DOME: DISTORTION-DRIVEN DOMAIN MIXTURE EXPERTS FOR ALL-WEATHER MULTIMODAL 3D OBJECT DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Reliable 3D perception under adverse weather is essential for autonomous driving, yet remains challenging due to weather-induced sensor degradation. Different adverse weather conditions introduce heterogeneous physical distortions into LiDAR observations, leading to latent domain shifts that are difficult to characterize with discrete weather categories. Existing weather-aware approaches typically rely on image-based weather classifiers to route features according to predefined weather labels, resulting in a mismatch between semantic weather categories and actual sensor degradation patterns, especially under degraded visual conditions. To address this challenge, we propose DOME, a Distortion-Driven Domain Mixture-of-Experts framework for all-weather multimodal 3D object detection. DOME models LiDAR–4D Radar fusion through latent physical-distortion domain decomposition, where experts are specialized according to distortion patterns rather than predefined weather identities. Specifically, the Physical Distortion-guided Router (PDR) extracts distortion fingerprints from multimodal features and guides Distortion-Specialized Experts (DSE) to perform distortion-aware feature adaptation. Furthermore, we introduce DOME-L with LoRA-based Distortion Experts (LDE), reducing the parameter overhead of distortion-driven MoE while preserving domain-specific adaptation capability. Extensive experiments on K-Radar and OmniHD-Scenes demonstrate strong performance. On K-Radar, DOME improves L4DR by 8.1 AP_{3D} points overall, while a capacity-controlled comparison isolates a 1.9-point gain from distortion-aware routing. Additional evaluations demonstrate robustness under continuously varying and held-out degradation conditions, as well as transferability to a different multimodal detection backbone. The open-source code will be publicly available.

1 INTRODUCTION

3D object detection (Lang et al., 2019; Yin et al., 2021) serves as a fundamental perception task for autonomous driving and robotic systems. While current LiDAR-based methods (Shi et al., 2020; Deng et al., 2021) achieve remarkable performance under clear conditions, their reliability drops sharply in adverse weather (Dong et al., 2023; Kong et al., 2023). Adverse weather conditions (e.g., rain, snow, and fog) introduce distinct physical and geometric distortions to LiDAR point clouds (Hahner et al., 2021; Huang et al., 2024). To mitigate this, integrating LiDAR's precise geometry with 4D Radar's weather-resilience has emerged as a promising paradigm for reliable 3D detection in harsh environments (Chae et al., 2024b; Huang et al., 2025; Paek & Kong, 2025).

Despite the significant perceptual advantages brought by multimodal fusion, existing methods typically adopt a mixed-weather training strategy to process diverse weather conditions (Chae et al., 2024b; Huang et al., 2025; Paek & Kong, 2025). However, this paradigm overlooks an important characteristic of real-world adverse weather: different weather conditions induce heterogeneous physical distortions with varying degradation patterns (see Fig. 2 (a)), leading to substantial cross-weather distribution shifts (see Fig. 2 (b)). Treating these diverse distortion patterns as a single shared distribution limits the model's ability to adapt to condition-specific degradation, leading to suboptimal performance across different weather scenarios (Yang et al., 2025; Lin et al., 2026).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

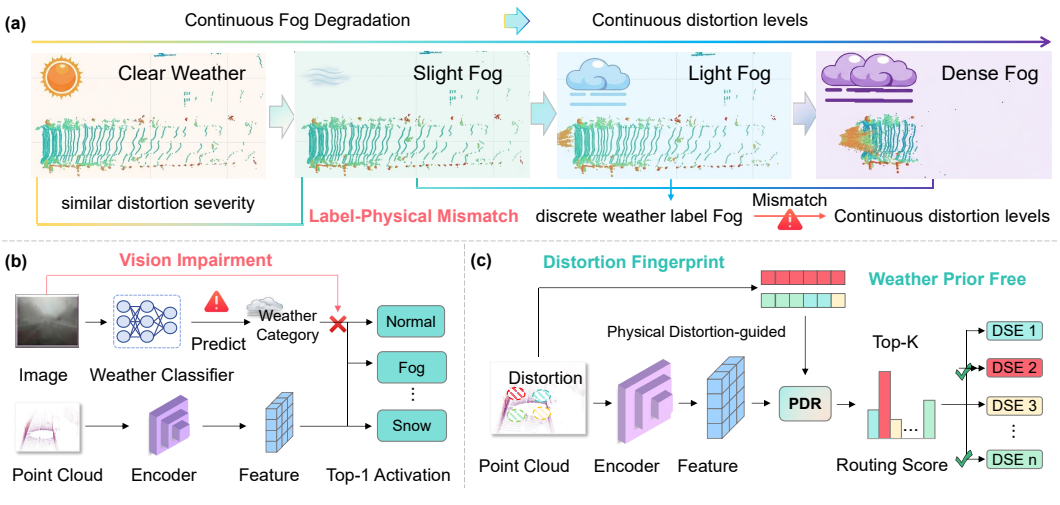


Figure 1: (a) Label-physical mismatch between discrete weather labels and continuous physical distortions. (b)(c) Comparison between the existing weather-aware pipeline using image-based weather priors and our distortion-driven approach using latent distortion fingerprints.

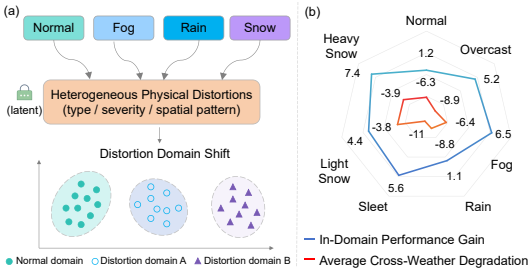


Figure 2: Motivation for distortion-aware domain modeling under adverse weather. (a) Different weather conditions induce heterogeneous physical distortions with varying types, severities, and spatial patterns. (b) Single-weather fine-tuning on L4DR improves the target condition but degrades others, illustrating cross-weather distribution shifts.

Recent studies attempt to address such cross-weather domain shifts through weather-aware adaptation strategies. However, existing approaches (Yang et al., 2025; Lin et al., 2026) typically rely on image-based weather classifiers to route point cloud features into discrete weather-specific branches (see Fig. 1 (b)), implicitly treating weather categories as proxies for underlying distortion domains. This approximation suffers from **label-physical mismatch**: discrete weather labels cannot accurately characterize the heterogeneous physical distortions in point clouds, where samples with the same weather condition may exhibit different degradation patterns (see Fig. 1 (a)). Consequently, weather categories provide only a coarse semantic description of environmental conditions rather than a faithful representation of the actual sensor degradation states. Moreover, such approaches introduce **strong dependence on visual weather recognition**, as the routing process relies on external visual weather recognition and corresponding annotations, which may become unreliable under degraded or unavailable visual conditions (see Fig. 1 (b)). Therefore, a key challenge is to discover latent physical-distortion domains directly from intrinsic multimodal observations, enabling adaptive modeling of diverse degradation patterns without relying on predefined weather categories.

To address this challenge, we propose DOME, a Distortion-Driven Domain Mixture-of-Experts framework that models all-weather LiDAR–4D Radar fusion with latent physical-distortion domain decomposition. Instead of relying on predefined weather categories, DOME learns distortion-aware representations from multimodal observations and specializes experts according to latent physical-distortion domains (see Fig. 1 (c)). Specifically, the Physical Distortion-guided Router (PDR) extracts distortion fingerprints by jointly modeling global multimodal context and local cross-modal discrepancies, which guide the Distortion-Specialized Experts (DSE) to learn sample-adaptive, domain-specific feature transformations (Shazeer et al., 2017; Lepikhin et al., 2020). Different from conventional weather-specific branches, DOME decomposes domain variations according to latent physical distortions rather than discrete weather identities.

Furthermore, we introduce DOME-L with LoRA-based Distortion Experts (LDE) to efficiently adapt distortion-specialized experts while preserving their domain-specific capability. By integrating

LoRA (Hu et al., 2022; Wu et al., 2024) into the distortion-domain MoE (Shazeer et al., 2017; Lepikhin et al., 2020) framework, DOME-L reduces the parameter count from 302M to 78M, providing a more parameter-efficient design for resource-constrained deployment. In summary, our main contributions include:

- We propose DOME, a Distortion-Driven Domain Mixture-of-Experts framework that models all-weather LiDAR–4D Radar fusion through latent physical-distortion domain decomposition, enabling expert specialization according to weather-induced physical distortions rather than predefined weather categories.
- We design the PDR to extract distortion fingerprints from multimodal features, enabling distortion-aware expert adaptation and addressing the mismatch between weather categories and physical distortions. Furthermore, we introduce LDE to reduce the parameter overhead of distortion-driven MoE, enabling efficient deployment while preserving distortion-specific adaptation capability.
- Extensive experiments on K-Radar and OmniHD-Scenes demonstrate DOME's effectiveness, robustness under ambiguous weather, and scalability across multimodal architectures.

2 Related Work

Robust Multimodal 3D Object Detection. 3D object detection is a fundamental task for autonomous driving and other unmanned systems (Xia et al., 2023; Wu et al., 2025; Lin et al., 2025; Xia et al., 2026; Zhong et al., 2026). Mainstream LiDAR-based methods achieve impressive performance in clear weather (Yan et al., 2018; Lang et al., 2019; Shi et al., 2020; Yin et al., 2021; Deng et al., 2021; Yue et al., 2025). However, adverse weather like fog and snow introduces atmospheric particles that scatter and refract LiDAR beams, leading to severe physical distortions and degraded detection accuracy (Dong et al., 2023; Kong et al., 2023; Huang et al., 2024). To achieve robust perception, recent studies such as 3D-LRF (Chae et al., 2024b) and L4DR (Huang et al., 2025) focus on fusing LiDAR and 4D Radar. These methods aim to jointly leverage the fine-grained geometric profiling capability of LiDAR and the strong atmospheric penetrability of 4D Radar (Wang et al., 2022; 2023; Chae et al., 2024a; Paek & Kong, 2025).

Cross-Weather Domain Adaptation. Although multimodal fusion improves robustness under adverse weather, existing approaches (Wang et al., 2022; 2023; Chae et al., 2024a) often treat diverse weather conditions as a shared data distribution and overlook the heterogeneous physical distortions introduced into point clouds. Such a strategy may hinder effective adaptation when different degradation patterns exist across weather scenarios (Chae et al., 2024b; Huang et al., 2025). Recent methods, including DA3D (Yang et al., 2025) and AW-MoE (Lin et al., 2026), attempt to alleviate cross-weather domain shifts through weather-aware routing with image-based classifiers. However, using semantic weather categories as adaptation cues cannot fully capture the underlying physical distortion variations and introduces a dependency on visual weather recognition. To address these limitations, we propose DOME, which introduces a distortion-driven Mixture-of-Experts paradigm to adapt multimodal representations according to latent physical-distortion domains rather than predefined weather categories.

