

Open-Weather Robust 3D Detection via Dual-Critic Diffusion Alignment

Shuyao Li¹, Chuanxing Geng^{1,2}*, Heyang Sun¹, Qiang Zhou¹, and Jingjing Gu¹

¹ Nanjing University of Aeronautics and Astronautics, Nanjing, China
{lsyao, gengchuanxing, zhouqnuaacs, gujingjing}@nuaa.edu.cn

² MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing, China

Abstract. Robust 3D object detection under adverse weather remains a critical hurdle for autonomous driving. Despite progress with LiDAR–4D radar fusion, most methods are constrained by a closed-world assumption, implicitly requiring training and test weather to align in both type and severity. This premise fails in practice: the open-ended nature of weather, and even variations within a single type like rain, cause dramatically different LiDAR degradation patterns, leading to significant performance drops in unseen conditions. To address this, we present Dual-Critic Guided Diffusion Alignment (DCDA), a *weather-agnostic* framework that learns to recover degraded LiDAR features toward a clean manifold. Rather than modeling specific weather types, DCDA employs a 4D radar-conditioned diffusion process to progressively refine features, guided by two complementary critics. (i) A detection-guided critic, anchored by a pre-trained clean-weather model, ensures that the refined features retain object-level discriminability and localization accuracy. (ii) A weather adversarial critic enforces holistic distributional consistency with clean-weather representations. By aligning features through semantic and distributional constraints rather than explicit weather modeling, DCDA generalizes effectively to unseen weather types and severities without requiring paired data or weather labels. We further introduce a structured open-weather benchmark with held-out type–severity combinations and extensive experiments verify DCDA’s advantages. Code is available at <https://github.com/Mangonn/DCDA>.

Keywords: 3D Object Detection · Adverse Weather Robustness · Unseen Weather Generalization · LiDAR–4D Radar Fusion

1 Introduction

Robust 3D object detection is fundamental to autonomous driving, where perception errors directly impact planning and safety-critical decisions [17]. While modern LiDAR-based detectors achieve strong performance in clear conditions,

* Corresponding author.

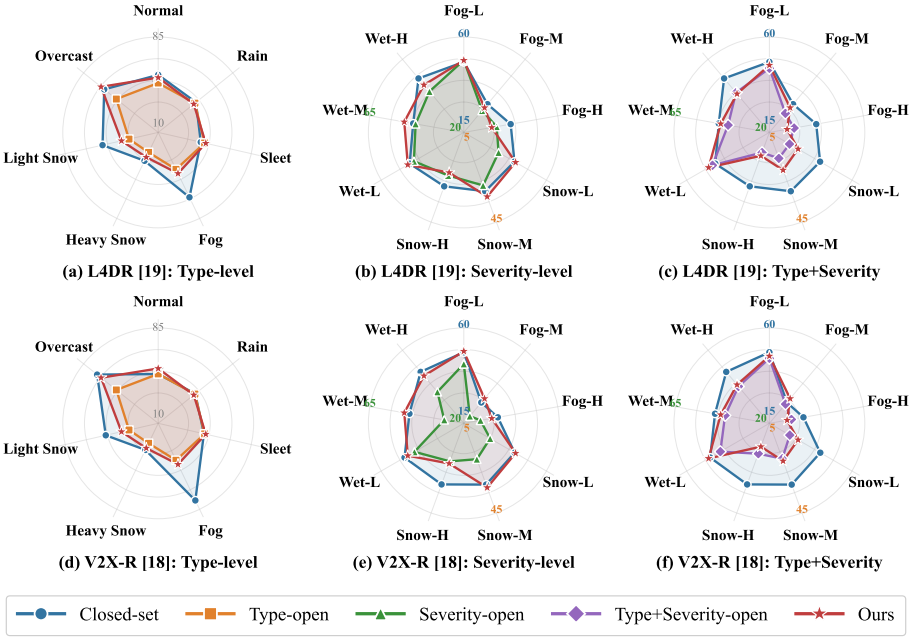


Fig. 1: Open-weather generalization gap on K-Radar. The top and bottom rows show BEV AP for L4DR and V2X-R, respectively, under the three protocols shown from left to right: *type-open*, *severity-open*, and *type+severity-open*. *Closed-set* denotes each detector’s conventional closed-set baseline, where the test weather conditions are seen during training. **Ours** denotes the proposed DCDA model under the corresponding open-weather protocol. Type-open holds out real weather types, severity-open holds out synthetic severity levels, and type+severity-open holds out both synthetic weather type and severity. The drop from *Closed-set* to the open-weather curves reveals a clear generalization gap for the two representative baselines.

their reliability degrades sharply under adverse weather such as fog, rain, and snow [2, 12, 13]. Weather-induced degradation is inherently asymmetric across sensing modalities. Fog and snow introduce severe backscatter and range attenuation in LiDAR, corrupting geometric structure [2]. In contrast, millimeter-wave radar remains comparatively robust and preserves long-range and Doppler cues [4, 10, 19]. Motivated by this complementarity, recent work increasingly adopts LiDAR–4D radar fusion [19, 25].

However, despite their improved robustness, most existing fusion methods [4, 18, 19, 33, 35] implicitly rely on a closed-world assumption: the weather conditions encountered at test time are assumed to be represented during training or restricted to predefined corruption settings [11, 20]. This assumption rarely holds in real deployment. Weather is open-ended and continuously varying in both type and severity. Even within a single category such as snow or fog, intensity variations can induce substantially different LiDAR degradation patterns [12, 13, 37].

As a result, models trained under limited weather configurations may fail when exposed to unseen type–severity combinations.

To examine how current state-of-the-art fusion detectors behave under such open-weather conditions, we construct a structured benchmark that jointly considers variations in weather type and severity. Based on K-Radar [25], we design systematic holdout protocols over (i) weather types, (ii) severity levels, and (iii) their combinations, simulating unseen environmental configurations at test time. Since K-Radar provides limited severity annotations, we further construct controlled intensity levels (Light/Medium/Heavy) by progressively synthesizing weather effects on *Normal* samples [20, 37], enabling systematic robustness evaluation across severity variations.

As shown in Fig. 1, our findings are striking: Both representative state-of-the-art baselines, L4DR [19] and V2X-R [18] exhibit consistent and often substantial performance degradation under these open settings compared to conventional closed-world evaluations. This reveals a clear open-weather generalization gap that has not been adequately addressed. These observations suggest that robustness cannot rely solely on enumerating weather types or simulating predefined degradations. Instead, models must learn weather-agnostic mechanisms capable of recovering stable geometric representations beyond seen conditions.

