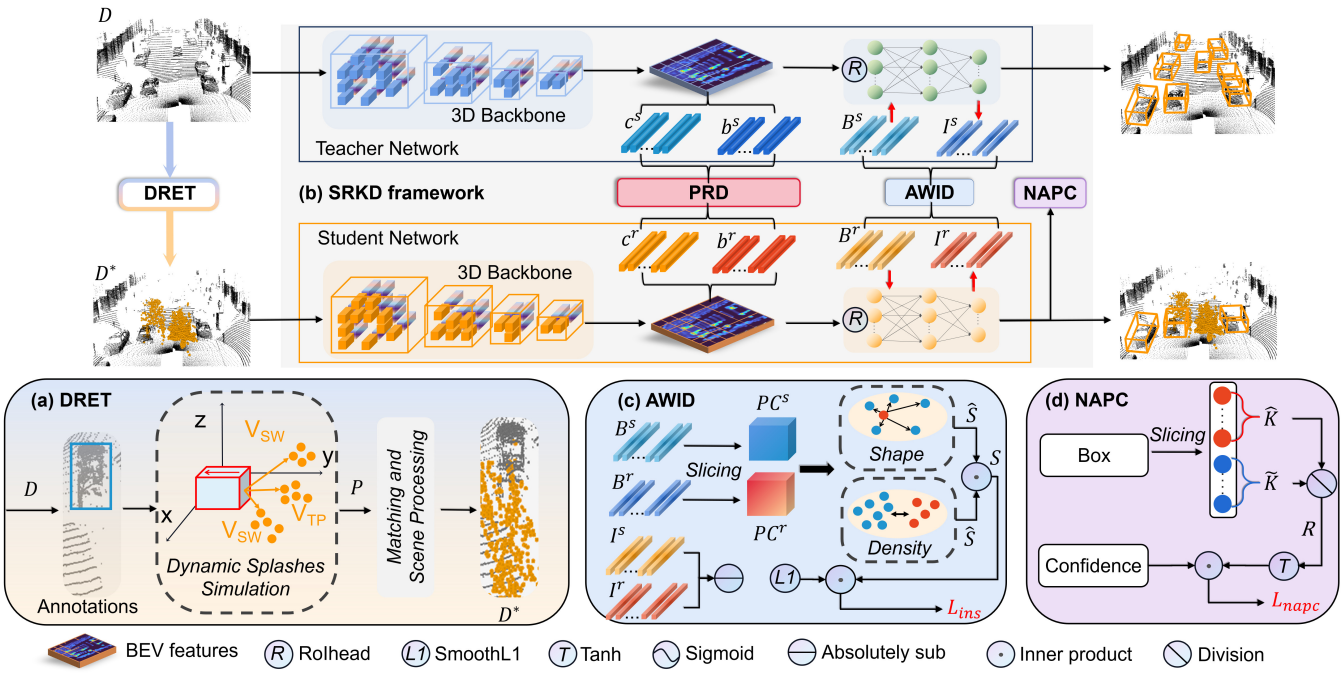

Figure 3: The overview of our method, including (a) DRET for rain simulation and (b) SRKD framework for 3D object detection. DRET involves a two-stage process. The first stage simulates the dynamic splashes to generate rain particles, and the second stage matches rain particles with the point cloud and processes the scene based on the rainy environment theory. SRKD enables sunny-to-rainy knowledge distillation with the help of (c) AWID and PRD, which avoids the problems associated with the weather domain gap. PRD is a classic response distillation with some adjustments, which will be introduced in the SRKD method. Additionally, the (d) NAPC module is designed to mitigate the influence of rain-induced noise.

new object detection tailored for rainy environments. As illustrated in Fig. 3, we introduce DRET (a), a method that generates more accurate and realistic simulated rain data for detector training. Additionally, we present SRKD (b), a general framework that enhances the detector's robustness in rainy conditions without compromising efficiency.

DRET Rain Simulation Method

Rain simulation models take a sunny point cloud data D as input and generate a simulated rainy point cloud D^*. Existing simulation models such as SPRAY (Shih et al. 2022) and LISA (Kilic et al. 2021) have not considered both the dynamic particle simulation and rainy environment theory in a unified pipeline. We propose a new integrated, two-stage simulation method. As depicted in Fig. 3 (a), its first stage employs the Unity3D engine to simulate dynamic splashing. It also utilizes Perlin noise (Bridson, Houriham, and Nordenstam 2007) to replicate wind interference. However, this particle simulation alone doesn't account for LIDAR reflection intensity for rain particles. Thus, we additionally introduce a second stage of performing a scene process following the rainy environment theory (Kilic et al. 2021; Hahner et al. 2021) to address this limitation.