Efficient Mixture-of-Experts. Mixture-of-Experts (MoE) (Jacobs et al., 1991; Fedus et al., 2022; Rajbhandari et al., 2022; Du et al., 2022; Shah et al., 2026) provides an effective paradigm for conditional computation by dynamically activating specialized experts through routing mechanisms (Lepikhin et al., 2020; Liu et al., 2025; Xu et al., 2026). However, conventional MoE architectures often introduce substantial parameter and memory overhead (Masoudnia & Ebrahimpour, 2014; Riquelme et al., 2021; Lin et al., 2024), limiting their practical deployment in resource-constrained autonomous systems. To improve efficiency, Parameter-Efficient Fine-Tuning (PEFT), particularly Low-Rank Adaptation (LoRA) (Bian et al., 2025; Ouyang et al., 2025; Che et al., 2026), has been incorporated into MoE frameworks. Recent studies further explore LoRA-based expert parameterization, enabling lightweight expert adaptation while preserving shared knowledge (Wu et al., 2024; Chen et al., 2024; Li et al., 2024; Gou et al., 2026). Building upon these advances, we introduce LoRA-based Distortion Experts (LDE) to efficiently realize distortion-specialized adaptation within the proposed distortion-driven MoE framework, reducing expert parameter overhead while preserving domain-specific adaptation capability.

3 METHODOLOGY

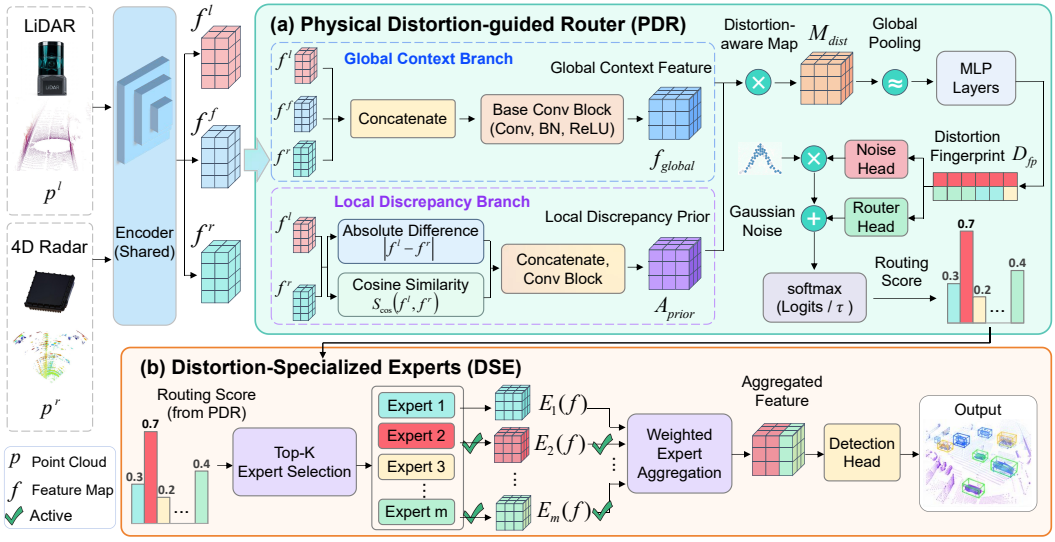


Figure 3: Overview of DOME. (a) The Physical Distortion-guided Router (PDR) learns distortion fingerprints from global multimodal context and local cross-modal discrepancies, generating distortion-aware routing decisions. (b) The Distortion-Specialized Experts (DSE) perform Top-K adaptive activation to learn distortion-specific feature transformations, enabling expert specialization according to latent physical-distortion domains.

3.1 PROBLEM STATEMENT AND OVERALL DESIGN

Robust Multimodal 3D Object Detection. To achieve robust 3D perception under adverse weather, LiDAR and 4D Radar sensors are commonly fused to exploit their complementary sensing characteristics. Given the point clouds generated by LiDAR and Radar, $\mathcal{P}^m = \{p_i^m\}_{i=1}^{N_m}$, where $m \in \{l, r\}$ denotes the modality, a multimodal encoder $\mathcal{M}$ extracts feature representations: $f^m = E_{\mathcal{M}}(\mathcal{P}^m; \Theta)$. The extracted LiDAR and Radar features are then fused by a multimodal fusion module $\Phi_{\mathcal{M}}$: $f^f = \Phi_{\mathcal{M}}(f^l, f^r)$, and the fused feature is finally processed by the detection head $H_{\mathcal{M}}$ to predict 3D bounding boxes $B = \{b_i\}_{i=1}^{N_b}$, $B \in \mathbb{R}^{N_b \times 7}$. Although multimodal fusion improves robustness, the extracted features still contain weather-dependent physical distortion variations, motivating distortion-aware domain modeling.

Overview. As illustrated in Fig. 3, DOME introduces a distortion-driven Mixture-of-Experts framework to model latent physical-distortion domains in all-weather LiDAR–4D Radar fusion. After extracting multimodal features, the Physical Distortion-guided Router (PDR) learns feature-level distortion fingerprints by capturing both global multimodal context and local cross-modal discrepancies. These fingerprints provide distortion-aware routing signals for activating the Distortion-Specialized Experts (DSE), where each expert learns domain-specific feature transformations corresponding to different latent distortion patterns. To improve the deployment efficiency of distortion-driven MoE, DOME-L further introduces LoRA-based Distortion Experts (LDE), which achieve lightweight expert adaptation with reduced parameter overhead (see Fig. 4).

3.2 PHYSICAL DISTORTION-GUIDED ROUTER (PDR)

Existing weather-aware routing approaches (Yang et al., 2025; Lin et al., 2026) rely on image-derived weather categories as domain indicators, but such semantic cues cannot fully represent the underlying physical distortion variations. Therefore, DOME discovers latent physical-distortion domains directly from multimodal observations. To this end, we design the Physical Distortion-guided Router (PDR), which extracts feature-level distortion fingerprints to characterize weather-induced degradation patterns and generate distortion-aware expert routing.

Dual-Branch Distortion Fingerprint Extraction. The key challenge of distortion-domain discovery is to capture both global degradation characteristics and local modality inconsistencies. As illustrated in Fig. 3 (a), PDR adopts a dual-branch design to construct distortion fingerprints from complementary perspectives. The global branch models the overall distortion distribution by aggregating LiDAR (f^l), 4D Radar (f^r), and fused (f^f) features. Specifically, the concatenated multimodal representation is processed by a base convolutional block (Conv, BN, ReLU) to obtain the global distortion feature:

$$f_{global} = \text{ConvBlock}(\text{Concat}(f^l, f^r, f^f)). \tag{1}$$

Meanwhile, the local branch captures modality-level distortion discrepancies by exploiting the complementary sensing characteristics of LiDAR and 4D Radar. Since LiDAR is generally more susceptible to weather-induced degradation while 4D Radar is comparatively robust, the discrepancy between the two modalities provides informative distortion cues. We concatenate the absolute feature difference and cosine similarity between f^l and f^r, followed by a convolutional block, to obtain the local distortion prior attention:

$$\begin{aligned} A_{prior} &= \text{ConvBlock}(\text{Concat}(|f^l - f^r|, S_{cos}(f^l, f^r))), \\ S_{cos}(f^l, f^r) &= \frac{f^l \cdot f^r}{\|f^l\|\|f^r\|} = \frac{\sum_{c=1}^{C} f_c^l \cdot f_c^r}{\sqrt{\sum_{c=1}^{C}(f_c^l)^2} \cdot \sqrt{\sum_{c=1}^{C}(f_c^r)^2}}. \end{aligned} \tag{2}$$

The local distortion prior is then combined with the global distortion feature through element-wise modulation: $M_{dist} = f_{global} \odot A_{prior}$. This operation preserves the global distortion context while emphasizing regions with strong cross-modal degradation responses. Finally, global pooling and MLP layers are applied to obtain the scene-level distortion fingerprint: $D_{fp} = \text{MLP}(\text{Global Pooling}(M_{dist}))$.

Uncertainty-aware Routing Head. Based on the extracted distortion fingerprint D_{fp}, PDR predicts distortion-aware routing scores for expert specialization. To improve routing exploration and reduce premature expert collapse during early training, we adopt an adaptive noisy-routing strategy, termed Adaptive Logit Exploration (ALE), following the general idea of noisy gating in MoE (Shazeer et al., 2017). Specifically, a noise head predicts an input-dependent noise scale, which modulates Gaussian perturbations added to the routing logits:

$$R = \text{Softmax}\left(\frac{H_{route}(D_{fp}) + H_{noise}(D_{fp}) \cdot \epsilon}{\tau}\right), \quad \epsilon \sim \mathcal{N}(0, I). \tag{3}$$

During inference, Gaussian perturbations are disabled for deterministic routing. Through distortion fingerprint extraction and uncertainty-aware routing, PDR learns latent distortion representations without relying on explicit weather priors, enabling distortion-specialized expert adaptation.

3.3 Distortion-Specialized Experts (DSE)

After extracting feature-level distortion fingerprints, PDR provides distortion-aware routing signals to specialize expert adaptation. Based on these latent distortion domains, we design Distortion-Specialized Experts (DSE) to learn domain-specific feature transformations for mitigating weather-induced distribution variations. As illustrated in Fig. 3 (b), each DSE focuses on a specific distortion pattern and transforms multimodal features accordingly, enabling the MoE framework to adapt beyond predefined weather categories.

Top-K Adaptive Activation. Since real-world physical distortions exhibit diverse patterns and severities, assigning each sample to a single expert may limit the model's ability to capture mixed distortion characteristics. Therefore, DSE adopts a Top-K adaptive activation strategy to aggregate multiple relevant experts according to the routing score R. Let $\mathcal{S}$ denote the indices of the selected Top-K experts. The normalized gating weight is defined as $G_i = R_i / \sum_{j \in \mathcal{S}} R_j$ if $i \in \mathcal{S}$ and $G_i = 0$ otherwise.

Adaptive Expert Aggregation. The selected experts perform distortion-specific feature transformations and are aggregated using the normalized Top-K gating weights:

$$f_{agg} = \sum_{i \in \mathcal{S}} G_i \cdot \text{DSE}_i(\{f^l, f^r, f^f\}). \tag{4}$$

This adaptive aggregation enables the model to combine complementary expert responses according to the inferred distortion distribution, producing refined multimodal representations for robust 3D object detection. The aggregated feature f_{agg} is subsequently fed into the detection head to regress the final 3D bounding boxes of target objects.