To this end, we propose Dual-Critic Guided Diffusion Alignment (DCDA), a weather-agnostic framework that restores degraded LiDAR features by aligning them to a clean (weather) manifold. Rather than explicitly modeling specific weather patterns, DCDA formulates feature recovery as a 4D radar-conditioned diffusion process, progressively refining corrupted LiDAR features [7, 18, 40, 44]. A central challenge is to enhance degraded features while mitigating negative transfer from imperfect radar cues and preserving detection semantics. We address this using two complementary critics that regulate the diffusion trajectory:

- ✓ **A detection-guided critic**, instantiated by a frozen *Normal* detector to enforce task-level semantic consistency and localization fidelity;
- ✓ **A weather adversarial critic**, which enforces global distributional alignment with *Normal* representations.

By jointly constraining task-space and representation-space alignment—without explicit weather modeling, paired clean–degraded supervision, or conditioning on specific weather type/severity labels—DCDA generalizes effectively to unseen weather types and severity levels. As shown in Fig. 1, rather than over-optimizing for specific conditions, our DCDA achieves stronger overall open-weather generalization than both representative baselines on K-Radar. Averaged over the plotted weather, DCDA improves mean 3D AP over the better baseline by 5.2 percentage points under held-out weather types (type-open), 3.0 percentage points under held-out severity levels (severity-open), and 1.8 percentage points under their combinations (type+severity-open).

Our contributions are threefold:

- We identify and formalize the open-weather generalization problem in LiDAR–4D radar fusion, and construct a structured benchmark that jointly evaluates

weather type and severity variations, revealing a substantial robustness gap in existing state-of-the-art detectors.

- We propose a weather-agnostic diffusion alignment paradigm (DCDA) that formulates degraded LiDAR feature recovery as a radar-conditioned manifold alignment process, avoiding explicit weather modeling, paired clean–degraded supervision, and weather type/severity labels as conditional inputs.
- We introduce a dual-critic regularization mechanism that stabilizes diffusion-based alignment by jointly enforcing task-space semantic consistency and representation-space distribution matching, enabling robust generalization to unseen weather types and severity levels.

2 Related Work

2.1 LiDAR-Based Robust 3D Object Detection in Adverse Weather

Adverse weather (e.g., fog, rain, and snow) introduces modality-specific degradations (e.g., attenuation, backscatter clutter, and point sparsification) that corrupt geometric evidence and substantially degrade 3D detection performance [2, 11–13]. Prior efforts improve robustness through three directions: (i) LiDAR restoration and filtering, including deweathering/denoising and clutter suppression (e.g., snow/rain outlier filters and adverse-effect detection) [16, 31]; (ii) robust learning, which enhances adverse-weather generalization via robust objectives (e.g., cross-weather distillation or geometry/structure constraints) [32, 38]; and (iii) adaptation and post-processing, which mitigates test-time shifts via deployment-time adaptation or calibrated post-processing [8, 21, 41]. When complementary camera cues are available, degradation-aware camera-LiDAR fusion exploits asymmetric corruption via uncertainty-guided feature exchange and fog-aware depth completion [2, 42, 43]. Context-gated fusion further improves robustness across day/night and clear/rainy conditions [34]. However, robustness often remains limited once LiDAR geometry is heavily corrupted, motivating the use of weather-tolerant millimeter-wave radar signals.

2.2 LiDAR–4D Radar Fusion for Robust 3D Detection in Adverse Weather

Millimeter-wave radar is generally less affected by scattering media than LiDAR [10, 14, 28], motivating LiDAR–4D radar fusion for adverse-weather 3D detection [4, 18, 19, 36]. K-Radar provides a large-scale benchmark with synchronized LiDAR and 4D radar across diverse real-world weather, enabling systematic study of such fusion [25]. Recent methods improve robustness via (i) 3D-centric fusion with weather-aware modulation (e.g., 3D-LRF with weather-conditional gating) [4], and (ii) denoising and gated fusion to handle sensor gaps and varying degradation (L4DR) [19]. Beyond fusion operator design, radar has also been used as auxiliary supervision to improve LiDAR-only robustness [3]. Other LiDAR–4D radar fusion lines explore sensor-adaptive multi-modal fusion [26] and interaction-based attention with adaptive weighting, leveraging

radar-specific cues (e.g., velocity/RCS), robustness to sensor failures, and multi-frame radar accumulation with motion-aware compensation [5, 27, 29, 35, 39]. Diffusion-based denoising further leverages radar as a condition to suppress weather noise in LiDAR features [18]. Despite these advances, many existing evaluations follow standard splits where weather categories overlap between training and test, leaving explicit generalization to held-out weather types/severity less explored.

2.3 Out-of-Distribution (OOD) Weather Generalization

A limited but growing body of work explicitly targets open-weather generalization and evaluates robustness under controlled OOD corruptions and severity levels [11, 20]. Representative approaches leverage uncertainty-aware mechanisms to handle asymmetric degradations (e.g., SeeingThroughFog) [2]. Diffusion-based weather adaptation has also been explored for robust 3D detection; for example, MonoWAD enhances features with a weather-adaptive diffusion model guided by weather-reference representations learned from clear/fog counterparts [23]. Recent studies further investigate diffusion for adverse-weather restoration and augmentation beyond monocular settings [15, 22]. In parallel, robustness under domain shift has been advanced via unsupervised domain adaptation [6, 30] and test-time adaptation [8, 9, 41]. Simulation-free robustness training has also been explored to improve weather-based OOD generalization without relying on costly simulators [1].

Despite emerging OOD evaluations and adaptation/restoration techniques, two limitations remain for practical LiDAR–4D radar fusion under open-weather deployment. First, several robustness pipelines still rely on explicit weather priors (e.g., domain identifiers or predefined corruption families) and/or additional signals such as weather annotations or paired clear-adverse references, which are costly to curate and not consistently available across diverse operating conditions. Second, robustness is commonly characterized under either type shift or severity variation in isolation, while generalization under their joint type–severity shifts is less systematically examined in LiDAR–4D radar fusion settings. These limitations motivate robustness mechanisms and evaluation protocols that operate under minimal supervision and explicitly account for joint shifts.

3 Method

We study open-weather generalization for LiDAR–4D radar 3D object detection, where test-time conditions may deviate from training in both weather *type* and *severity*. To sidestep explicit weather modeling and reliance on paired clean-degraded supervision, we pursue weather-agnostic feature recovery, restoring stable LiDAR geometric representations beyond seen conditions. Specifically, we propose **Dual-Critic Guided Diffusion Alignment (DCDA)**, which refines the LiDAR BEV feature $\mathbf{F}_L$ into $\tilde{\mathbf{F}}_L$ conditioned on the 4D radar feature $\mathbf{F}_R$ as