In the first stage, similar to SPRAY (Shih et al. 2022), we utilize the particle emitter in the dynamics particle system to simulate water splashes caused by moving vehicles. We incorporate three splash mechanisms: bow wave (BW), side wave (SW), and tread pickup (TP) following SPRAY. But unlike SPRAY, we use Perlin noise to introduce random accelerations along the x, y, and z axes. This extra design better imitates wind perturbations and contributes to a broader range of realistic rain particle behaviors.

After the initial stage, we obtain sets of rain particles. However, merging them directly into the sunny point clouds, as done in SPRAY, is not a good idea. Because it results in the absence of LiDAR intensity information for rain particles, and the intensity of the original points would remain unadjusted for rainy conditions. Therefore, we introduce an additional scene processing step based on the rainy environment theory in the second stage.

Before conducting the scene processing, we collect the location, number, inclination and maximum distance parameters of the LiDARs used in the dataset. These parameters are used to establish a correspondence between the point cloud data D and the rain particle set P. This gives us matched pairs $\{(D_i, P_j)\}$, in which each pair of points originates from the same laser beam reflection.

Next, unlike LISA's random generation of rain particles, we employ more realistic rain particle sets P for scene processing. We also adopt more accurate formulas from fog simulation (Hahner et al. 2021). Specifically, for a matched pair (D_i, P_j), where D_i is a point with reflection intensity I_i and distance R_i, and P_j is a rain particle with distance R_j, we calculate the intensity I_j of P_j using

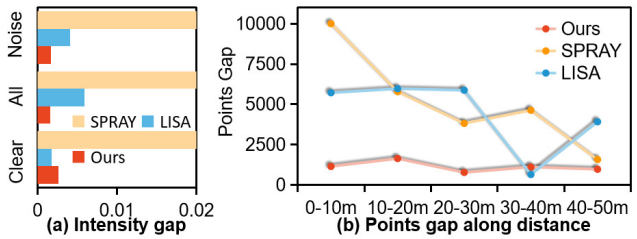

Figure 4: The average intensity gap. (b) The average points gap in each distance interval.

$$I_j = C_A P_0 \beta \int_0^{2\tau_H} \sin^2\left(\frac{\pi}{2\tau_H} t\right) \frac{\exp\left(-2\alpha\left(R_j - \frac{ct}{2}\right)\right)}{\left(R_j - \frac{ct}{2}\right)^2} \times \times \gamma\left(R_j - \frac{ct}{2}\right) U\left(R_i - R_j + \frac{ct}{2}\right) dt, \tag{1}$$

where R_1, R_2 are the sensor optical parameters, β_0 is the differential reflectance of point, β is the scattering rate, α is the attenuation coefficient, c is the speed of light, τ_H is the half-power pulse width, U is the Heaviside function, and

$$C_A P_0 = I_i \frac{R_i^2}{\beta_0}, \gamma(R) = \begin{cases} 0, & R \le R_1 \\ \frac{R-R_1}{R_2-R_1}, & R_1 < R < R_2 \\ 1, & R_2 \le R \end{cases}.$$

The calculation is performed using Simpson's $^1/_3$ rule for numerical integration.

After calculating the intensity I_j of P_j, we use P_j to replace D_i in D to simulate the occlusion of dense noise in real rain. This process yields a new point cloud D'. Then, following LISA (Kilic et al. 2021), we adjust the part of original sunny points in D' to reflect rainy weather conditions. We collect points where the LiDAR received power $\mathcal{P}$ exceeds the minimum LiDAR received power $\mathcal{P}_{min}$ to get the simulated rainy point cloud D^*:

$$D^* = \{D'_i \mid \mathcal{P}_i \ge \mathcal{P}_{min}\}. \tag{2}$$

where $\mathcal{P}_i = \frac{I_i}{R_i^2}$, $\mathcal{P}_{min} = 0.9 R_{max}^2$, and R_{max} is the max distance of LiDAR.

Comparison with other rain simulation methods. We simulated 10k rainy point cloud data using LISA (Kilic et al. 2021), SPRAY (Shih et al. 2022), and our DRET method, respectively. We then analyzed the average intensity gap between these simulated data and real rain data (see Fig. 4 (a)). The SPRAY has a serious gap in intensity with the real data, mainly because of the missing rain particle intensity calculations and original clear point corrections, as mentioned earlier. Although LISA seems to have a smaller intensity gap in the clear points this is a fluctuation caused by the randomization of the rain rate parameter. The intensity gap of our method is significantly less than LISA in terms of noise and all points, especially on all points. By analyzing the average points gap along distance (see Fig. 4 (b)), we can also find that SPRAY has a huge gap because it can not filter the low received power points. While LISA has a small points gap at 30-40m, the overall points gap fluctuates greatly and the overall gap is significantly larger than our method. In contrast, our simulated data is smaller in both the intensity gap and points gap, indicating that our method is more realistic.