3.4 LoRA-based Distortion Experts (LDE)

Although DSE enables distortion-specific expert adaptation, deploying multiple full experts introduces considerable parameter overhead. To improve the efficiency of distortion-driven MoE, we introduce DOME-L, a lightweight variant of DOME that replaces full DSE with LoRA-based Distortion Experts (LDE) by incorporating LoRA (Hu et al., 2022) into the DSE architecture (Fig. 4). Instead of optimizing independent full-rank experts, LDE retains a shared feature extraction module and learns lightweight LoRA-based residual experts, allowing each expert to focus on distortion-specific adaptation with substantially fewer parameters.

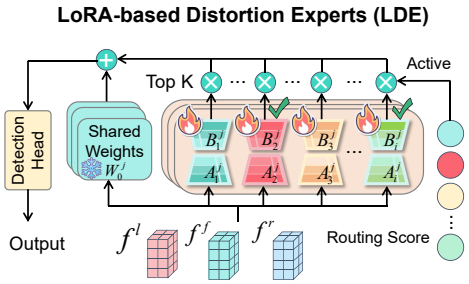


Figure 4: LDE Framework.

Specifically, the shared module captures general 3D perception knowledge, while the LoRA branches specialize in latent distortion domains by learning expert-specific residual updates. DOME-L adopts a two-stage optimization strategy. In the first stage, the shared encoder, shared module W_0, and detection head are trained with PDR and LDE inactive to establish general 3D perception. During the second-stage optimization, the shared parameters W_0 remain frozen, while the low-rank expert parameters are optimized jointly with PDR and the detection head. The LDE aggregation follows the same Top-K activation and normalized gating as DSE, with an additional shared feature pathway:

$$f_{agg} = \text{Shared Block}(\{f^l, f^r, f^f\}) + \sum_{i \in \mathcal{S}} G_i \cdot \text{LDE}_i(\{f^l, f^r, f^f\}). \tag{5}$$

DOME-L preserves the distortion-driven adaptation capability of DOME while substantially reducing expert parameter overhead, enabling more efficient deployment.

3.5 Training Strategy and Loss Function

Training distortion-driven MoE models requires stable expert specialization, as premature routing convergence may cause expert collapse and prevent effective learning of distortion-specific representations (Masoudnia & Ebrahimpour, 2014). To address this issue, we introduce the Progressive Expert Maturation (PEM) strategy, which stabilizes expert optimization through the Temperature Annealing Scheme (TAS), Asymptotic Backpropagation Scaling (ABS), and load balancing.

Progressive Expert Maturation. During the early training stage, PEM encourages sufficient expert exploration before establishing distortion-aware specialization. Specifically, the Temperature Annealing Scheme (TAS) gradually adjusts the routing temperature from a smoother distribution to a sharper selection process, allowing the router to explore diverse experts before forming stable distortion-specific assignments. Meanwhile, the Asymptotic Backpropagation Scaling (ABS) progressively increases the influence of PDR optimization, preventing unstable routing decisions caused by premature router updates. The detailed scheduling functions and transition analysis are provided in the supplementary material.

Loss Functions. To further encourage balanced expert utilization, we adopt the auxiliary load-balancing loss (Fedus et al., 2022):

$$\mathcal{L}_{bal} = N_{exp} \cdot \sum_{i=1}^{N_{exp}} F_i P_i, \quad \mathcal{L}_{total} = \mathcal{L}_{det} + \alpha \mathcal{L}_{bal}, \tag{6}$$

where F_i denotes the fraction of Top-K assignments routed to expert i, and P_i denotes its average routing probability. This optimization strategy stabilizes expert training and facilitates effective specialization across latent distortion domains.

4 EXPERIMENTS

Table 1: Performance of different 3D object detection methods on the K-Radar dataset at IoU=0.5. Best in **bold**, second in underline. † denotes the published AW-MoE result, while ‡ denotes our reproduction without its proposed data augmentation.

Method	Modality	Metric	Total	Normal	Overcast	Fog	Rain	Sleet	Light Snow	Heavy Snow
InterFusion (IROS 2022)	L+4DR	AP_{BEV}	66.1	70.5	82.0	81.8	67.2	33.9	62.9	46.0
		AP_{3D}	41.7	44.6	53.5	64.8	37.2	25.5	35.4	27.0
3D-LRF (CVPR 2024)	L+4DR+C	AP_{BEV}	73.6	72.3	88.4	86.6	76.6	47.5	79.6	64.1
		AP_{3D}	45.2	45.3	55.8	51.8	38.3	23.4	60.2	36.9
L4DR (AAAI 2025)	L+4DR	AP_{BEV}	77.5	76.8	88.6	89.7	78.2	59.3	80.9	53.8
		AP_{3D}	53.5	53.0	64.1	73.2	53.8	46.2	52.4	37.0
L4DR-DA3D (MM 2025)	L+4DR+C	AP_{BEV}	78.5	77.4	89.1	90.1	79.3	58.8	88.9	60.6
		AP_{3D}	**61.9**	58.9	66.4	79.2	63.0	48.2	64.6	47.6
L4DR-RAF (ECCV 2026)	L+4DR+C	AP_{BEV}	82.0	82.0	88.2	93.9	82.6	63.0	88.5	60.4
		AP_{3D}	57.4	59.4	67.6	79.3	53.3	45.9	56.3	37.8
AW-MoE† (arXiv 2026)	L+4DR+C	AP_{BEV}	**84.2**	82.8	91.6	96.3	**85.3**	**75.0**	**94.7**	66.4
		AP_{3D}	61.5	59.0	67.2	**85.7**	**63.5**	43.3	**70.1**	**53.1**
AW-MoE‡ (arXiv 2026)	L+4DR+C	AP_{BEV}	81.7	80.9	89.1	95.2	79.8	64.6	89.7	60.7
		AP_{3D}	59.6	56.1	67.0	84.2	60.3	47.0	59.9	45.4
DOME (Ours)	L+4DR	AP_{BEV}	83.2	**84.1**	**94.2**	**97.2**	81.3	68.9	93.7	**68.2**
		AP_{3D}	61.6	**60.5**	**75.6**	85.3	58.2	**48.4**	66.7	47.6

Table 2: Performance of different 3D object detection methods on the OmniHD-Scenes benchmark. We present the modality of each method (L: LiDAR, 4DR: 4D Radar, C: Camera). Best in **bold**.

Method	Modality	Vehicle $IoU = 0.5$		Pedestrian $IoU = 0.3$		Cyclist $IoU = 0.3$	
		AP	APH	AP	APH	AP	APH
PointPillars (CVPR 2019)	L	73.11	71.29	23.63	13.98	60.70	53.81
L4DR (AAAI 2025)	L+4DR	75.08	73.92	24.68	14.06	62.96	57.01
L4DR-DA3D (MM 2025)	L+4DR+C	75.47	74.11	24.70	14.06	63.39	57.62
DOME (Ours)	L+4DR	**76.33**	**74.62**	**25.81**	**14.51**	**64.17**	**58.51**

4.1 DATASETS AND IMPLEMENTATION DETAILS

Datasets. We evaluate DOME on K-Radar (Paek et al., 2022) and OmniHD-Scenes (Zheng et al., 2026), two multimodal autonomous-driving benchmarks with LiDAR, cameras, and 4D Radar. K-Radar contains 34,944 synchronized frames covering normal and six adverse weather conditions. Following the standard protocol, we evaluate the Sedan category using AP_{3D} and AP_{BEV} at IoU thresholds of 0.3 and 0.5. OmniHD-Scenes contains over 450K synchronized frames across sunny, cloudy, and rainy conditions. Following its official protocol, we report AP_{3D} and heading-weighted AP (AP_H), with IoU thresholds of 0.5 for Vehicles and 0.3 for Pedestrians and Cyclists.

Implementation Details. We instantiate DOME on L4DR (Huang et al., 2025). We use $N_{exp} = 6$ distortion experts with Top-K routing ($K = 2$), and set the LoRA rank of LDE to $r = 8$. For Progressive Expert Maturation, we set $E_{trans} = 5$, $\tau_{start} = 2.0$, $\lambda_{start} = 0.1$, and the load-balancing weight $\alpha = 0.01$. All experiments use a total batch size of 4 on four NVIDIA RTX 3090 GPUs. Additional implementation details are provided in Appendix A.

4.2 MAIN RESULTS

Main Results on K-Radar under Diverse Weather Conditions. We compare DOME with representative multimodal 3D detectors, including L4DR (Huang et al., 2025), L4DR-DA3D (Yang et al., 2025), L4DR-RAF (Park et al., 2026) and AW-MoE (Lin et al., 2026), among others. As shown in Table 1, DOME achieves strong performance across diverse weather conditions, with an overall 8.1-point AP_{3D} gain over L4DR at $IoU = 0.5$. For the closely related weather-aware AW-MoE, we report both its original result with the proposed data augmentation (†) and the augmentation-free result (‡). DOME achieves 83.2/61.6 AP_{BEV}/AP_{3D} using only LiDAR and 4D Radar, comparable to the augmented AW-MoE at 84.2/61.5, without requiring the camera modality. Under the augmentation-free setting, DOME further obtains 1.5/2.0-point higher AP_{BEV}/AP_{3D} than AW-MoE.

Details of AW-MoE's augmentation strategy and an augmentation-matched comparison are provided in Appendix B.6.

Generalization on OmniHD-Scenes. To evaluate the generalization capability of DOME beyond K-Radar, we further conduct experiments on the OmniHD-Scenes dataset (Zheng et al., 2026). As reported in Table 2, DOME consistently improves over the L4DR baseline across all object categories, including Vehicle, Pedestrian, and Cyclist. Despite fewer weather variations than K-Radar, DOME achieves consistent gains, demonstrating its generalization to different multimodal perception scenarios. These results demonstrate effective transfer of DOME to a second multimodal benchmark with different sensor configurations and object categories beyond K-Radar.

4.3 Validating Learned Distortion Fingerprints

Table 3: Capacity-controlled comparison on K-Radar at IoU=0.5.

Routing	Params [M]	FLOPs [G]	AP_{3D}
Generic-MoE	301.78	195.82	59.7
PDR (DOME)	301.82	195.83	**61.6**

Table 4: Comparison of routing cues on K-Radar at IoU=0.5.