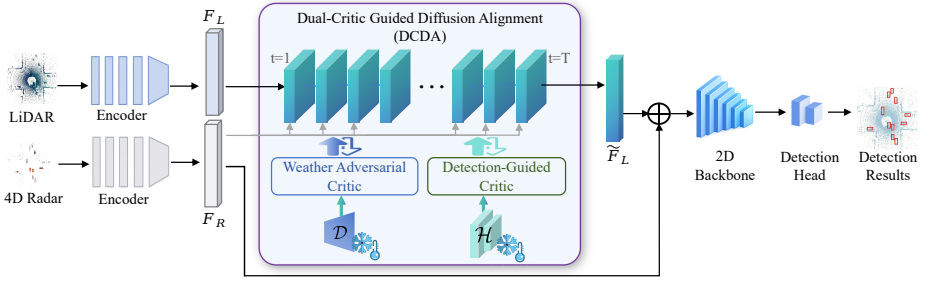


Fig. 2: Overview of DCDA. The radar-conditioned diffusion alignment refines LiDAR BEV features, guided by a frozen detection critic and a frozen normal-weather discriminator.

a comparatively stable signal. DCDA consists of (i) a **radar-conditioned diffusion alignment** $\mathcal{A}_\theta$ for iterative refinement, and (ii) a **dual-critic guidance** toward the clean manifold during training: a detection-guided critic $\mathcal{H}$ that enforces semantic and localization fidelity, and a weather adversarial critic $\mathcal{D}$ that promotes alignment to *Normal*-weather representations. At inference, $\mathcal{D}$ can be optionally reused as a lightweight router to bypass refinement for confidently *Normal* inputs. An overview is shown in Fig. 2.

3.1 Radar-Conditioned Diffusion Alignment

We model feature refinement as a radar-conditioned diffusion process operating on LiDAR BEV features. Given the input LiDAR feature $\mathbf{F}_0 \equiv \mathbf{F}_L$ (possibly degraded), the forward diffusion process progressively corrupts it with Gaussian noise over T steps:

$$\mathbf{F}_t = \sqrt{\bar{\alpha}_t} \mathbf{F}_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \quad (1)$$

where $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$ follows a predefined variance schedule.

Unlike generative diffusion models that start from pure Gaussian noise, we initialize the reverse chain from a noised observation $\mathbf{F}_T \sim q(\mathbf{F}_T | \mathbf{F}_0)$. This retains the coarse structure of the input and confines denoising to a local neighborhood around $\mathbf{F}_0$, serving as a conditional refiner.

At each reverse step, a radar-conditioned U-Net $\mathcal{A}_\theta$ directly predicts a *Normal*-weather feature estimate:

$$\hat{\mathbf{F}}_0^{(t)} = \mathcal{A}_\theta([\mathbf{F}_t; \mathbf{F}_R], t). \quad (2)$$

This prediction is used to compute the reverse transition via the standard DDPM posterior:

$$\mu_\theta = \frac{\sqrt{\bar{\alpha}_{t-1}} \beta_t}{1 - \bar{\alpha}_t} \hat{\mathbf{F}}_0^{(t)} + \frac{\sqrt{\bar{\alpha}_t} (1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \mathbf{F}_t, \quad (3)$$

where $\beta_t = 1 - \alpha_t$ and the posterior variance is $\tilde{\beta}_t = \frac{(1 - \tilde{\alpha}_t - 1)}{1 - \tilde{\alpha}_t} \beta_t$. After T reverse steps, we obtain the refined feature $\tilde{\mathbf{F}}_L = \hat{\mathbf{F}}_L^{(0)}$, which replaces the original $\mathbf{F}_L$ for downstream fusion.

We apply a self-reconstruction loss to maintain the geometric structure of the input features:

$$\mathcal{L}_{\text{diff}} = \left\| \tilde{\mathbf{F}}_L - \mathbf{F}_L \right\|_2^2. \quad (4)$$

It serves as a content-preserving regularizer to stabilize optimization and prevent over-correction, while the alignment toward the clean manifold is driven by the dual critics, as described in Sec. 3.2.

3.2 Dual-Critic Guidance Toward the Clean Manifold

We construct two *frozen* critics to steer diffusion alignment toward the clean manifold: (1) a **detection-guided critic** $\mathcal{H}$, trained on *Normal*-only data and then fixed to provide *Normal*-domain semantic guidance; and (2) a **weather adversarial critic** $\mathcal{D}$, trained to distinguish *Normal* vs. non-*Normal* LiDAR BEV features and then fixed to enforce distributional alignment to the *Normal*-domain.

Detection-guided critic (semantic constraint). To preserve object-level semantics, we reuse the standard detection loss computed by the frozen head $\mathcal{H}$:

$$\mathcal{L}_{\text{det}} = \mathcal{L}_{\text{cls}} + \mathcal{L}_{\text{loc}}. \quad (5)$$

Since $\mathcal{H}$ is pretrained on only *Normal* data, minimizing $\mathcal{L}_{\text{det}}$ requires $\tilde{\mathbf{F}}_L$ to remain in regions of the feature space that the *Normal* detector recognizes as object-discriminative, providing a task-driven *semantic pull* toward the clean manifold.

Weather adversarial critic (distributional constraint). We enforce distributional consistency with clean features via $\mathcal{D}(\cdot)$ on LiDAR BEV features. Let $\mathbf{F}_L^N$ and $\mathbf{F}_L^{\text{non-N}}$ denote LiDAR BEV features from the *Normal* and non-*Normal* training samples, respectively. We train the discriminator $\mathcal{D}$ to predict whether a feature comes from the *Normal* subset using the binary cross-entropy loss:

$$\mathcal{L}_{\mathcal{D}} = -\mathbb{E}_{\mathbf{F}_L^N} [\log \mathcal{D}(\mathbf{F}_L^N)] - \mathbb{E}_{\mathbf{F}_L^{\text{non-N}}} [\log (1 - \mathcal{D}(\mathbf{F}_L^{\text{non-N}}))]. \quad (6)$$

After pretraining, $\mathcal{D}$ is kept frozen and used to define an adversarial guidance term for refined features:

$$\mathcal{L}_{\text{adv}} = -\mathbb{E}_{\tilde{\mathbf{F}}_L} [\log \mathcal{D}(\tilde{\mathbf{F}}_L)]. \quad (7)$$

Critic-guided refinement. We apply both critics on the refined LiDAR feature $\tilde{\mathbf{F}}_L$ produced by the diffusion aligner $\mathcal{A}_\theta$ and form

$$\mathcal{L}_{\text{crit}} = \lambda_1 \mathcal{L}_{\text{det}} + \lambda_2 \mathcal{L}_{\text{adv}}, \quad (8)$$

whose gradients are backpropagated only to $\mathcal{A}_\theta$ while keeping $\mathcal{H}$ and $\mathcal{D}$ frozen. $\mathcal{L}_{\text{crit}}$ is combined with the diffusion objective to form the overall training loss, as described in Sec. 3.3 and formalized in Eq. (10). These *Normal*-domain critics provide complementary semantic and distributional guidance, regularizing the refinement and mitigating potential drift.

3.3 Training and Inference

Training. Optimizing the aligner $\mathcal{A}_\theta$ proceeds in two stages, while keeping both critics $\mathcal{H}$ and $\mathcal{D}$ frozen.