In summary, our DRET simulation method can offer a lot of realistic simulated rain data, thereby broadening the possibilities for research into robust 3D object detection in rain.

Weather	Precision			Recall		
	0.3	0.5	0.7	0.3	0.5	0.7
Sunny	46.8	44.8	34.6	87.3	84.1	69.7
Rainy	33.6	32.3	27.3	83.2	80.0	66.7
Degradation	*-13.2*	*-12.5*	*-7.3*	*-4.1*	*-4.1*	*-3.0*

Table 1: Precision and recall of vehicle in different weather. 0.3, 0.5, and 0.7 are the IoU thresholds.

Sunny-to-Rainy Knowledge Distillation

Expanding rainy data using DRET as a data augmentation during training is valuable. However, this strategy alone does not address the gaps existing between sunny and rainy data. Even applying distillation techniques as seen in previous works (Zheng et al. 2022a,b; Do and Yoo 2022) cannot effectively bridge the weather-related gaps. In light of this, we undertake an initial analysis of how these weather disparities impact 3D object detection. Based on these analyses, we design the SRKD framework to overcome the challenges posed by the weather gaps when distilling knowledge from sunny to rainy conditions. As shown in Fig.3 (b), the SRKD framework brings the following technical contributions. Firstly, we devise a knowledge distillation approach capable of facilitating the transition of 3D detectors from sunny to rainy scenarios. This includes the integration of *Adaptively Weighted Instance Distillation (AWID)* and *Precise Response Distillation (PRD)*. Additionally, we introduce *Noise-Aware Prediction Correction (NAPC)* to mitigate the influence of rain-induced noise.

Analysis on impact of rain on 3D object detection. As previously discussed, the presence of dense rain noise and the absence of points in rainy conditions pose significant challenges for 3D object detection. Table 1 highlights the substantial decrease in both precision and recall for Voxel-RCNN (Deng et al. 2021) on the WOD dataset during rainy conditions. This effect is further illustrated in Fig. 5, which depicts the impact of rainy weather on the detector's performance: 1) *Dense rain noise.* The presence of dense rain noise often results in numerous false positives (Fig. 5①). Additionally, these noise induced detections can erod the confidence of predictions, potentially leading to false negatives (Fig.5②). 2) *Missing points.* The sparse clustered background points due to rain induce false positives (Fig.5③). Meanwhile, the absence of foreground points creates sparse point representations for many objects, making them difficult to recall accurately(Fig.5④).

Thus, instead of naively utilizing simulated data as augmentation, it is crucial to propose a novel framework to ad-

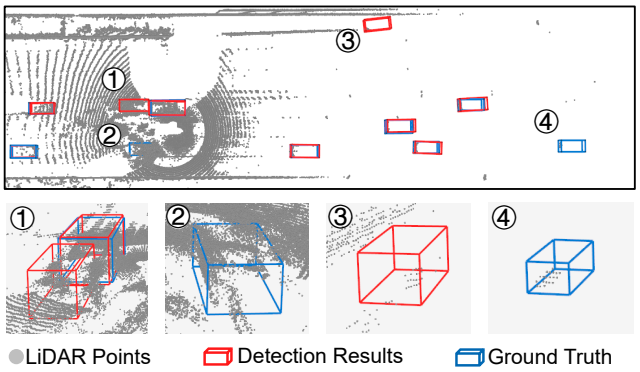

Figure 5: Examples of rainy condition effects on 3D object detection. The blue and red boxes represent groundtruths and predictions.

dress the domain gap between rainy and sunny conditions.

AWID. Distilling sunny instances into rainy instances can help students efficiently extract features from indistinguishable and sparse objects. However, directly distilling instance features from sunny teachers to rainy students, following (Zheng et al. 2022a,b), poses a great challenge due to the gap between sunny and rainy object variations. By comparing and analyzing the data of instances under rainy and sunny conditions, we found that the changes in instances are mainly reflected in differences in density and shape. To overcome this challenge, our proposed AWID utilizes the similarity between corresponding sunny and rainy objects as distillation weights for instance features. Our adaptive weighting reduces the domain gap caused by variations in density and shape under different weather conditions. The similarity weight consists of two components: density similarity and shape similarity. For a groundtruth object box B_i in sunny and rainy conditions, with point clouds PC_i^s and PC_i^r of d_i^s and d_i^r points, the density similarity $\hat{\mathcal{S}}_i$ is:

$$\hat{\mathcal{S}}_i = tanh\left(\frac{min\left(d_i^s, d_i^r\right)}{|d_i^s - d_i^r| + \epsilon}\right), \tag{3}$$