Routing Cue	Weather	Camera	AP_{3D}
Weather-fixed routing	✓		60.2
Image weather posterior	✓	✓	59.4
PDR fingerprint			**61.6**

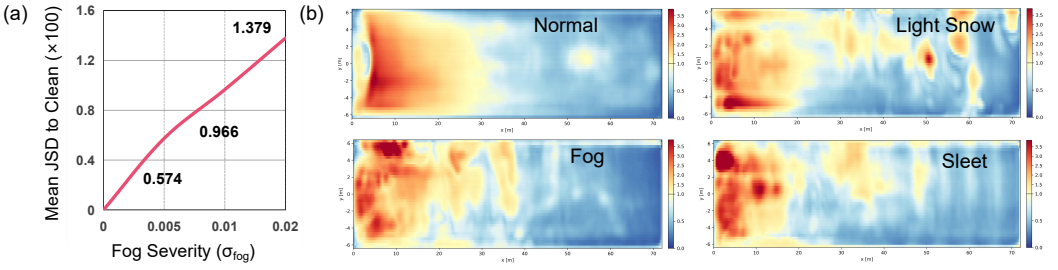


Figure 5: Visualization of learned distortion fingerprints. (a) Routing divergence increases with fog severity. (b) PDR spatial attribution exhibits smooth responses under normal weather and clustered patterns under adverse weather.

Capacity-Controlled Generic-MoE Comparison. To disentangle MoE capacity from distortion-aware routing, we replace PDR with a generic router using globally pooled multimodal features while keeping the number of experts, Top-K activation, and training protocol unchanged. As shown in Table 3, Generic-MoE achieves 59.7 AP_{3D} and PDR 61.6 under nearly identical parameters and FLOPs, isolating a 1.9-point gain from distortion-aware routing.

Routing Cue Comparison. Table 4 compares PDR with two weather-based routing alternatives. *Weather-fixed routing* averages the learned routing distributions within each ground-truth weather category and uses the ground-truth category at inference, while *Image weather posterior* uses weather probabilities predicted from camera images. They achieve 60.2 and 59.4 AP_{3D}, respectively, compared with 61.6 for the sample-adaptive PDR fingerprint. The 1.4-point gap to weather-fixed routing indicates that PDR captures sample-specific degradation variations beyond fixed weather identities, while the lower performance of image-based routing suggests weaker visual weather cues.

Continuous Distortion Response. To examine whether the learned fingerprint responds to continuous physical degradation, we progressively increase fog severity on the same clean scenes while keeping scene content and Radar observations unchanged. As shown in Fig. 5 (a), the routing divergence from the clean counterpart increases monotonically with fog severity. Since scene semantics remain fixed, this trend indicates that PDR responds to progressive sensor degradation rather than merely encoding discrete weather identities or scene-specific content.

Spatial Interpretation of Distortion Fingerprints. We further visualize the spatial attribution of PDR to understand the degradation cues underlying its routing decisions. As shown in Fig. 5 (b), normal weather exhibits a smooth distance-decaying response, consistent with regular LiDAR sensing characteristics. In contrast, adverse weather produces irregular and spatially clustered attribution patterns, reflecting heterogeneous weather-induced distortions. These distinct spatial responses

provide complementary evidence that PDR captures structured, physically meaningful degradation patterns beyond discrete weather categories.

Overall, these complementary analyses provide consistent evidence that the learned PDR fingerprints reflect sample-specific physical degradation beyond discrete weather semantics. Further analyses of routing behavior and distortion fingerprints are provided in Appendix B.1.

4.4 Robustness and Practical Evaluation

Robustness across Fog Degradation Levels. Beyond analyzing routing behavior, we evaluate whether distortion sensitivity yields task-level robustness by comparing DOME with two weather-aware methods, L4DR-DA3D and AW-MoE, across simulated fog intensities (Hahner et al., 2021). All methods use the same simulated-fog training data and held-out test subsets, with camera images degraded at the same severity as the corresponding LiDAR observations. As shown in Table 5, DOME consistently achieves higher AP_{BEV} and AP_{3D} across all fog levels while using only LiDAR and 4D Radar. DOME maintains nearly stable AP_{3D} across fog intensities, demonstrating robustness across varying degradation levels without relying on discrete weather cues.

Table 5: Robustness under different simulated fog intensities at IoU=0.5.

Method	Modality	σ_{fog}	AP_{BEV}	AP_{3D}
L4DR-DA3D	L+4DR+C		70.4	42.3
AW-MoE	L+4DR+C	0.005	69.5	40.2
DOME	L+4DR		**75.7**	**51.0**
L4DR-DA3D	L+4DR+C		74.0	45.9
AW-MoE	L+4DR+C	0.010	72.1	43.6
DOME	L+4DR		**77.5**	**51.1**
L4DR-DA3D	L+4DR+C		73.8	46.3
AW-MoE	L+4DR+C	0.020	72.5	43.4
DOME	L+4DR		**76.1**	**51.2**

Generalization to Held-out Snow Severities. We separately remove Light Snow or Heavy Snow from training while retaining the other snow severity, and evaluate directly on the corresponding held-out subset without test-time adaptation. As shown in Table 6, DOME retains 87.2/56.7 AP_{3D}, outperforming L4DR by 24.1/19.7 points. Its performance drops are also limited to 2.6/2.5 points, compared with 15.8/14.9 for L4DR, indicating stronger cross-severity generalization within the snow family.

Table 6: Generalization to Held-out Snow Severities. AP_{3D} is reported at IoU=0.3.

Method	Light Snow			Heavy Snow		
	Seen	Held-out	Δ	Seen	Held-out	Δ
L4DR	78.9	63.1	-15.8	51.9	37.0	-14.9
DOME	**89.8**	**87.2**	**-2.6**	**59.2**	**56.7**	**-2.5**

Parameter-efficient Distortion Experts. Table 7 compares DOME with lightweight expert designs in terms of accuracy and computational efficiency. DOME-L reduces the parameter count from 301.82M to 78.17M and FLOPs from 195.83G to 164.87G, with only a 0.3-point drop in AP_{3D}. We further construct two capacity-matched baselines: *UDGA-A*, which replaces the LoRA branches with UDGA-style bottleneck adapters (Chang et al., 2024), and *LoRA-MoE-G*, which retains the same low-rank experts but uses generic MoE routing. Under comparable parameters and FLOPs, DOME-L achieves 61.3 AP_{3D}, outperforming UDGA-A (59.8) and LoRA-MoE-G (60.0). These results suggest that its gains arise from combining parameter-efficient distortion experts with distortion-aware routing, rather than lightweight parameterization alone.

Table 7: Accuracy–efficiency comparison of DOME and lightweight expert designs on L4DR.

Method	AP_{3D}	FPS [Hz]	FLOPs [G]	Params[M]
L4DR	53.5	13.94	142.65	59.73
DOME	**61.6**	11.92	195.83	301.82
UDGA-A	59.8	13.52	162.67	78.28
LoRA-MoE-G	60.0	13.63	162.52	78.15
DOME-L	**61.3**	13.40	164.87	78.17

Additional analyses of sensor-degradation robustness and random-seed stability, cross-backbone scalability and parameter-efficient DOME-L, PEM and hyper-parameter sensitivity, and qualitative detection results are provided in Appendix B.2, B.3, B.4, B.5, and C.

5 Conclusion

This paper presents DOME, a Distortion-Driven Domain Mixture-of-Experts framework that reformulates all-weather LiDAR–4D Radar fusion from weather-aware adaptation to latent physical-distortion domain modeling. Instead of relying on external image-based weather priors, DOME discovers distortion fingerprints through the PDR and enables expert specialization according to degradation patterns. Extensive experiments demonstrate that DOME effectively mitigates weather-induced domain shifts and remains robust under ambiguous weather. Furthermore, DOME-L provides an efficient realization of distortion-specialized adaptation with reduced expert parameter overhead.

AI USE STATEMENT

Generative AI tools were used for language editing, presentation improvement, and consistency checking during manuscript preparation. The methodology, experimental design, implementation, and experimental results were produced by the authors. All scientific interpretations and conclusions were independently reviewed and approved by the authors, who take full responsibility for the final manuscript.

ETHICS STATEMENT

This work uses publicly released autonomous-driving benchmarks and does not collect new human-subject or private data. We follow the usage terms of the respective datasets.

REPRODUCIBILITY STATEMENT

We provide detailed descriptions of the proposed DOME framework, including the model architecture, training strategy, and implementation details in the main paper and supplementary materials. Additional experimental settings, ablation studies, robustness analyses, and qualitative evaluations are provided in the supplementary material. The implementation code will be released upon acceptance to facilitate reproduction and further research.

REFERENCES

Jieming Bian, Lei Wang, Letian Zhang, and Jie Xu. Lora-fair: Federated lora fine-tuning with aggregation and initialization refinement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3737–3746, 2025.

Yujeong Chae, Hyeonseong Kim, Changgyoon Oh, Minseok Kim, and Kuk-Jin Yoon. Lidar-based all-weather 3d object detection via prompting and distilling 4d radar. In *European Conference on Computer Vision*, pp. 368–385. Springer, 2024a.

Yujeong Chae, Hyeonseong Kim, and Kuk-Jin Yoon. 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, 2024b.

Gyusam Chang, Jiwon Lee, Donghyun Kim, Jinkyu Kim, Dongwook Lee, Daehyun Ji, Sujin Jang, and Sangpil Kim. Unified domain generalization and adaptation for multi-view 3d object detection. *Advances in Neural Information Processing Systems*, 37:58498–58524, 2024.

Chang Che, Ziqi Wang, Pengwan Yang, Cheems Wang, Hui Ma, and Zenglin Shi. Lora in lora: Towards parameter-efficient architecture expansion for continual visual instruction tuning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 19978–19986, 2026.

Shaoxiang Chen, Zequn Jie, and Lin Ma. Llava-mole: Sparse mixture of lora experts for mitigating data conflicts in instruction finetuning mllms. *arXiv preprint arXiv:2401.16160*, 2024.

Jiajun Deng, Shaoshuai Shi, Peiwei Li, Wengang Zhou, Yanyong Zhang, and Houqiang Li. Voxel r-cnn: Towards high performance voxel-based 3d object detection. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 1201–1209, 2021.

Yinpeng Dong, Caixin Kang, Jinlai Zhang, Zijian Zhu, Yikai Wang, Xiao Yang, Hang Su, Xingxing Wei, and Jun Zhu. 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.

Nan Du, Yanping Huang, Andrew M Dai, Simon Tong, Dmitry Lepikhin, Yuanzhong Xu, Maxim Krikun, Yanqi Zhou, Adams Wei Yu, Orhan Firat, et al. Glam: Efficient scaling of language models with mixture-of-experts. In *International conference on machine learning*, pp. 5547–5569. PMLR, 2022.

William Fedus, Barret Zoph, and Noam Shazeer. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity. *Journal of Machine Learning Research*, 23(120):1–39, 2022.

Yunhao Gou, Zhili Liu, Kai Chen, Lanqing Hong, Hang Xu, Zhenguo Li, Dit-Yan Yeung, James T Kwok, and Yu Zhang. Mixture of cluster-conditional lora experts for vision-language instruction tuning. *IEEE Transactions on Image Processing*, 35:3881–3892, 2026.