Stage I: normal weather prior warm-up. We train $\mathcal{A}_\theta$ on *Normal* samples to learn an identity-preserving refinement under radar conditioning and to stabilize optimization:

$$\mathcal{L}_I = \mathcal{L}_{\text{diff}} + \lambda_1 \mathcal{L}_{\text{det}}. \quad (9)$$

Stage II: adversarial alignment. We fine-tune $\mathcal{A}_\theta$ on the full training set $\mathcal{W}_{\text{train}}$. The discriminator term encourages refined features to match the *clean*-domain distribution, while the detection critic preserves task semantics:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{diff}} + \lambda_1 \mathcal{L}_{\text{det}} + \lambda_2 \mathcal{L}_{\text{adv}}. \quad (10)$$

Over Stage II, the weight on $\mathcal{L}_{\text{diff}}$ is annealed toward 0, so the frozen critics increasingly drive the alignment.

Inference. As shown in Fig. 3, we reuse the discriminator $\mathcal{D}$ as a lightweight router to decide whether to activate DCDA. Let $s = \mathcal{D}(\mathbf{F}_L) \in [0, 1]$ denote the confidence of *Normal*. We activate DCDA only when $s < \tau$ and otherwise bypass it:

$$\mathbf{F}_L^{\text{route}} = \begin{cases} \mathbf{F}_L, & s \geq \tau \quad (\text{bypass DCDA}), \\ \tilde{\mathbf{F}}_L, & s < \tau \quad (\text{activate DCDA}), \end{cases} \quad (11)$$

where $\tilde{\mathbf{F}}_L$ is the aligned LiDAR feature produced by DCDA when activated. The routed LiDAR features $\mathbf{F}_L^{\text{route}}$ are then fused with radar features and fed into the downstream detection network. We set τ on a held-out split from the training-weather domain without using any test data. Pseudo-code is provided in Sec. B of the supplementary material.

4 Experiments

4.1 Experimental Setup

Dataset. To evaluate the generalization capabilities of models beyond their training distributions, we establish a structured open-weather benchmark based on the K-Radar dataset. K-Radar provides synchronized LiDAR point clouds and 4D radar data across seven real-world conditions: Normal, Overcast, Fog, Rain, Sleet, LightSnow, and HeavySnow. Building on these data, we define three evaluation protocols along the dimensions of weather type and severity. Detailed statistics and split implementation details are provided in Sec. A of the supplementary material.

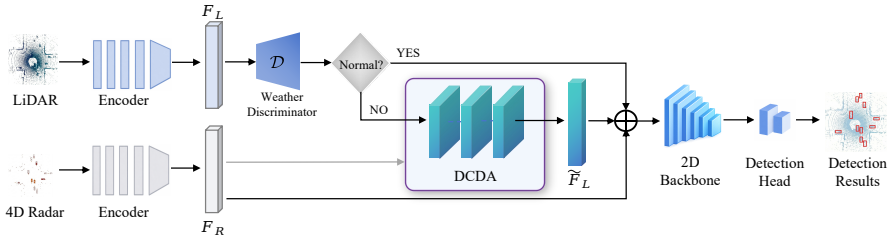


Fig. 3: Inference of DCDA with optional routing. When the input is confidently *Normal*, DCDA can be bypassed to reduce unnecessary refinement.

Table 1: Results on K-Radar **type-open**: per-weather BEV and 3D AP (%). Best and second-best are **bold** and underlined.

Method	Seen			Unseen			
	Normal	Rain	Sleet	Fog	HeavySnow	LightSnow	Overcast
BEV AP (%)							
RTNH [25]	35.64	38.50	30.41	72.43	31.63	61.93	64.96
InterFusion [35]	67.65	66.71	56.93	62.05	24.36	56.40	71.22
SpikingRTNH [24]	30.48	32.75	29.75	66.72	31.27	47.66	56.25
V2X-R [18]	65.96	<u>68.00</u>	<u>57.46</u>	69.43	30.37	66.72	78.99
L4DR [19]	<u>68.33</u>	68.16	56.02	<u>81.32</u>	<u>31.74</u>	<u>74.81</u>	<u>80.42</u>
DCDA (Ours)	68.39	67.15	64.76	88.81	35.94	77.07	81.76
3D AP (%)							
RTNH [25]	21.69	20.98	14.78	10.36	11.12	9.79	28.57
InterFusion [35]	42.88	41.04	39.51	34.26	11.52	28.96	40.51
SpikingRTNH [24]	11.50	19.03	10.46	23.40	19.23	8.07	29.96
V2X-R [18]	41.61	40.70	47.09	28.65	20.70	21.87	36.44
L4DR [19]	<u>44.71</u>	43.60	36.46	<u>36.73</u>	<u>21.88</u>	<u>29.60</u>	<u>49.85</u>
DCDA (Ours)	49.92	<u>41.81</u>	<u>44.69</u>	41.91	26.29	35.09	65.70

- **Type-open (real-world weather type holdout).** We train on *Normal* together with a subset of adverse weather types (*seen* types), and evaluate on the full set of real-world conditions. Adverse types not included in training (*unseen* types) are encountered only at test time, directly measuring generalization to novel weather categories. Concretely, we use the two most frequent adverse types in K-Radar—*Rain* and *Sleet*—as *seen*, and hold out *Overcast*, *Fog*, *LightSnow*, and *HeavySnow*, which span mild to severe degradations not represented by the seen set.
- **Severity-open (synthetic severity-level holdout).** Severity variation is known to induce substantially different LiDAR degradation patterns even within the same weather type (e.g., fog/snow scattering) [12, 13]. Since K-Radar does not provide controlled severity annotations, we construct a synthetic severity setting from *Normal* frames following Robo3D [20]. We consider three synthetic weather types (*Fog*, *Snow*, *Wet Ground*) with three severity levels (*Light/Medium/Heavy*). Training observes only the *Light* level for each synthetic type, while *Medium* and *Heavy* are held out for test-

Table 2: Results on synthetic **severity-open**: BEV and 3D AP (%) on Fog/Snow/Wet Ground at Light/Medium/Heavy levels. Training sees *Normal* and *Light* weather only.