ϵ = 1e-6 prevents division by zero. Then calculate the shape similarity $\tilde{\mathcal{S}}_i$ by Chamfer Distance d_{CD} (Borgefors 1986):

$$\tilde{\mathcal{S}}_i = 1 - tanh\left(d_{CD}\left(PC_i^s, PC_i^r\right)\right). \tag{4}$$

Combining Eqs.3 and Eqs.4 we obtain the similarity $\mathcal{S}_i$:

$$\mathcal{S}_i = \hat{\mathcal{S}}_i \cdot \tilde{\mathcal{S}}_i. \tag{5}$$

So we can transfer instance feature knowledge for each box from sunny to rainy by using similarity $\mathcal{S}$ as weight:

$$\mathcal{L}_{ins} = \frac{1}{|B|}\sum_{i \in B} \mathcal{S}_i \cdot \mathcal{L}_{sml1}\left(I_i^s, I_i^r\right), \tag{6}$$

where B is the groundtruth boxes; I_i is the i-th instance feature, superscript s, r represent sunny and rainy.

PRD. Our AWID greatly aids in sunny-to-rainy knowledge distillation. However, it remains specific to the RoI head in two-stage works and lacks comprehensiveness. And we expect the model to predict consistent results for the same scene on sunny and rainy weather.

To address the remaining challenges, we introduce response distillation as a classical distillation approach (Yang et al. 2022). This approach promotes prediction consistency between sunny teachers and rainy students. We focus on the teacher's output in positions where the confidence exceeds the threshold $\mathcal{T}$ (0.5), aiming to prevent negative effects from already difficult sunny objects. Specifically, at estimated foreground positions G:

$$\mathcal{L}_{cls}^r = \frac{1}{|G|}\sum_{i \in G} \mathbb{I}\left(\phi\left(c_i^s\right) \geq \mathcal{T}\right) \cdot \mathcal{L}_{mse}\left(c_i^r, c_i^s\right), \tag{7}$$

$$\text{and} \quad \mathcal{L}_{reg}^r = \mathcal{L}_{reg}^{3D}\left(b^r, b^s\right), \tag{8}$$

where the superscripts r and s indicate rainy student and sunny teacher, respectively; c and b represent classification and box regression predictions; ϕ is the sigmoid function; $\mathbb{I}$ filters the teacher's classification output, focusing on high-confidence predictions; $\mathcal{L}_{cls}^r$ and $\mathcal{L}_{reg}^r$ are the classification and regression response distillation losses. By combining Eqs.7 and Eqs.8, the total precise response distillation loss is obtained as:

$$\mathcal{L}_{rsp} = \lambda_1 \mathcal{L}_{cls}^r + \lambda_2 \mathcal{L}_{reg}^r. \tag{9}$$

where set λ_1 = 15, λ_2 = 0.2 according to (Yang et al. 2022).

NAPC. The previous modules improve model robustness in rainy conditions via sunny-to-rainy distillation. However, this implicit processing alone cannot mitigate false positives caused by dense rain noise effectively. The rain noise caused by splashing caused by the high speed movement of vehicle tires is heavily distributed around the instance. We leverage the key advantage of simulated rainy data, which provides self-contained rain noise labels for correcting prediction inaccuracies. We introduce a Noise-Aware Prediction Correction head, NAPC, which focuses on suppressing rain noise. For a predicted box $\mathcal{B}_i$ with $\mathcal{K}$ points, including $\hat{\mathcal{K}}$ noise points and $\tilde{\mathcal{K}}$ non-noise points ($\hat{\mathcal{K}} + \tilde{\mathcal{K}} = \mathcal{K}$), the noise ratio $\mathcal{R}_i$ is calculated as:

$$\mathcal{R}_i = \frac{\hat{\mathcal{K}}}{\tilde{\mathcal{K}} + \epsilon}, \tag{10}$$

where ϵ = 1e-6 is a small constant to prevent division by zero. The noise-aware prediction correction loss $\mathcal{L}_{napc}$ is calculated as:

$$\mathcal{L}_{napc} = \frac{1}{|\mathcal{B}|}\sum_{i \in \mathcal{B}} tanh(\mathcal{R}_i) \cdot \mathcal{C}_i, \tag{11}$$

where $\mathcal{B}$ represents prediction boxes, and $\mathcal{C}_i$ is the confidence of the i-th prediction box.

Overall Loss Function. We train the rainy student while keeping the pre-trained sunny teacher fixed (see Fig.3) with the following supervision losses and distillation losses:

$$\mathcal{L} = \mathcal{L}_{cls} + \mathcal{L}_{reg} + \eta_1 \mathcal{L}_{ins} + \eta_2 \mathcal{L}_{rsp} + \eta_3 \mathcal{L}_{napc} \tag{12}$$

where $\mathcal{L}_{cls}$ is classification loss and $\mathcal{L}_{reg}$ is regression loss; η_1, η_2 and η_3 are hyper-parameters.