Martin Hahner, Christos Sakaridis, Dengxin Dai, and Luc Van Gool. 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.

Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nZeVKeeFYf9.

Xun Huang, Hai Wu, Xin Li, Xiaoliang Fan, Chenglu Wen, and Cheng Wang. Sunshine to rainstorm: Cross-weather knowledge distillation for robust 3d object detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 2409–2416, 2024.

Xun Huang, Ziyu Xu, Hai Wu, Jinlong Wang, Qiming Xia, Yan Xia, Jonathan Li, Kyle Gao, Chenglu Wen, and Cheng Wang. L4dr: Lidar-4dradar fusion for weather-robust 3d object detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 3806–3814, 2025.

Robert A Jacobs, Michael I Jordan, Steven J Nowlan, and Geoffrey E Hinton. Adaptive mixtures of local experts. *Neural computation*, 3(1):79–87, 1991.

Lingdong Kong, Youquan Liu, Xin Li, Runnan Chen, Wenwei Zhang, Jiawei Ren, Liang Pan, Kai Chen, and Ziwei Liu. Robo3d: Towards robust and reliable 3d perception against corruptions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 19994–20006, 2023.

Alex H Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, and Oscar Beijbom. Pointpillars: Fast encoders for object detection from point clouds. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12697–12705, 2019.

Dmitry Lepikhin, HyoukJoong Lee, Yuanzhong Xu, Dehao Chen, Orhan Firat, Yanping Huang, Maxim Krikun, Noam Shazeer, and Zhifeng Chen. Gshard: Scaling giant models with conditional computation and automatic sharding. *arXiv preprint arXiv:2006.16668*, 2020.

Dengchun Li, Yingzi Ma, Naizheng Wang, Zhengmao Ye, Zhiyuan Cheng, Yinghao Tang, Yan Zhang, Lei Duan, Jie Zuo, Cal Yang, et al. Mixlora: Enhancing large language models fine-tuning with lora-based mixture of experts. *arXiv preprint arXiv:2404.15159*, 2024.

Bin Lin, Zhenyu Tang, Yang Ye, Jiaxi Cui, Bin Zhu, Peng Jin, Jinfa Huang, Junwu Zhang, Yatian Pang, Munan Ning, et al. Moe-llava: Mixture of experts for large vision-language models. *arXiv preprint arXiv:2401.15947*, 2024.

Hongwei Lin, Dongyu Pan, Qiming Xia, Hai Wu, Cheng Wang, Siqi Shen, and Chenglu Wen. Pretend benign: A stealthy adversarial attack by exploiting vulnerabilities in cooperative perception. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 19947–19956, 2025.

Hongwei Lin, Xun Huang, Chenglu Wen, and Cheng Wang. Aw-moe: All-weather mixture of experts for robust multi-modal 3d object detection. *arXiv preprint arXiv:2603.16261*, 2026.

Xinyi Liu, Yujie Wang, Fangcheng Fu, Xupeng Miao, Shenhan Zhu, Xiaonan Nie, and Bin Cui. Netmoe: Accelerating moe training through dynamic sample placement. In *International Conference on Learning Representations*, volume 2025, pp. 90113–90131, 2025.

Saeed Masoudnia and Reza Ebrahimpour. Mixture of experts: a literature survey. *Artificial Intelligence Review*, 42(2):275–293, 2014.

Ziheng Ouyang, Zhen Li, and Qibin Hou. K-lora: Unlocking training-free fusion of any subject and style loras. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13041–13050. IEEE, 2025.

Dong-Hee Paek and Seung-Hyun Kong. Availability-aware sensor fusion via unified canonical space for 4d radar, lidar, and camera. *arXiv preprint arXiv:2503.07029*, 2025.

Dong-Hee Paek, Seung-Hyun Kong, and Kevin Tirta Wijaya. K-radar: 4d radar object detection for autonomous driving in various weather conditions. *Advances in Neural Information Processing Systems*, 35:3819–3829, 2022.

Heejun Park, Jaeseok Jeong, and Kuk-Jin Yoon. Raf: Reliability-aware fusion of camera, lidar, and 4d radar for robust 3d object detection. In *Computer Vision–ECCV 2026: 19th European Conference, Malmö, Sweden, September 8–12, 2026, Proceedings, Part XVIII*, pp. 201. Springer Nature, 2026.

Samyam Rajbhandari, Conglong Li, Zhewei Yao, Minjia Zhang, Reza Yazdani Aminabadi, Ammar Ahmad Awan, Jeff Rasley, and Yuxiong He. Deepspeed-moe: Advancing mixture-of-experts inference and training to power next-generation ai scale. In *International conference on machine learning*, pp. 18332–18346. PMLR, 2022.

Carlos Riquelme, Joan Puigcerver, Basil Mustafa, Maxim Neumann, Rodolphe Jenatton, André Susano Pinto, Daniel Keysers, and Neil Houlsby. Scaling vision with sparse mixture of experts. *Advances in Neural Information Processing Systems*, 34:8583–8595, 2021.

Aamir Shah, Yu Wen, Renjie Hu, Jiefu Chen, Jose Contreras-Vidal, Xuqing Wu, and Xin Fu. Cogmoe: Signal-quality–guided multimodal moe for cognitive load prediction. In *International Conference on Learning Representations*, volume 2026, pp. 110563–110584, 2026.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *arXiv preprint arXiv:1701.06538*, 2017.

Shaoshuai Shi, Chaoxu Guo, Li Jiang, Zhe Wang, Jianping Shi, Xiaogang Wang, and Hongsheng Li. Pv-rcnn: Point-voxel feature set abstraction for 3d object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10529–10538, 2020.

Li Wang, Xinyu Zhang, Baowei Xv, Jinzhao Zhang, Rong Fu, Xiaoyu Wang, Lei Zhu, Haibing Ren, Pingping Lu, Jun Li, 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.

Yingjie Wang, Jiajun Deng, Yao Li, Jinshui Hu, Cong Liu, Yu Zhang, Jianmin Ji, Wanli Ouyang, and Yanyong Zhang. 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.

Hai Wu, Hongwei Lin, Xusheng Guo, Xin Li, Mingming Wang, Cheng Wang, and Chenglu Wen. Motal: Unsupervised 3d object detection by modality and task-specific knowledge transfer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6284–6293, 2025.

Xun Wu, Shaohan Huang, and Furu Wei. Mixture of lora experts. *arXiv preprint arXiv:2404.13628*, 2024.

Qiming Xia, Jinhao Deng, Chenglu Wen, Hai Wu, Shaoshuai Shi, Xin Li, and Cheng Wang. Coin: Contrastive instance feature mining for outdoor 3d object detection with very limited annotations. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6254–6263, 2023.

Qiming Xia, Hongwei Lin, Wei Ye, Yadan Luo, Shijia Zhao, Hai Wu, Chenglu Wen, and Cheng Wang. Label-efficient outdoor 3d object detection via single click annotation from lidar point cloud. *ISPRS Journal of Photogrammetry and Remote Sensing*, 234:151–166, 2026.

Zukang Xu, Zhixiong Zhao, Xing Hu, Zhixuan Chen, and Dawei Yang. KBVQ-moe: KLT-guided SVD with bias-corrected vector quantization for moe large language models. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=veFs5UfYq9.

Yan Yan, Yuxing Mao, and Bo Li. Second: Sparsely embedded convolutional detection. *Sensors*, 18 (10):3337, 2018.

Haochen Yang, Lei Li, Jiacheng Guo, Baolu Li, Minghai Qin, Hongkai Yu, and Tianyun Zhang. Da3d: Domain-aware dynamic adaptation for all-weather multimodal 3d detection. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pp. 2150–2158, 2025.

Tianwei Yin, Xingyi Zhou, and Philipp Krahenbuhl. Center-based 3d object detection and tracking. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11784–11793, 2021.

Jingtong Yue, Zhiwei Lin, Xin Lin, Xiaoyu Zhou, Xiangtai Li, Lu Qi, Yongtao Wang, and Ming-Hsuan Yang. Roburcdet: Enhancing robustness of radar-camera fusion in bird's eye view for 3d object detection. In *International Conference on Learning Representations*, volume 2025, pp. 12726–12741, 2025.

Lianqing Zheng, Long Yang, Qunshu Lin, Wenjin Ai, Minghao Liu, Shouyi Lu, Jianan Liu, Hongze Ren, Jingyue Mo, Xiaokai Bai, et al. Omnihd-scenes: A next-generation multimodal dataset for autonomous driving. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026.

Jianping Zhong, Zhaobo Qi, Kaiwen Duan, Xinyan Liu, Beichen Zhang, Weigang Zhang, and Qingming Huang. Ptnet: A proposal-centric transformer net-work for 3d object detection. In *International Conference on Learning Representations*, volume 2026, pp. 85386–85402, 2026.

A Detailed Experimental Setups

A.1 Ambiguous Weather Evaluation

To evaluate robustness under ambiguous and continuously varying weather degradation, we conduct controlled fog simulation experiments on the "Normal" subset of K-Radar. Following the physics-based fog simulation method (Hahner et al., 2021), we use $\sigma_{\text{fog}} \in \{0.005, 0.01, 0.02\}$ for both simulated-fog training augmentation and evaluation. Training and test scenes are disjoint, while the fog severity values themselves are shared. All compared methods use the same simulated-fog training data and held-out test subsets. For camera-dependent baselines, the corresponding camera images are degraded using the same fog severity as the LiDAR observations, ensuring synchronized multimodal degradation and a fair comparison.

No fog-intensity information or weather-category labels are provided to DOME during training or inference. PDR predicts routing scores solely from the multimodal observations and learned distortion fingerprints, allowing DOME to adapt to different degradation levels without explicit weather priors.

A.2 Model Instantiation and Expert Architectures

DOME is instantiated on the L4DR (Huang et al., 2025) architecture. The early backbone layers are retained as a shared encoder to extract preliminary LiDAR and 4D Radar representations before the Physical Distortion-guided Router (PDR).

After routing, the remaining backbone components are instantiated as Distortion-Specialized Experts (DSE). Each DSE inherits the Inter-Modal and Intra-Modal (IM^2) design of L4DR, including its multi-scale convolutional and deconvolutional operations. Thus, all experts preserve the original LiDAR–4D Radar fusion architecture while learning distortion-specific feature transformations.

For parameter-efficient adaptation, DOME-L replaces the full DSE branches with LoRA-based Distortion Experts (LDE). One full-rank expert block is retained as a shared module for general multimodal perception, while the remaining expert-specific transformations are parameterized using lightweight LoRA residual updates. During the second-stage optimization of DOME-L, the shared encoder and shared expert block remain frozen, while PDR, the LoRA-based distortion experts, and the detection head are optimized. DOME-L otherwise follows the same Top-K routing and expert aggregation mechanism as DOME.