Method	Fog			Snow			Wet Ground		
	L	M	H	L	M	H	L	M	H
BEV AP (%)									
RTNH [25]	39.39	30.32	34.60	33.55	26.20	<u>33.19</u>	40.88	27.25	32.55
InterFusion [35]	60.92	39.78	34.22	28.74	32.40	<u>24.92</u>	74.12	65.72	66.47
L4DR [19]	<u>65.55</u>	<u>44.43</u>	<u>35.13</u>	34.16	<u>35.98</u>	34.19	73.54	<u>67.19</u>	<u>67.08</u>
V2X-R [18]	65.45	39.53	32.20	28.16	34.78	30.96	79.50	62.44	<u>64.43</u>
DCDA (Ours)	68.81	44.70	35.16	42.35	44.27	32.77	<u>75.13</u>	67.40	67.37
3D AP (%)									
RTNH [25]	21.66	11.97	17.13	19.50	17.42	15.49	15.09	17.10	16.91
InterFusion [35]	<u>45.94</u>	19.04	22.86	19.01	21.61	15.21	41.89	36.63	41.14
L4DR [19]	45.76	<u>26.63</u>	26.18	<u>21.59</u>	<u>28.03</u>	22.55	43.02	<u>39.26</u>	<u>43.70</u>
V2X-R [18]	41.30	15.38	19.29	15.02	18.60	19.67	<u>45.06</u>	25.79	36.80
DCDA (Ours)	47.89	27.24	<u>25.23</u>	28.56	32.53	<u>20.73</u>	49.05	46.86	47.66

ing, evaluating robustness to unseen severity. Although synthesis starts from Normal frames, DCDA does not exploit paired clean–corrupted supervision; weather labels are used only to define train/test splits.

- **Type+severity-open.** We further evaluate the most challenging intersection setting, where training observes only a single type–severity combination, and all remaining type–severity combinations are evaluated at test time.

Metrics. Following the KITTI 3D detection protocol used in prior K-Radar works [19, 25], we report BEV AP (%) and 3D AP (%) for the *sedan* category at IoU = 0.5. Unless otherwise noted, all methods are evaluated using the same preprocessing and evaluation configuration.

Implementation details. The detection-guided critic $\mathcal{H}$ is pretrained on the *Normal*-only subset with a learning rate of 8×10^{-3} and kept frozen thereafter. The weather discriminator $\mathcal{D}$ is trained to distinguish *Normal* from non-*Normal* training samples with a learning rate of 5×10^{-4} , and is also frozen when training DCDA $\mathcal{A}_\theta$. We optimize $\mathcal{A}_\theta$ with AdamW and a cosine-annealing schedule for 40 epochs, using a batch size of 8 and a learning rate of 1×10^{-3} , following the two-stage procedure in Sec. 3.3. We set the number of diffusion steps to 3 for both training and inference, and use $(\lambda_1, \lambda_2) = (0.5, 0.05)$ in Eq. (10).

4.2 Main Results

To assess open-weather robustness for LiDAR–4D radar fusion 3D detection, we compare DCDA with representative baselines, including RTNH [25], InterFusion [35], V2X-R [18], and L4DR [19], under the structured *open-weather* protocols in Sec. 4.1. The detailed results for each protocol are reported and analyzed in the following sections.

Type-open (real weather types). To evaluate generalization to entirely unseen real-world weather types, we first conduct experiments under the type-open

Table 3: Results on the **type+severity-open**: BEV and 3D AP (%) over held-out type-severity pairs. Training sees *Normal* and *Fog-Light* only.

Method	Fog			Snow			Wet Ground		
	L	M	H	L	M	H	L	M	H
BEV AP (%)									
RTNH [25]	39.20	27.60	22.11	20.72	20.37	20.49	33.13	19.78	17.99
InterFusion [35]	58.09	<u>40.52</u>	30.21	20.73	16.29	17.85	70.03	64.62	<u>66.97</u>
L4DR [19]	<u>64.41</u>	41.03	<u>31.49</u>	21.43	21.88	<u>22.36</u>	<u>72.16</u>	64.55	69.12
V2X-R [18]	62.89	38.46	32.79	20.99	<u>22.96</u>	23.83	71.37	<u>64.78</u>	64.91
DCDA (Ours)	68.68	39.15	28.65	25.54	28.15	19.17	74.09	65.78	66.24
3D AP (%)									
RTNH [25]	10.50	12.39	10.21	<u>14.53</u>	8.90	11.77	13.12	13.64	12.20
InterFusion [35]	39.64	<u>24.36</u>	18.75	11.16	14.49	<u>15.89</u>	39.89	<u>40.15</u>	37.44
L4DR [19]	43.75	23.41	23.53	11.72	13.74	10.88	<u>49.59</u>	37.25	42.44
V2X-R [18]	<u>44.21</u>	23.50	<u>22.11</u>	11.87	<u>18.36</u>	15.91	44.88	38.85	40.46
DCDA (Ours)	45.38	27.25	19.83	16.27	19.48	12.46	51.92	41.15	<u>41.62</u>

protocol. Table 1 shows that while strong baselines (e.g., L4DR) remain competitive on the seen split, their performance degrades noticeably on unseen adverse types (e.g., *Fog* and *HeavySnow* in our split). In contrast, DCDA improves robustness on unseen types and reduces the seen-unseen gap. Averaged over all unseen weather types, DCDA surpasses the best baseline (L4DR) by **3.83** BEV AP and **7.73** 3D AP, while maintaining comparable accuracy on the seen split. Furthermore, to examine robustness under different evaluation operating points, we sweep IoU thresholds over $\{0.3, 0.5, 0.7\}$ and score thresholds over $\{0.1, 0.2, 0.3\}$. Across all nine combinations, DCDA has the highest *mean* unseen-weather BEV AP, leading the best baseline by **1.6** to **7.3** points. The largest gain is observed under the strict IoU = 0.7 setting, and the full grid is provided in Sec. C.1 of the supplementary material.

Severity-open (synthetic severity). To evaluate robustness to unseen severity levels within known synthetic weather types, we train on *Normal* and *Light*, and evaluate on *Light* and held-out *Medium/Heavy* levels. As reported in Table 2, DCDA achieves the strongest mean performance over the unseen combinations, improving unseen BEV/3D AP by **1.28/2.32** percentage points over L4DR. Across *Fog*, *Snow*, and *Wet Ground*, DCDA stays competitive as degradation increases. Under the most severe cases (e.g., *Heavy Snow*), performance remains challenging for all methods due to extreme LiDAR corruption, but DCDA still provides the best overall robustness trend.

Type+severity-open (synthetic type & severity). To evaluate the most challenging joint holdout where both weather type and severity may be unseen at test time, we train only on *Normal* and a single adverse configuration (*Fog-Light*), and evaluate on all type-severity combinations. As reported in Table 3, despite the extremely limited adverse exposure, DCDA attains the highest mean unseen BEV AP and the highest mean unseen 3D AP, exceeding V2X-R by **1.76** percentage points and L4DR by **2.18** percentage points on mean unseen 3D AP.

Table 4: Synthetic-trained models evaluated on real K-Radar weather (BEV/3D AP %). *Non-Normal* denotes the macro mean over the six adverse conditions.

Method	Severity-open		Type+severity-open	
	Normal	non-Normal	Normal	non-Normal
InterFusion	67.9/41.7	53.5/28.2	66.9/41.3	45.9/18.1
V2X-R	66.5/35.2	55.5/20.5	65.6/42.2	49.8/23.3
L4DR	67.9/ 49.1	56.1/ 30.0	67.3/43.4	50.6/25.5
DCDA (Ours)	73.3/48.2	58.8/27.3	67.1/ 43.4	54.2/26.8

Table 5: Real *LightSnow*→*HeavySnow* severity transfer (BEV/3D AP %): models trained on real *Normal* + *LightSnow* and tested on real *Normal* + *HeavySnow*.