Methods	All(L1) mAP/mAPH	All(L2) mAP/mAPH	Vehicle(L1) mAP/mAPH	Vehicle(L2) mAP/mAPH	Ped.(L1) mAP/mAPH	Ped.(L2) mAP/mAPH
Voxel-RCNN (Deng et al. 2021)	65.3/62.6	55.8/53.7	73.2/72.8	66.6/66.2	57.4/52.4	45.1/41.1
+ Our DRET-Aug&SRKD	**69.5/66.2**	**60.1/57.3**	**75.0/74.6**	**68.5/68.1**	**63.9/57.8**	**51.6/46.4**
Improvement	*+4.2/+3.6*	*+4.3/+3.6*	*+1.8/+1.8*	*+1.9/+1.9*	*+6.5/+5.4*	*+6.5/+5.3*
PV-RCNN++ (Shi et al. 2022)	66.2/63.4	56.8/54.5	73.7/73.2	67.1/66.7	58.6/53.5	46.5/42.2
+ Our DRET-Aug&SRKD	**69.2/66.5**	**59.8/57.6**	**76.7/76.2**	**69.4/69.0**	**61.6/56.7**	**50.1/46.2**
Improvement	*+3.0/+3.1*	*+3.0/+3.2*	*+3.0/+3.0*	*+2.3/+2.3*	*+3.0/+3.2*	*+3.6/+4.0*
DSVT (Wang et al. 2023)	69.3/66.5	60.4/57.9	75.1/74.7	68.4/68.0	63.4/58.3	52.3/47.8
+ Our DRET-Aug&SRKD	**72.7/69.7**	**63.6/61.1**	**76.8/76.4**	**70.2/69.8**	**68.5/63.0**	**57.0/52.3**
Improvement	*+3.4/+3.2*	*+3.2/+3.2*	*+1.7/+1.7*	*+1.8/+1.8*	*+5.1/+4.7*	*+4.7/+4.5*

Table 2: The rainy testing results of 3D detectors within our proposed DRET-Aug and SRKD on the WOD-DA (100% training data). All the results are based on our implementation following the open-source code. The best results are in bold.

Methods	All(L1) mAP/mAPH	All(L2) mAP/mAPH	Vehicle(L1) mAP/mAPH	Vehicle(L2) mAP/mAPH	Ped.(L1) mAP/mAPH	Ped.(L2) mAP/mAPH	Cyc.(L1) mAP/mAPH	Cyc.(L2) mAP/mAPH
Voxel-RCNN	77.3/74.9	70.6/68.4	78.4/77.9	69.8/69.4	80.5/74.7	71.7/66.4	73.1/72.0	70.4/69.3
+ Ours	**77.6/75.2**	**71.0/68.6**	**78.5/78.0**	**70.0/69.5**	**81.0/75.1**	**72.2/66.7**	**73.4/72.4**	**70.7/69.7**
PV-RCNN++	77.6/75.0	70.8/68.6	79.1/78.6	70.3/69.9	81.1/75.0	72.5/66.8	72.5/71.4	69.7/69.2
+ Ours	**78.1/75.6**	**71.4/69.0**	**79.3/78.8**	**70.6/70.1**	**81.2/75.2**	**72.6/66.9**	**73.7/72.7**	**70.9/69.9**
DSVT	79.5/76.9	73.4/70.9	79.1/78.7	71.3/70.8	82.5/76.2	75.2/69.2	**76.9/75.8**	73.8/72.8
+ Ours	**79.6/77.1**	**73.7/71.2**	**79.5/79.0**	**71.6/71.2**	**82.6/76.5**	**75.3/69.5**	76.8/75.7	**74.1/73.0**

Table 3: The sunny testing results of 3D detectors within our proposed DRET-Aug and SRKD on the WOD-P (100% training data). All the results are based on our implementation following the open-source code. The best results are in bold.

Experiments

We conducted evaluations of our proposed DRET and SRKD on the Waymo Open Dataset (WOD) (Sun et al. 2020). To demonstrate the effectiveness of 3D detection under rainy condition, we compared the performance of state-of-the-art model DSVT (Wang et al. 2023) and PV-RCNN++ (Shi et al. 2022), Voxel-RCNN (Deng et al. 2021).

Dataset. The WOD comprises two subsets: Perception (WOD-P) and Domain Adaptation (WOD-DA). The WOD-P contains ∼158k training frames and ∼40k validation frames, predominantly sunny conditions (99.4%). We directly use the WOD-P validation for sunny testing. the WOD-DA includes various weather conditions such as foggy, cloudy, and rainy. For our rainy testing, we selected ∼3k rainy frames for rainy testing. The training process of all models used only the training set of WOD-P.