A.3 Training Protocol and Progressive Expert Maturation

Training Protocol. DOME is trained end-to-end by jointly optimizing the encoder, PDR, DSE, and detection head on the complete mixed-weather K-Radar training split. Only the original 3D bounding-box annotations are used; no weather-category labels, target-weather subsets, or additional adaptation annotations are introduced.

DOME-L follows a two-stage training protocol. In the first stage, the encoder, shared expert block, and detection head are trained while PDR and the LoRA-based distortion experts remain inactive, yielding an L4DR-style pretrained detector. In the second stage, the encoder and shared expert block are frozen, and PDR, LDE, and the detection head are jointly optimized for distortion-specific adaptation.

Progressive Expert Maturation. To encourage stable expert specialization, we progressively adjust the routing temperature τ and the PDR gradient scaling factor λ_{PDR} during the transition period of E_{trans} epochs. At epoch e, Temperature Annealing Scheme (TAS) follows

$$\tau(e) = \begin{cases} \tau_{\text{start}} - \dfrac{e}{E_{\text{trans}}}(\tau_{\text{start}} - 1.0), & e < E_{\text{trans}}, \\ 1.0, & e \geq E_{\text{trans}}, \end{cases} \tag{7}$$

while Asymptotic Backpropagation Scaling (ABS) uses

$$\lambda_{\text{PDR}}(e) = \begin{cases} \lambda_{\text{start}} + \dfrac{e}{E_{\text{trans}}}(1.0 - \lambda_{\text{start}}), & e < E_{\text{trans}}, \\ 1.0, & e \geq E_{\text{trans}}. \end{cases} \tag{8}$$

At the beginning of training, TAS uses a relatively high temperature τ_{start} to flatten the routing distribution and encourage broader expert exploration. Meanwhile, ABS applies a reduced gradient scale λ_{start} to suppress premature updates of PDR. As training progresses, τ decreases toward 1.0 while λ_{PDR} increases toward 1.0, gradually transitioning from expert exploration to stable distortion-aware specialization.

Both DOME and DOME-L are optimized using AdamW with a learning rate of 0.001. Unless otherwise specified, all experiments use the same dataset split, data-processing protocol, and optimization settings as the corresponding baseline.

A.4 Fusion Module Details

DOME directly inherits the LiDAR–4D Radar fusion pipeline from L4DR (Huang et al., 2025) without modifying its fusion architecture. LiDAR and 4D Radar features are first spatially aligned in the shared BEV representation and then processed by the original multimodal fusion module.

The resulting fused feature, together with the modality-specific LiDAR and Radar features, is used by PDR to construct the distortion representation. Therefore, the performance gains of DOME do not arise from introducing a stronger fusion backbone, but from distortion-aware routing and expert specialization built on top of the original L4DR representation.

A.5 Complete Training Configuration of DOME and DOME-L

DOME Training. DOME is trained end-to-end for 35 epochs using AdamW with an initial learning rate of 1×10^{-3}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and a weight decay of 0.01. A cosine annealing scheduler is applied every epoch with a minimum learning rate of 1×10^{-4}. The total batch size is 4, and 16 data-loading workers are used without balanced sampling. Following the Progressive Expert Maturation strategy, the router transition lasts for the first 5 epochs, with the routing temperature initialized as $\tau_{\text{start}} = 2.0$ and the PDR gradient scaling factor initialized as $\lambda_{\text{start}} = 0.1$. All modules are jointly optimized throughout the end-to-end training process.

DOME-L Training. DOME-L follows a two-stage optimization protocol. In Stage 1, the base perception model is trained for 35 epochs while PDR and LDE remain inactive. We use AdamW with an initial learning rate of 1×10^{-3}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, weight decay 0.01, and cosine annealing to a minimum learning rate of 1×10^{-4}. The batch size and number of data-loading workers are 4 and 16, respectively, with balanced sampling disabled.

Stage 2 is initialized from the Stage 1 checkpoint and trained for an additional 15 epochs. During this stage, the shared encoder and shared expert block are frozen, while PDR, the low-rank expert branches, and the detection head are jointly optimized with the same AdamW configuration and cosine learning-rate schedule. The router transition is applied over the first 3 epochs of Stage 2, with $\tau_{\text{start}} = 2.0$ and $\lambda_{\text{start}} = 0.1$. All experiments are conducted on four NVIDIA RTX 3090 GPUs.

B More Experimental Results

B.1 Additional Analysis of Distortion-Driven Routing

B.1.1 Expert Routing Distributions across Weather Conditions

To better understand how DOME utilizes distortion-specialized experts, we visualize the average routing distributions across different weather conditions in Fig. 6. Distinct conditions exhibit different expert-utilization patterns. For example, Normal and Overcast assign relatively larger weights to DSE 4, whereas several adverse conditions place stronger emphasis on DSE 3 or DSE 5. Importantly, the routing does not exhibit a one-to-one correspondence between semantic weather categories and individual experts: the same expert can be shared across multiple conditions, while each condition activates a weighted combination of experts. This behavior is consistent with DOME organizing expert specialization according to latent distortion patterns rather than explicitly assigning one expert to each weather category.

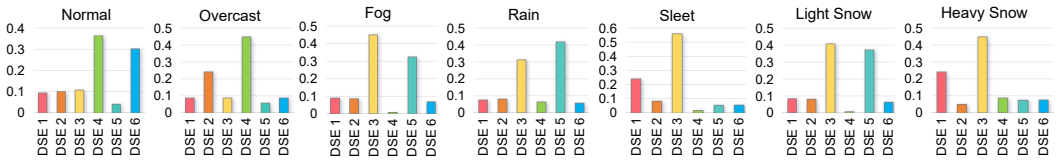


Figure 6: Average expert routing distributions of DOME across different weather conditions on K-Radar. Different conditions exhibit distinct expert-utilization patterns, while multiple conditions may share experts with different routing weights.

B.1.2 PDR Component Ablation

To examine the two complementary cues within PDR, we retain only the global-context or local-discrepancy branch while keeping the PDR formulation and all other settings unchanged. As shown in Table 8, *Global-only* and *Local-only* achieve 59.7 and 58.8 AP_{3D}, respectively, whereas the full PDR reaches 61.6. These results show that global multimodal context and local LiDAR–Radar discrepancy provide complementary information for distortion-sensitive routing.

Table 8: Ablation of the PDR design on K-Radar at IoU=0.5.

Router	Global	Local	AP_{3D}
L4DR (baseline)	–	–	53.5
Global-only	✓		59.7
Local-only		✓	58.8
PDR	✓	✓	**61.6**

B.1.3 Quantitative Continuous Distortion Response

We further quantify whether PDR responds continuously to changes in physical degradation. Increasing fog intensity is applied to the same Normal scenes while keeping scene content, Radar observations, and annotations unchanged. We measure the Jensen–Shannon divergence (JSD) between the routing distribution of each fog-degraded sample and that of its clean counterpart. As shown in Table 9, the routing divergence increases monotonically from 0.574 to 1.379 as fog severity increases. Since scene semantics and Radar observations remain fixed, this controlled trend provides evidence that the learned routing representation is sensitive to progressive sensor degradation rather than only to discrete weather identities or scene-specific content.

Table 9: Routing divergence under continuously increasing simulated fog severity.

σ_{fog}	JSD to Normal ($\times 100$)
0.005	0.574
0.010	0.966
0.020	1.379

B.1.4 Physical Interpretation of Distortion Fingerprints

To further investigate the spatial cues underlying distortion-driven routing, Fig. 7 presents PDR attribution maps across all weather conditions. Normal and Overcast exhibit relatively smooth, distance-dependent responses, whereas adverse conditions produce more irregular and spatially clustered patterns. Moreover, the locations and structures of these responses differ across Fog, Rain, Sleet, Light Snow, and Heavy Snow, suggesting that PDR is sensitive to heterogeneous spatial manifestations of sensor degradation.

We next quantify the spatial concentration of the learned attribution. As reported in Table 10, foreground regions occupy only 1.3–3.1% of the BEV grid but receive 4.7–9.8% of the total routing attribution, corresponding to a 1.78–4.97× enrichment. Similarly, the near field within 20 m covers 27.31% of the spatial grid but contributes 47.20–57.26% of the attribution, yielding a 1.73–2.10× enrichment. These results show that PDR concentrates disproportionately on informative foreground and near-field regions rather than distributing its response uniformly across the scene.

Table 10: Spatial enrichment statistics of PDR routing attribution.

Region	Grid area	Attribution	Enrichment
Foreground	1.3–3.1%	4.7–9.8%	1.78–4.97×
Near field (0–20 m)	27.31%	47.20–57.26%	1.73–2.10×

To relate this near-field preference to observable sensor characteristics, Fig. 8 reports the normalized near-to-far LiDAR point-density ratio under different weather conditions. Compared with Normal

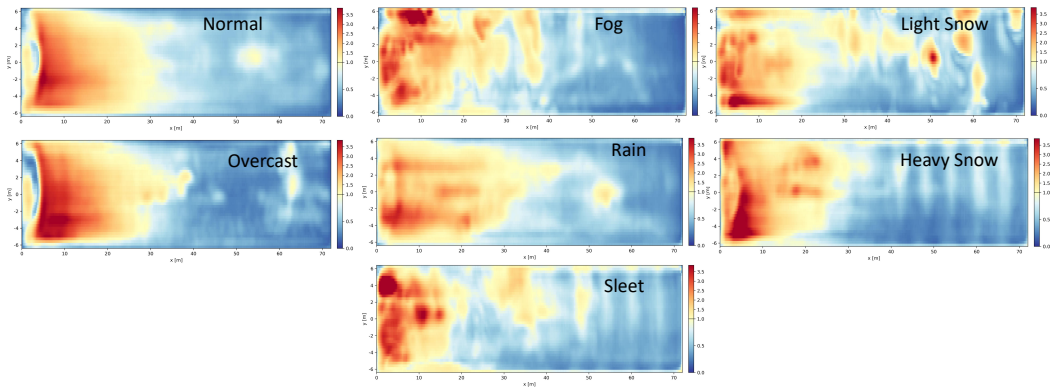


Figure 7: Spatial attribution maps of PDR across different weather conditions. Normal and Overcast exhibit relatively smooth distance-dependent responses, whereas adverse conditions produce more irregular and spatially clustered attribution patterns.

weather, adverse conditions exhibit stronger near-field concentration, indicating weather-associated changes in the spatial distribution of LiDAR observations. The consistency between this density shift and the learned attribution provides additional evidence that PDR responds to physically meaningful sensor-degradation patterns.