Method	Normal	HeavySnow
L4DR	71.61/51.20	30.57/16.57
DCDA (Ours)	68.39/50.01	37.92/25.61

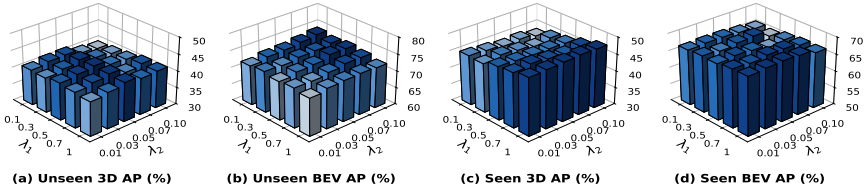
These results suggest DCDA can transfer beyond the single observed adverse setting.

Generalization to real weather. Since the severity-open and type+severity-open evaluations rely on *synthetic* weather, we further test DCDA on *real* adverse weather through the following two experiments on K-Radar: (1) *synthetic-to-real direct testing*, where models trained under the synthetic protocols are evaluated on real K-Radar weather conditions without further training; and (2) *real LightSnow-to-HeavySnow generalization*, where models are trained on real light snow and tested on real heavy snow.

- **Synthetic-to-real direct testing.** To examine whether models trained with synthetic adverse weather can generalize to real weather conditions, we take the models already trained under the severity-open and type+severity-open protocols and evaluate them, as is, on the seven real K-Radar conditions. Table 4 reports the averaged results, and the per-condition results are provided in Sec. C.2 of the supplementary material. Averaged over Rain, Sleet, Fog, Overcast, LightSnow, and HeavySnow, DCDA achieves the best BEV AP under both protocols, exceeding L4DR by **2.7** AP in the severity-open setting and by **3.6** AP in the type+severity-open setting, while maintaining competitive *Normal* accuracy.
- **Real LightSnow-to-HeavySnow generalization.** To further assess generalization across real severity levels, we train the models on real *Normal* + *LightSnow* and test them on real *Normal* + *HeavySnow*. As reported in Table 5, DCDA improves this challenging transfer case over L4DR by **7.35** BEV AP and **9.04** 3D AP on *HeavySnow*.

Table 6: Ablation study on K-Radar **type-open**: mean BEV and 3D AP (%) on seen and unseen weather types.

$\mathcal{L}_{\text{diff}}$	$\mathcal{L}_{\text{det}}$	$\mathcal{L}_{\text{adv}}$	Seen BEV	Seen 3D	Unseen BEV	Unseen 3D
✓			61.76	43.12	66.23	39.11
✓	✓		66.68	44.95	70.47	41.50
✓		✓	63.97	44.06	68.22	40.99
✓	✓	✓	66.77	45.47	72.65	42.25

**Fig. 4:** Parameter sensitivity analysis on K-Radar **type-open**: mean BEV/3D AP (%) on seen and unseen weather types.

4.3 Ablation Study

We perform ablation studies to analyze the contribution of each component in DCDA. Table 6 reports the results under the *type-open* protocol. Using structural diffusion alone ($\mathcal{L}_{\text{diff}}$) yields limited cross-weather robustness. Integrating the detection-guided critic ($\mathcal{L}_{\text{det}}$) significantly boosts performance, confirming the necessity of task-level semantic constraints. Adding the weather adversarial critic ($\mathcal{L}_{\text{adv}}$) encourages distributional alignment, further improving unseen-weather generalization. Adding both critics to the diffusion objective achieves the highest AP, showing that semantic preservation and distributional alignment are complementary for recovering a clean feature manifold.

4.4 Other Experiments

Parameter Sensitivity. We evaluate DCDA’s sensitivity to the critic weights λ_1 and λ_2 in Eq. (10) under the type-open protocol. We sweep $\lambda_1 \in \{0.1, 0.3, 0.5, 0.7, 1.0\}$ and $\lambda_2 \in \{0.01, 0.03, 0.05, 0.07, 0.1\}$ around the default (0.5, 0.05). Fig. 4 shows that DCDA remains stable across a broad hyperparameter neighborhood, suggesting that its dual-critic guidance is robust to moderate weight variations.

Efficiency and Routing. We further analyze the robustness–efficiency trade-off of DCDA under the type-open protocol, with results reported in Table 7. Applying DCDA to every frame yields the largest robustness gain of **6.69** unseen BEV AP, increasing latency from 29.51 to 70.86 ms. The routed variant skips DCDA for frames classified as *Normal* with high confidence by the discriminator $\mathcal{D}$. With 99.8% routing accuracy, it retains most of the robustness gain, improving unseen BEV AP by **4.94** points while reducing latency to 59.59 ms.

Table 7: Robustness–efficiency trade-off on K-Radar **type-open** (BEV AP %). Latency (ms), throughput (FPS), and peak memory (MiB) measured on a single GPU.

Method	Normal	Seen	Unseen	All	ms	FPS	MiB
w/o DCDA	68.41	64.28	65.96	65.24	29.51	33.89	470.7
+DCDA	68.43	66.87	72.65	70.13	70.86	14.11	573.7
+DCDA + routing	68.39	66.77	<u>70.90</u>	<u>69.13</u>	<u>59.59</u>	<u>16.78</u>	710.1

Table 8: Same-detector and same-budget generative-refiner comparison on K-Radar **type-open**. BEV/3D AP (%).

Refiner	Seen BEV/3D	Unseen BEV/3D
GAN refiner	61.40/43.01	65.73/40.52
DCDA (Ours)	66.77/45.47	72.65/42.25

Capacity Control. To isolate the effect of diffusion refinement from model capacity, we conduct a capacity-controlled comparison by replacing the diffusion aligner with a same-budget GAN refiner. Table 8 shows that DCDA exceeds this refiner by **6.92** BEV AP, indicating that the gain comes from coupling diffusion refinement with dual-critic guidance rather than added capacity.

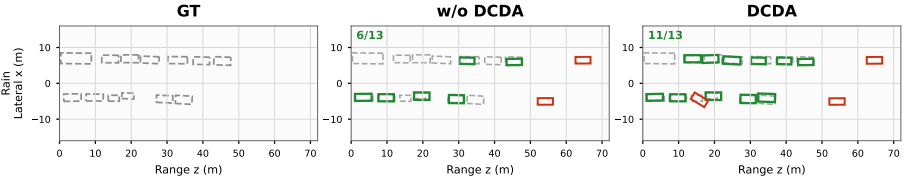


Fig. 5: Qualitative BEV detections on a real K-Radar *Rain* scene (**type-open**). Columns: ground truth (dashed gray), w/o DCDA, and DCDA (Ours). Green: matched predictions; red: false positives. Numbers: matched-GT out of total GT objects.