Metrics. The official metrics of WOD are mean Average Precision (mAP(L1), mAP(L2)) with Heading (mAPH(L1), mAPH(L2)), where L1 and L2 denote the difficulty level. Additionally, for the WOD-DA, only vehicle and pedestrian classes have labeled annotations for evaluation.

Implementation and training details. We trained models using two LiDAR returns and a single frame. The specific implementation and training details for each baseline are provided in the supplementary material. In the DRET, we used atmospheric parameters following (Hahner et al. 2021) and (Kilic et al. 2021). In the SRKD framework, we set the hyper-parameters η_1, η_2, and η_3 (eq.12) as 2.0, 0.5, and 2.0.

Main Experiment Results

We conducted three model comparison groups. The models were trained on the WOD-P training set based on DRET augmentation (DRET-Aug) and evaluated on the WOD-P validation set and the rainy data of the WOD-DA.

Performance comparison under rainy weather. We first evaluated rainy performance on the WOD-DA. Table 2 shows our proposed DRET-Aug and SRKD significantly improve the performance of all 3D detectors in rain. With the SRKD training using DRET-Aug augmentation, Voxel-RCNN, PV-RCNN++, and DSVT achieved improvements of 4.3%, 3.0%, and 3.2% on All(L2-mAP), respectively. Notably, all detectors show substantial gains in the pedestrian class, particularly Voxel-RCNN with a remarkable improvement of 6.5% on Ped.(L2-mAP). This demonstrates the capability of our framework to improve challenging small object detection under rainy conditions. The results demonstrate the effectiveness and broad applicability of our design.

Performance comparison under sunny weather. In contrast to existing methods (Bijelic et al. 2020; Lin et al. 2022) that only focus on performance under adverse weather, it is also essential to maintain performance in sunny weather. Table 3 demonstrates that our framework also slightly improves the performance of 3D detectors in sunny conditions. The performance gains mainly come from the enhanced robustness to handle sparse and indistinguishable objects.

Rain simulation methods. In Table 4, we further compared our DRET with the leading rain simulation methods, LISA-Aug and SPRAY-Aug. We observed that plain

Methods	All(L1) mAP/mAPH	All(L2) mAP/mAPH
Baseline	65.3/62.6	55.8/53.7
+ LISA-Aug (Kilic et al. 2021)	65.5/63.2	56.1/54.3
+ SPRAY-Aug (Shih et al. 2022)	**66.0**/63.0	56.0/53.9
+ DRET-Aug (Ours)	65.9/**63.4**	**56.6/54.7**

Table 4: Comparison with state-of-the-art rain simulation methods on the WOD-DA.(100% training data).

Methods	All(L1) mAP/mAPH	All(L2) mAP/mAPH
Baseline	65.3/62.6	55.8/53.7
+ De-Noising (Heinzler et al. 2020)	66.9/64.1	56.1/55.4
+ SPG (Xu et al. 2021)	67.5/64.7	58.0/55.9
+ SRKD (Ours)	**69.5/66.2**	**60.1/57.3**

Table 5: Comparison with state-of-the-art 3D object detection in rain methods on the WOD-DA. (100% training data).

Methods	All(L1) mAP/mAPH	All(L2) mAP/mAPH
Voxel-RCNN (Deng et al. 2021)	65.3/62.6	55.8/53.7
+ Our DRET-Aug	**65.9/63.4**	**56.6/54.7**
PV-RCNN++ (Shi et al. 2022)	66.2/63.4	56.8/54.5
+ Our DRET-Aug	**67.0/64.1**	**57.8/55.4**
DSVT (Wang et al. 2023)	69.3/66.5	60.4/57.9
+ Our DRET-Aug	**69.9/67.3**	**61.0/58.7**

Table 6: The rainy testing results of 3D detectors with our proposed DRET-Aug on the WOD-DA (100% training data).

rain data augmentation alone does not provide significant improvement, as mentioned earlier. Although SPRAY-Aug achieves the best performance on All(L1-mAP), slightly outperforming our DRET-Aug. However, combining all metrics, our method still performs better, especially at the more comprehensive L2 difficulty. Additionally, our DRET-Aug achieves an improvement of 0.5% and 0.6% on All(L2-mAP) compared to LISA-Aug and SPRAY-Aug. These results again prove DRET's realism by analyzing the rainy performance gained from simulated data.

Robust 3D object detection methods in rain. Given the limited prior work on 3D object detection under rainy conditions, we compared our method to the de-noising (Heinzler et al. 2020) and SPG (Xu et al. 2021). As shown in Table 5, our SRKD has improved by 4.0% and 2.1% All(L2-mAP) respectively compared to De-Noising and SPG.