Taken together, the expert-utilization patterns, PDR component ablation, continuous-severity response, and spatial analyses provide complementary evidence for the learned distortion fingerprints. PDR benefits from complementary global and local cues, varies continuously with controlled degradation, and assigns increased importance to spatial regions whose LiDAR statistics are affected under adverse weather.

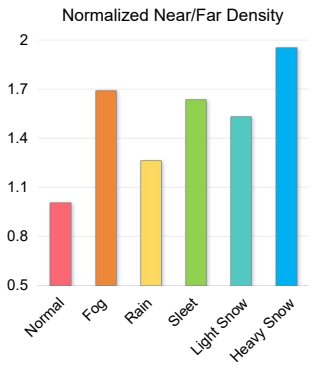


Figure 8: Normalized near-to-far LiDAR density ratios across weather conditions.

B.2 Additional Robustness and Statistical Stability

Beyond the controlled degradation experiments in the main paper, we further evaluate DOME under impaired visual observations, individual sensor perturbations, and different random initializations. These experiments examine robustness to additional sensing disturbances and the statistical stability of the reported performance.

B.2.1 Robustness under Vision Impairment

Table 11: Performance comparison between DOME and L4DR-DA3D under different vision impairment probabilities (VIP). AP_{3D} is reported at IoU=0.5.

Method	Modality	VIP	Total	Normal	Overcast	Fog	Rain	Sleet	Light Snow	Heavy Snow
L4DR-DA3D (Yang et al., 2025)	L+4DR+C	0.00	61.9	58.9	66.4	79.2	63.0	48.2	64.6	47.6
		0.25	60.5 (-1.4)	57.6 (-1.3)	65.3 (-1.1)	76.6 (-2.6)	62.2 (-0.8)	47.0 (-1.2)	61.7 (-2.9)	45.0 (-2.6)
		0.50	59.2 (-2.7)	56.5 (-2.4)	64.5 (-1.9)	74.1 (-5.1)	61.0 (-2.0)	46.5 (-1.7)	60.3 (-4.3)	42.8 (-4.8)
DOME (Ours)	L+4DR	-	61.6	60.5	75.6	85.3	58.2	48.4	66.7	47.6

Existing weather-aware adaptation methods may rely on camera-based weather recognition to determine weather-specific processing strategies. To examine their sensitivity to degraded visual observations, we compare DOME with the image-dependent L4DR-DA3D (Yang et al., 2025) under different vision impairment probabilities (VIP). Specifically, with probability VIP, the camera observation is impaired by either strong additive noise or random input dropout to simulate unreliable visual sensing. As shown in Table 11, the overall AP_{3D} of L4DR-DA3D decreases from 61.9 to 59.2 as VIP increases from 0 to 0.50, with consistent degradation across most weather conditions. In contrast, DOME performs routing directly from LiDAR–4D Radar distortion representations and

therefore does not require camera-based weather recognition. This comparison highlights the robustness advantage of avoiding an external visual weather-routing dependency when camera observations become unreliable.

B.2.2 Robustness under Sensor Degradation

We further test the clean-trained DOME model under individual LiDAR and Radar perturbations without additional retraining. Specifically, we apply a 20% random point dropout to LiDAR and perturb Radar point coordinates with random jitter within one voxel. As shown in Table 12, randomly dropping 20% of LiDAR points reduces AP_{3D} from 61.6 to 60.4, while Radar coordinate jitter results in 61.1 AP_{3D}. The limited degradation under both perturbations indicates that DOME retains effective multi-modal perception when one sensing stream becomes moderately corrupted, rather than depending exclusively on a single modality.

Table 12: Robustness of DOME under sensor perturbations at IoU=0.5.

Condition	AP_{3D}
Clean	61.6
LiDAR dropout (20%)	60.4
Radar coordinate jitter	61.1

B.2.3 Stability across Random Seeds

To evaluate statistical stability, we train DOME with five different random seeds while keeping the dataset split, model configuration, and optimization settings unchanged. As shown in Table 13, DOME achieves 87.12±0.70/83.62±0.52 AP_{BEV}/AP_{3D} at IoU=0.3 and 82.96±0.36/61.46±0.55 at IoU=0.5. The relatively small standard deviations across all four metrics indicate that the performance of DOME is stable across different random initializations.

Table 13: Performance of DOME across different random seeds on K-Radar. All entries report Total AP (%).

Seed	IoU = 0.3		IoU = 0.5	
	AP_{BEV}	AP_{3D}	AP_{BEV}	AP_{3D}
0	86.1	82.9	82.7	61.0
1	87.6	83.7	83.1	61.8
2	87.9	84.3	83.5	62.2
3	86.8	83.4	82.6	60.9
4	87.2	83.8	82.9	61.4
Mean ± Std.	87.12±0.70	83.62±0.52	82.96±0.36	61.46±0.55

B.3 Architectural Scalability

To examine whether the proposed distortion-driven framework is tied to a specific detection backbone, we further integrate DOME and DOME-L with InterFusion (Wang et al., 2022). The original InterFusion fusion pipeline and detection architecture are retained, while the proposed distortion-aware routing and expert adaptation mechanisms are incorporated into its multimodal representation.

B.3.1 DOME on InterFusion

We first evaluate the full DOME design on InterFusion. As shown in Table 14, integrating DOME improves the overall AP_{BEV} from 66.1 to 75.8 and AP_{3D} from 41.7 to 53.2 at IoU=0.5. Consistent improvements are also observed across most weather conditions, indicating that the proposed DOME framework can be effectively transferred beyond the L4DR backbone to a different LiDAR–4D Radar fusion architecture.

Table 14: Performance of InterFusion with and without DOME integration on K-Radar at IoU=0.5.

Method	Modality	Metric	Total	Normal	Overcast	Fog	Rain	Sleet	Light Snow	Heavy Snow
InterFusion (Wang et al., 2022)	L+4DR	AP_{BEV}	66.1	70.5	82.0	81.8	67.2	33.9	62.9	46.0
		AP_{3D}	41.7	44.6	53.5	64.8	37.2	25.5	35.4	27.0
InterFusion-DOME	L+4DR	AP_{BEV}	**75.8**	**81.9**	**88.3**	**88.6**	**74.2**	**49.5**	**79.1**	**48.3**
		AP_{3D}	**53.2**	**57.7**	**55.7**	**73.2**	**52.5**	**30.9**	**59.7**	**29.7**

B.3.2 PARAMETER-EFFICIENT DOME-L ON INTERFUSION

We further examine whether the parameter-efficient LDE design remains effective when transferred to InterFusion. As shown in Table 15, InterFusion-DOME-L achieves 52.5 AP_{3D} while reducing the parameter count from 20.88M to 8.59M and FLOPs from 41.95G to 21.43G relative to the full InterFusion-DOME. It also retains nearly the original InterFusion inference speed. These results show that the lightweight distortion-expert design generalizes to a different multimodal backbone while substantially reducing the overhead of full experts.

Table 15: Accuracy–efficiency comparison of DOME variants on InterFusion. AP_{3D} is reported at IoU=0.5.

Method	AP_{3D}	FPS [Hz]	FLOPs [G]	Params [M]
InterFusion (Wang et al., 2022)	41.7	31.64	16.72	4.40
InterFusion-DOME	**53.2**	30.82	41.95	20.88
InterFusion-DOME-L	**52.5**	31.17	21.43	**8.59**

B.4 ADDITIONAL ANALYSIS OF PARAMETER-EFFICIENT DOME-L

B.4.1 FULL WEATHER RESULTS OF DOME-L

To further examine whether the parameter-efficient DOME-L preserves distortion-aware adaptation across different weather conditions, we report its complete K-Radar results at IoU=0.5. As shown in Table 16, DOME-L improves the overall AP_{BEV}/AP_{3D} from 77.5/53.5 to 83.5/61.3 over the L4DR baseline. DOME-L consistently improves both AP_{BEV} and AP_{3D} across all reported weather conditions, indicating that the lightweight LoRA-based distortion experts preserve effective distortion-aware adaptation while substantially reducing the parameter overhead of full experts.

Table 16: Performance of DOME-L on K-Radar across different weather conditions at IoU=0.5.

Method	Modality	Metric	Total	Normal	Overcast	Fog	Rain	Sleet	Light Snow	Heavy Snow
L4DR (Huang et al., 2025)	L+4DR	AP_{BEV}	77.5	76.8	88.6	89.7	78.2	59.3	80.9	53.8
		AP_{3D}	53.5	53.0	64.1	73.2	53.8	46.2	52.4	37.0
DOME-L (Ours)	L+4DR	AP_{BEV}	**83.5**	**81.7**	**92.2**	**96.8**	**84.7**	**69.7**	**94.7**	**67.5**
		AP_{3D}	**61.3**	**60.5**	**70.1**	**81.9**	**62.6**	**48.4**	**63.3**	**50.8**

B.5 TRAINING STRATEGY AND HYPER-PARAMETER ANALYSIS

We further analyze the optimization strategy and major hyper-parameters of the proposed framework. We first examine the progressive contribution of the training components in Progressive Expert Maturation (PEM), and then study the sensitivity to architectural and optimization hyper-parameters.

B.5.1 ABLATION OF PROGRESSIVE EXPERT MATURATION

We analyze the progressive contribution of the optimization components while keeping the PDR+DSE architecture fixed. As shown in Table 17, introducing Adaptive Logit Exploration (ALE) improves AP_{3D} from 58.3 to 60.0. Further incorporating the Temperature Annealing Scheme (TAS) and Asymptotic Backpropagation Scaling (ABS) increases performance to 61.6. These results show that the progressive optimization strategy provides additional gains beyond the distortion-aware expert architecture itself, with TAS and ABS jointly complementing ALE during expert maturation.

Table 17: Ablation of the optimization components in PEM on K-Radar. AP_{3D} is reported at IoU=0.5.

Configuration	ALE	TAS+ABS	AP_{3D}
PDR+DSE			58.3
PDR+DSE	✓		60.0
PDR+DSE	✓	✓	**61.6**

B.5.2 HYPER-PARAMETER SENSITIVITY

To examine the sensitivity of the proposed framework to its architectural and optimization choices, we vary the number of distortion experts N_{exp}, the Top-K activation value, the LoRA rank r in DOME-L, and the PEM-related parameters E_{trans}, τ_{start}, and λ_{start}. As shown in Fig. 9, the proposed framework remains effective over a range of configurations, while the selected settings

provide the strongest overall performance. Based on these experiments, we use $N_{exp} = 6$, $K = 2$, $E_{trans} = 5$, $\tau_{start} = 2.0$, and $\lambda_{start} = 0.1$ throughout the main experiments, and set the LoRA rank to $r = 8$ for DOME-L.