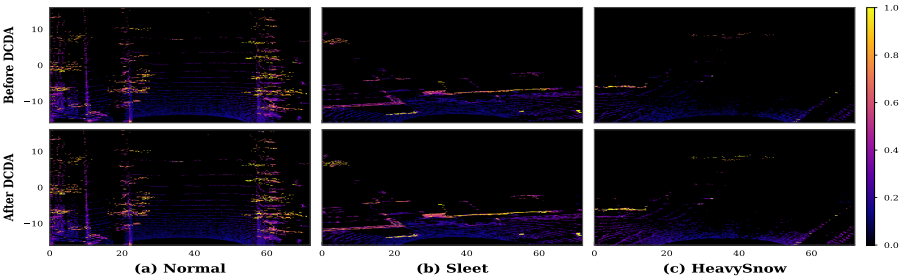


Fig. 6: Visualization on K-Radar **type-open**: intermediate BEV feature maps before (top) and after (bottom) applying DCDA under (a) *Normal*, (b) *Sleet*, (c) *HeavySnow* weather conditions.

Visualization Results. To qualitatively understand how DCDA helps under adverse weather, we visualize its effect at both the detection and feature levels. At the detection level, Fig. 5 shows BEV detections on a real *Rain* scene, where DCDA correctly matches 11 of the 13 ground-truth objects, versus only 6 for the baseline. At the feature level, Fig. 6 shows that DCDA reshapes the intermediate LiDAR BEV feature maps by strengthening weak object-related responses and partially recovering suppressed evidence, while preserving the LiDAR-centric layout.

4.5 Limitations

While DCDA delivers consistent open-weather gains, several limitations remain and point to future work. First, iterative refinement adds inference overhead, leaving a speed–robustness trade-off that routing alleviates but does not eliminate. Second, DCDA’s added margin is smaller at the most extreme severities (e.g., *Snow-Heavy*, *Fog-Heavy*), where LiDAR geometry is largely degraded and little structure remains for the anchored refinement to recover—a regime that is intrinsically challenging for most fusion methods. Finally, our controlled severity evaluation is synthetic and DCDA currently assumes *Normal* as the clean reference; extending to richer real-severity annotations and heterogeneous clean references is future work.

5 Conclusion

We investigate open-weather generalization for LiDAR–4D radar fusion 3D detection, where test-time weather may differ from training conditions in both type and severity. We propose DCDA, a weather-agnostic diffusion alignment framework that restores degraded LiDAR features through 4D radar-guided refinement with complementary semantic and distributional supervision. Without paired clean–degraded data or weather type/severity labels as conditional inputs, DCDA consistently improves robustness under unseen weather while remaining competitive on seen conditions. We hope this work provides a practical step toward robust and generalizable multi-modal 3D perception in real-world weather.

Acknowledgements

This research was supported in part by the National Natural Science Foundation of China under Grant 62576167, the Young Scientists Fund of the National Natural Science Foundation of China under Grant 62402220, and the Natural Science Foundation of Jiangsu Province under Grant BK20241402.

References

1. Batten, B., Lomuscio, A.: Improving weather-based ood generalisation in lidar-based object detection models via adversarial training. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 4321–4329 (2025)
2. Bijelic, M., Gruber, T., Mannan, F., Kraus, F., Ritter, W., Dietmayer, K., Heide, F.: Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 11682–11692 (2020)
3. Chae, Y., Kim, H., Oh, C., Kim, M., Yoon, K.J.: Lidar-based all-weather 3d object detection via prompting and distilling 4d radar. In: *European Conference on Computer Vision*. pp. 368–385. Springer (2024)
4. Chae, Y., Kim, H., Yoon, K.J.: Towards robust 3d object detection with lidar and 4d radar fusion in various weather conditions. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 15162–15172 (2024)
5. Chae, Y., Park, H., Kim, H., Yoon, K.J.: Doppler-aware lidar-radar fusion for weather-robust 3d detection. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 27197–27208 (2025)
6. Chang, G., Roh, W., Jang, S., Lee, D., Ji, D., Oh, G., Park, J., Kim, J., Kim, S.: Cmda: Cross-modal and domain adversarial adaptation for lidar-based 3d object detection. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 38, pp. 972–980 (2024)
7. Chen, X., Liu, Z., Luo, K., Datta, S., Polavaram, A., Wang, Y., You, Y., Li, B., Pavone, M., Chao, W.L.H., et al.: Diffubox: Refining 3d object detection with point diffusion. *Advances in Neural Information Processing Systems* **37**, 103681–103705 (2024)
8. Chen, Z., Meng, J., Baktashmotlagh, M., Zhang, Y., Huang, Z., Luo, Y.: Mos: Model synergy for test-time adaptation on lidar-based 3d object detection. *arXiv preprint arXiv:2406.14878* (2024)
9. Chen, Z., Wang, Z., Luo, Y., Wang, S., Huang, Z.: Dpo: dual-perturbation optimization for test-time adaptation in 3d object detection. In: *Proceedings of the 32nd ACM International Conference on Multimedia*. pp. 4138–4147 (2024)
10. Ding, F., Wen, X., Zhu, Y., Li, Y., Lu, C.X.: Radarocc: Robust 3d occupancy prediction with 4d imaging radar. *Advances in Neural Information Processing Systems* **37**, 101589–101617 (2024)
11. Dong, Y., Kang, C., Zhang, J., Zhu, Z., Wang, Y., Yang, X., Su, H., Wei, X., Zhu, J.: Benchmarking robustness of 3d object detection to common corruptions. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 1022–1032 (2023)
12. Hahner, M., Sakaridis, C., Bijelic, M., Heide, F., Yu, F., Dai, D., Van Gool, L.: Lidar snowfall simulation for robust 3d object detection. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 16364–16374 (2022)
13. Hahner, M., Sakaridis, C., Dai, D., Van Gool, L.: Fog simulation on real lidar point clouds for 3d object detection in adverse weather. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 15283–15292 (2021)
14. Han, Z., Wang, J., Xu, Z., Yang, S., He, L., Xu, S., Wang, J., Li, K.: 4d millimeter-wave radar in autonomous driving: A survey. *arXiv preprint arXiv:2306.04242* (2023)