Ablation Study

Data augmentation based on DRET-Aug. To further explore the improvements achieved solely by our DRET-Aug data augmentation, we trained the above baseline 3D detectors with DRET-Aug alone. As shown in Table 6, DRET-Aug improves Voxel-RCNN, PV-RCNN++ and DSVT by 0.8%, 1.0% and 0.6% on All(L2-mAP), respectively.

Component of SRKD analysis. Furthermore, to validate the efficacy of each component in our proposed SRKD, we conducted an ablation study using 20% training data. The baseline model used is Voxel-RCNN (Deng et al. 2021) trained with our DRET-Aug. As shown in Table 7, the inclusion of our proposed components AWID, PRD, and NAPC significantly improves the performance of the Voxel-RCNN baseline. Specifically, AWID, PRD, and NAPC yield gains of 2.0%, 0.9%, and 0.7% on All(L1-mAP), respectively. This highlights the effectiveness of our SRKD components.

Method	AWID	PRD	NAPC	All(L1) mAP/mAPH
Voxel-RCNN				63.3/60.7
	✓			65.3/62.7
	✓	✓		66.2/63.7
	✓	✓	✓	**66.9/64.2**

Table 7: Effects of the different components of SRKD.

Method	$\hat{\mathcal{S}}$	$\tilde{\mathcal{S}}$	All(L1) mAP/mAPH
Voxel-RCNN + AWID			62.8/60.2
	✓		64.0/61.4
		✓	64.8/62.3
	✓	✓	**65.3/62.7**

Table 8: Effects of the different weighting strategy.

Similarity weighting strategies. Finally, we tested the different weighting strategies in our AWID module using shape ($\hat{\mathcal{S}}$) and density ($\tilde{\mathcal{S}}$) similarity. As shown in Table 8, directly distilling rainy-sunny instance features without any weighting performs poorly. This confirms that it is not reasonable to distill all instances equally. Moreover, using both $\hat{\mathcal{S}}$ and $\tilde{\mathcal{S}}$ as weighting strategies yields the best performance, improving upon the no weighting strategy by 2.5% on All (L1-mAP). This validates the importance of our designed adaptive weighting for distilling instance features.

Conclusion

We propose a novel design including DRET and SRKD for 3D detection in sunny and rainy conditions. It uses DRET, a realistic rain simulation, to generate rainy data and mitigate data scarcity. Our SRKD framework then transfers sunny knowledge to rainy detectors via sunny-to-rainy knowledge distillation. Sufficient experimental results of the WOD show that our design can enhance the robustness of 3D detectors in rain. Moreover, our design can even slightly improve the sunny performance of 3D detectors.

Limitations. Our two-stage DRET requires preprocessing to generate particle sets and cannot be end-to-end. While inference efficiency is unchanged, training time increases unavoidably mainly due to shape similarity calculation.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (No.62171393), and the Fundamental Research Funds for the Central Universities (No.20720220064).

References

Bijelic, M.; Gruber, T.; Mannan, F.; Kraus, F.; Ritter, W.; Dietmayer, K.; and Heide, F. 2020. Seeing Through Fog Without Seeing Fog: Deep Multimodal Sensor Fusion in Unseen Adverse Weather. In *CVPR*.

Borgefors, G. 1986. Distance transformations in digital images. *Computer vision, graphics, and image processing*, 34(3): 344–371.

Bridson, R.; Houriham, J.; and Nordenstam, M. 2007. Curl-Noise for Procedural Fluid Flow. *ACM Trans. Graph.*, 26(3): 46–es.

Charron, N.; Phillips, S.; and Waslander, S. L. 2018. De-Noising of Lidar Point Clouds Corrupted by Snowfall. In *CRV*.

Deng, J.; Shi, S.; Li, P.; Zhou, W.; Zhang, Y.; and Li, H. 2021. Voxel R-CNN: Towards High Performance Voxel-based 3D Object Detection. *AAAI*, 35.

Do, A. T.; and Yoo, M. 2022. LossDistillNet: 3D Object Detection in Point Cloud Under Harsh Weather Conditions. *IEEE Access*, 10: 84882–84893.

Hahner, M.; Sakaridis, C.; Bijelic, M.; Heide, F.; Yu, F.; Dai, D.; and Van Gool, L. 2022. LiDAR Snowfall Simulation for Robust 3D Object Detection. In *CVPR*.

Hahner, M.; Sakaridis, C.; Dai, D.; and Van Gool, L. 2021. Fog Simulation on Real LiDAR Point Clouds for 3D Object Detection in Adverse Weather. In *ICCV*.

He, C.; Zeng, H.; Huang, J.; Hua, X.-S.; and Zhang, L. 2020. Structure Aware Single-Stage 3D Object Detection From Point Cloud. In *CVPR*.

Heinzler, R.; Piewak, F.; Schindler, P.; and Stork, W. 2020. CNN-Based Lidar Point Cloud De-Noising in Adverse Weather. *IEEE Robotics and Automation Letters*, 5.