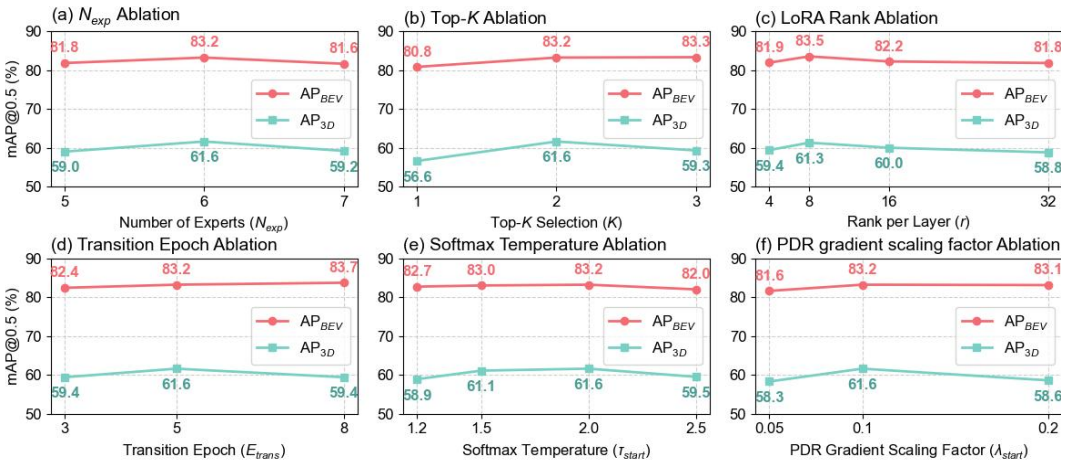


Figure 9: Hyper-parameter sensitivity of DOME and DOME-L on K-Radar. We vary the number of distortion experts N_{exp}, Top-K activation value, LoRA rank r in DOME-L, transition epochs E_{trans}, initial routing temperature τ_{start}, and initial PDR gradient scaling factor λ_{start}.

B.6 Augmentation-Matched Comparison with AW-MoE

Table 18: Augmentation-matched comparison between AW-MoE and DOME on K-Radar at IoU=0.5. Both methods use the Unified Dual-Modal Augmentation (UDMA) proposed by AW-MoE (Lin et al., 2026), including Weather-Specific GT Sampling (WSGTS). Best results are shown in **bold**.

Method	Modality	Metric	Total	Normal	Overcast	Fog	Rain	Sleet	Light Snow	Heavy Snow
AW-MoE (arXiv 2026)	L+4DR+C	AP_{BEV}	84.2	82.8	91.6	96.3	**85.3**	**75.0**	94.7	66.4
		AP_{3D}	61.5	59.0	67.2	85.7	**63.5**	43.3	70.1	**53.1**
DOME (Ours)	L+4DR	AP_{BEV}	**85.0**	**84.6**	**94.3**	**97.4**	84.6	72.9	**95.2**	**68.7**
		AP_{3D}	**62.7**	**61.5**	**75.5**	**86.0**	62.1	**49.3**	**70.8**	51.3

To provide an augmentation-matched comparison with AW-MoE (Lin et al., 2026), we further evaluate DOME using the Unified Dual-Modal Augmentation (UDMA) proposed by AW-MoE, which applies synchronized flipping, rotation, scaling, and ground-truth (GT) sampling to LiDAR and 4D Radar. Its Weather-Specific GT Sampling (WSGTS) restricts sampled objects to scenes with the same weather condition as the target scene, reducing cross-weather mismatches in sensor characteristics.

Because WSGTS inserts objects from different scenes, it may alter the local LiDAR–4D Radar discrepancies used by PDR for routing. We therefore first train DOME under its standard setting and record the PDR routing scores for the original training samples. The detector is then fine-tuned with UDMA while reusing these routing scores for expert activation and aggregation. This preserves routing based on the original scene-level distortion cues while allowing the detector to benefit from the additional GT samples introduced by UDMA.

As shown in Table 18, under the matched augmentation setting, DOME achieves 85.0/62.7 AP_{BEV}/AP_{3D}, compared with 84.2/61.5 for AW-MoE. Thus, DOME retains higher overall performance under the same augmentation protocol. Table 1 reports DOME under its standard training setting, while this experiment provides the augmentation-matched comparison with AW-MoE.

C VISUALIZATION COMPARISON

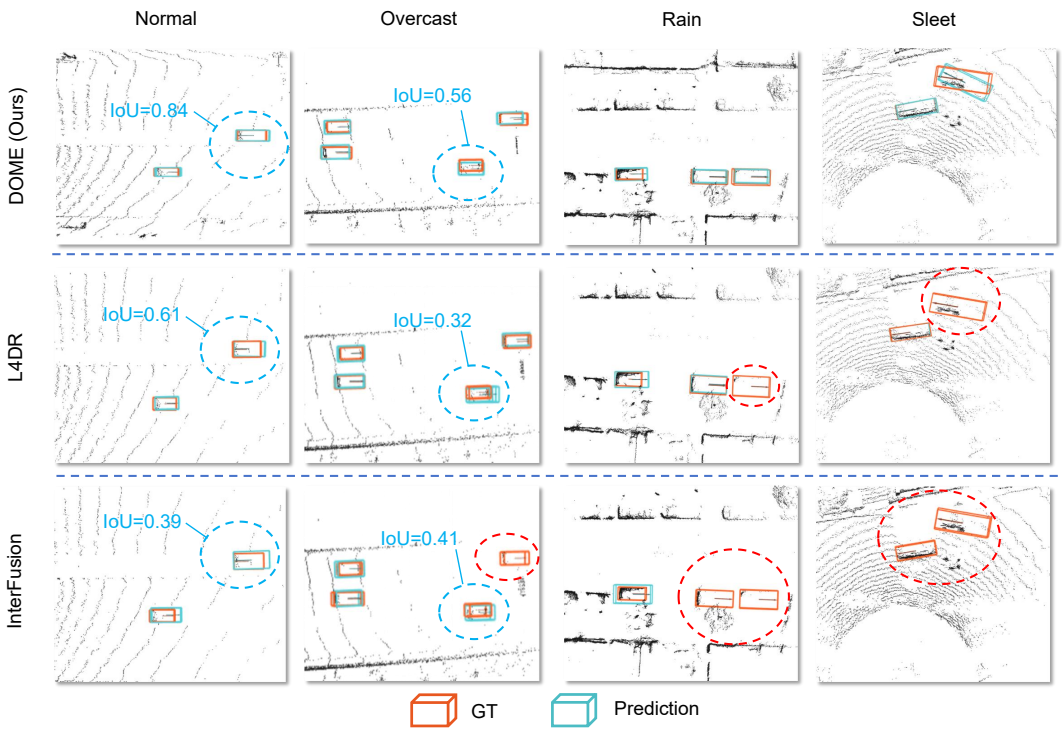


Figure 10: Qualitative comparison of DOME, L4DR (Huang et al., 2025), and InterFusion (Wang et al., 2022) under Normal, Overcast, Rain, and Sleet weather conditions.

To provide a more intuitive qualitative assessment, we compare DOME with the L4DR (Huang et al., 2025) and InterFusion (Wang et al., 2022) baselines under diverse weather conditions in Fig. 10 and Fig. 11. The visualizations reveal two recurring qualitative patterns. First, DOME reduces missed detections frequently observed under adverse weather conditions, as highlighted by the red circles. Second, its predicted 3D bounding boxes exhibit improved spatial alignment with the ground truth, as highlighted by the blue circles. These qualitative results further illustrate the robustness of DOME across diverse weather conditions and complement the quantitative comparisons reported in the main paper.

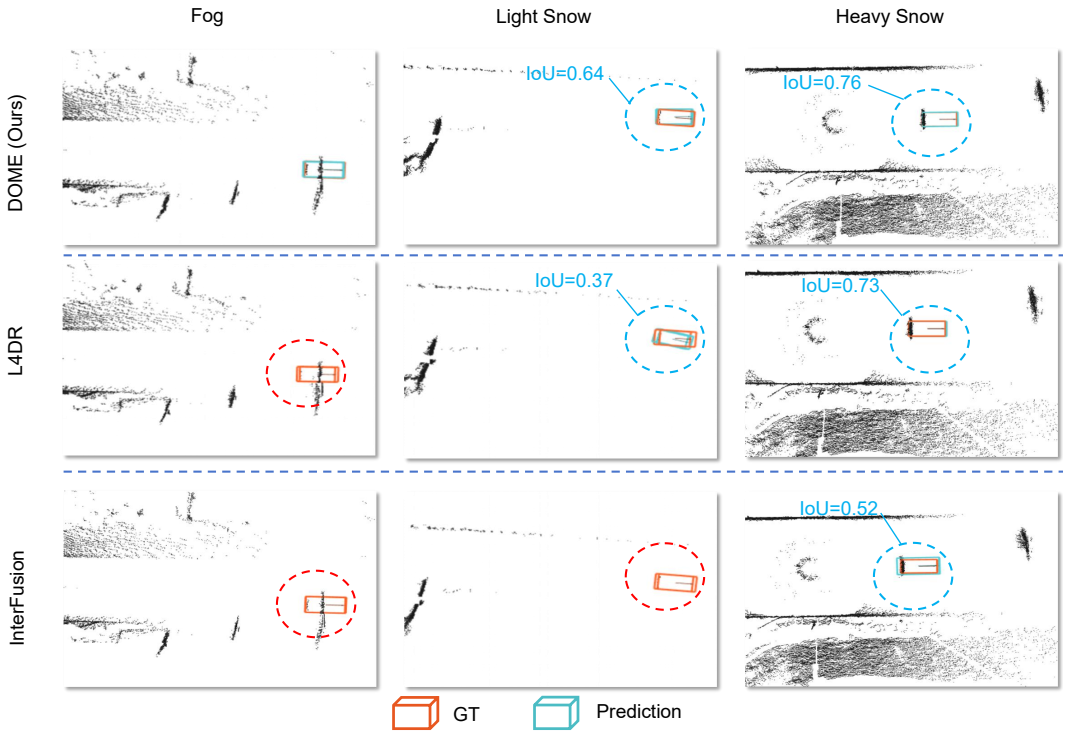


Figure 11: Qualitative comparison of DOME, L4DR (Huang et al., 2025), and InterFusion (Wang et al., 2022) under Fog, Light Snow, and Heavy Snow weather conditions.

D LIMITATIONS

While effective, DOME utilizes identical DSE architectures, overlooking the inherent variance in domain complexity. This structural homogeneity can lead to parameter redundancy in simpler domains while creating representation bottlenecks in more challenging ones. Future research could explore heterogeneous parameter allocation, adaptively scaling expert capacity to optimize the accuracy-efficiency trade-off within a fixed parameter budget.