15. He, Z., Liu, F., Li, Y., Luo, Z., Cheng, J., Chen, X., Tang, X.: Diffusion-based restoration for multi-modal 3d object detection in adverse weather. *arXiv preprint arXiv:2512.13107* (2025)
16. Huang, H., Yan, X., Yang, J., Cao, Y., Zhang, X.: Lidsor: A filter for removing rain and snow noise points from lidar point clouds in rainy and snowy weather. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **48**, 733–740 (2023)
17. Huang, K., Shi, B., Li, X., Li, X., Huang, S., Li, Y.: Multi-modal sensor fusion for auto driving perception: A survey. *arXiv preprint arXiv:2202.02703* (2022)
18. Huang, X., Wang, J., Xia, Q., Chen, S., Yang, B., Li, X., Wang, C., Wen, C.: V2x-r: Cooperative lidar-4d radar fusion with denoising diffusion for 3d object detection. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 27390–27400 (2025)
19. Huang, X., Xu, Z., Wu, H., Wang, J., Xia, Q., Xia, Y., Li, J., Gao, K., Wen, C., Wang, C.: L4dr: Lidar-4dradar fusion for weather-robust 3d object detection. In: *Proceedings of the AAAI conference on artificial intelligence*. vol. 39, pp. 3806–3814 (2025)
20. Kong, L., Liu, Y., Li, X., Chen, R., Zhang, W., Ren, J., Pan, L., Chen, K., Liu, Z.: Robo3d: Towards robust and reliable 3d perception against corruptions. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 19994–20006 (2023)
21. Lee, E., Jung, M., Kim, A.: Toward robust lidar based 3d object detection via density-aware adaptive thresholding. *arXiv preprint arXiv:2404.13852* (2024)
22. Matteazzi, A., Arnold, M., Tutsch, D.: Augmentation of lidar scenes with adverse weather conditions using latent diffusion models. In: *International Conference on Neural Information Processing*. pp. 198–212. Springer (2025)
23. Oh, Y., Kim, H.I., Kim, S.T., Kim, J.U.: Monowad: Weather-adaptive diffusion model for robust monocular 3d object detection. In: *European Conference on Computer Vision*. pp. 326–345. Springer (2024)
24. Paek, D.H., Kong, S.H.: Spikingtrnh: Spiking neural network for 4d radar object detection. *arXiv preprint arXiv:2502.00074* (2025)
25. Paek, D.H., Kong, S.H., Wijaya, K.T.: K-radar: 4d radar object detection for autonomous driving in various weather conditions. *Advances in Neural Information Processing Systems* **35**, 3819–3829 (2022)
26. Palladin, E., Dietze, R., Narayanan, P., Bijelic, M., Heide, F.: Samfusion: Sensor-adaptive multimodal fusion for 3d object detection in adverse weather. In: *European Conference on Computer Vision*. pp. 484–503. Springer (2024)
27. Peng, X., Sun, H., Bierzynski, K., Fischbacher, A., Servadei, L., Wille, R.: Mutualforce: Mutual-aware enhancement for 4d radar-lidar 3d object detection. In: *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 1–5. IEEE (2025)
28. Peng, X., Tang, M., Sun, H., Servadei, L., Wille, R.: 4d mmwave radar in adverse environments for autonomous driving: A survey. *arXiv e-prints* pp. arXiv–2503 (2025)
29. Peng, X., Wang, Y., Tang, M., Kay, B., Servadei, L., Wille, R.: Moral: Motion-aware multi-frame 4d radar and lidar fusion for robust 3d object detection. *arXiv preprint arXiv:2505.09422* (2025)
30. Peng, X., Zhu, X., Ma, Y.: Cl3d: Unsupervised domain adaptation for cross-lidar 3d detection. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 37, pp. 2047–2055 (2023)

31. Pirolì, A., Dallabetta, V., Kopp, J., Walessa, M., Meissner, D., Dietmayer, K.: Energy-based detection of adverse weather effects in lidar data. *IEEE Robotics and Automation Letters* **8**(7), 4322–4329 (2023)
32. Qi, Y., Liu, C., Scaioni, M., Li, Y., Qiao, Y., Ma, X., Wu, H., Zhang, K., Wang, D.: Geometric information constraint 3d object detection from lidar point cloud for autonomous vehicles under adverse weather. *Transportation research part C: emerging technologies* **161**, 104555 (2024)
33. Qi, Y., Liu, C., Wu, H., Chen, R., Wen, C., Huang, X., Jia, S., Zhang, K.: Fusionbev: Lidar and 4d radar fusion for 3d object detection. *Information Fusion* p. 104240 (2026)
34. Sural, S., Sahu, N., Rajkumar, R.R.: Contextualfusion: Context-based multi-sensor fusion for 3d object detection in adverse operating conditions. In: 2024 IEEE intelligent vehicles symposium (IV). pp. 1534–1541. IEEE (2024)
35. Wang, L., Zhang, X., Xv, B., Zhang, J., Fu, R., Wang, X., Zhu, L., Ren, H., Lu, P., Li, J., et al.: Interfusion: Interaction-based 4d radar and lidar fusion for 3d object detection. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 12247–12253. IEEE (2022)
36. Wang, Y., Deng, J., Li, Y., Hu, J., Liu, C., Zhang, Y., Ji, J., Ouyang, W., Zhang, Y.: Bi-lrfusion: Bi-directional lidar-radar fusion for 3d dynamic object detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 13394–13403 (2023)
37. Wu, Y., Zhu, Y., Zhang, K., Qian, J., Xie, J., Yang, J.: Weathergen: A unified diverse weather generator for lidar point clouds via spider mamba diffusion. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 17019–17028 (2025)
38. Wu, Y., Sun, W., Zhong, Z., Li, Q.: Pasetnet: Snowy scene 3d object detection with pillar-wise attention and semantic enhancement. *IET Image Processing* **20**(1), e70257 (2026)
39. Xu, R., Xiang, Z.: Rlnet: Adaptive fusion of 4d radar and lidar for 3d object detection. In: *European Conference on Computer Vision*. pp. 181–194. Springer (2024)
40. Ye, X., Yaman, B., Cheng, S., Tao, F., Mallik, A., Ren, L.: Bevdiffuser: Plug-and-play diffusion model for bev denoising with ground-truth guidance. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 1495–1504 (2025)
41. Yuan, J., Zhang, B., Gong, K., Yue, X., Shi, B., Qiao, Y., Chen, T.: Reg-tta3d: Better regression makes better test-time adaptive 3d object detection. In: *European conference on computer vision*. pp. 197–213. Springer (2024)
42. Zhang, B., Wang, Y., Zhang, C., Jiang, J., Luo, X., Wang, X., Zhang, Y., Liu, Z., Shen, G., Ye, Y., et al.: Fogfusion: Robust 3d object detection based on camera-lidar fusion for autonomous driving in foggy weather conditions. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* p. 09544070251327229 (2025)
43. Zhang, Z., Gong, H., Feng, Y., Chu, Z., Liu, H.: Enhancing object detection in adverse weather conditions through entropy and guided multimodal fusion. In: *Proceedings of the Asian Conference on Computer Vision*. pp. 2922–2938 (2024)
44. Zou, J., Tian, K., Zhu, Z., Ye, Y., Wang, X.: Diffbev: Conditional diffusion model for bird’s eye view perception. In: *Proceedings of the AAAI conference on artificial intelligence*. vol. 38, pp. 7846–7854 (2024)