Kilic, V.; Hegde, D.; Sindagi, V. A.; Cooper, A.; Foster, M.; and Patel, V. M. 2021. Lidar Light Scattering Augmentation (LISA): Physics-based Simulation of Adverse Weather Conditions for 3D Object Detection. *ArXiv*.

Lin, J.; Yin, H.; Yan, J.; Ge, W.; Zhang, H.; and Rigoll, G. 2022. Improved 3D Object Detector Under Snowfall Weather Condition Based on LiDAR Point Cloud. *IEEE Sensors Journal*, 22.

Sheng, H.; Cai, S.; Liu, Y.; Deng, B.; Huang, J.; Hua, X.-S.; and Zhao, M.-J. 2021. Improving 3D Object Detection With Channel-Wise Transformer. In *ICCV*.

Shi, S.; Guo, C.; Jiang, L.; Wang, Z.; Shi, J.; Wang, X.; and Li, H. 2020. PV-RCNN: Point-Voxel Feature Set Abstraction for 3D Object Detection. In *CVPR*.

Shi, S.; Jiang, L.; Deng, J.; Wang, Z.; Guo, C.; Shi, J.; Wang, X.; and Li, H. 2022. PV-RCNN++: Point-Voxel Feature Set Abstraction With Local Vector Representation for 3D Object Detection. *Int. J. Comput. Vision*, 131.

Shi, S.; Wang, X.; and Li, H. 2019. PointRCNN: 3D Object Proposal Generation and Detection From Point Cloud. In *CVPR*.

Shih, Y.-C.; Liao, W.-H.; Lin, W.-C.; Wong, S.-K.; and Wang, C.-C. 2022. Reconstruction and Synthesis of Lidar Point Clouds of Spray. *IEEE Robotics and Automation Letters*, 7.

Sun, P.; Kretzschmar, H.; Dotiwalla, X.; Chouard, A.; Patnaik, V.; Tsui, P.; Guo, J.; Zhou, Y.; Chai, Y.; Caine, B.; Vasudevan, V.; Han, W.; Ngiam, J.; Zhao, H.; Timofeev, A.; Ettinger, S.; Krivokon, M.; Gao, A.; Joshi, A.; Zhang, Y.; Shlens, J.; Chen, Z.; and Anguelov, D. 2020. Scalability in Perception for Autonomous Driving: Waymo Open Dataset. In *CVPR*.

Teufel, S.; Volk, G.; Von Bernuth, A.; and Bringmann, O. 2022. Simulating Realistic Rain, Snow, and Fog Variations For Comprehensive Performance Characterization of LiDAR Perception. In *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*.

Wang, H.; Shi, C.; Shi, S.; Lei, M.; Wang, S.; He, D.; Schiele, B.; and Wang, L. 2023. DSVT: Dynamic Sparse Voxel Transformer With Rotated Sets. In *CVPR*.

Wu, H.; Wen, C.; Shi, S.; Li, X.; and Wang, C. 2023. Virtual Sparse Convolution for Multimodal 3D Object Detection. In *CVPR*.

Wu, X.; Peng, L.; Yang, H.; Xie, L.; Huang, C.; Deng, C.; Liu, H.; and Cai, D. 2022. Sparse Fuse Dense: Towards High Quality 3D Detection with Depth Completion. In *CVPR*.

Xu, Q.; Zhou, Y.; Wang, W.; Qi, C. R.; and Anguelov, D. 2021. SPG: Unsupervised Domain Adaptation for 3D Object Detection via Semantic Point Generation. In *ICCV*.

Yan, Y.; Mao, Y.; and Li, B. 2018. SECOND: Sparsely Embedded Convolutional Detection. *Sensors*, 18.

Yang, J.; Shi, S.; Ding, R.; Wang, Z.; and Qi, X. 2022. Towards Efficient 3D Object Detection with Knowledge Distillation. *NIPS*, 35.

Yang, Z.; Sun, Y.; Liu, S.; Shen, X.; and Jia, J. 2019. STD: Sparse-to-Dense 3D Object Detector for Point Cloud. In *ICCV*.

Zheng, W.; Hong, M.; Jiang, L.; and Fu, C.-W. 2022a. Boosting 3D Object Detection by Simulating Multimodality on Point Clouds. In *CVPR*.

Zheng, W.; Jiang, L.; Lu, F.; Ye, Y.; and Fu, C.-W. 2022b. Boosting Single-Frame 3D Object Detection by Simulating Multi-Frame Point Clouds. In *ACMMM*.

Zheng, W.; Tang, W.; Jiang, L.; and Fu, C.-W. 2021. SE-SSD: Self-Ensembling Single-Stage Object Detector From Point Cloud. In *CVPR*.