Structure-to-Intensity Diffusion for Adverse-Weather LiDAR Generation

Peiyang Ni[1] Longyu Yang[1] Lu Zhang[2] Kuniaki Saito[3] Yap-Peng Tan[4]
Fumin Shen[1] Heng Tao Shen[5] Xiaofeng Zhu[1,6] Ping Hu[1,4*]
[1]University of Electronic Science and Technology of China [2]Dalian University of Technology
[3]OMRON SINIC X [4]VinUniversity [5]Tongji University [6]Hainan University

Abstract

Adverse-weather LiDAR point cloud generation is challenged by complex weather-induced degradations. These degradations affect geometry and reflectance in fundamentally different ways, making joint modeling difficult and ambiguous, especially when diverse real-world training data is limited. To address this, we propose ***Structure-to-Intensity Diffusion*** *(SiD), a diffusion-based framework that explicitly factorizes the denoising process at each time step: it first reconstructs the geometric structure, then conditions reflectance intensity denoising on the estimated structure. This structure-conditioned design decomposes the joint distribution, reduces modeling ambiguity, and leads to point clouds that are both geometrically coherent and radiometrically realistic. To mitigate data scarcity, we introduce* ***Real-Prior Weather Simulation*** *(RPWS), a degradation module that leverages real-world sensor statistics to synthesize physically plausible adverse-weather point clouds from clear scans. Extensive experiments demonstrate that, with similar model complexity, our approach outperforms the previous state-of-the-art in generating adverse-weather LiDAR scans with both structural and radiometric properties more closely aligned with real-world data.*

1. Introduction

Light Detection and Ranging (LiDAR) is important for autonomous systems and 3D world modeling [9, 22], offering precise structural data that is robust to lighting changes [10]. However, adverse weather, such as rain, snow, and fog, can degrade both geometry and reflectance, leading to incomplete or noisy point clouds [7, 15], which undermines perception reliability and safety. While weather-robust perception has advanced [21, 24, 47, 50], real-world adverse-weather LiDAR remains scarce due to the difficulty of data collection. Synthesizing diverse weather-degraded LiDAR is thus a scalable solution for tackling the challenge.

*Corresponding author.

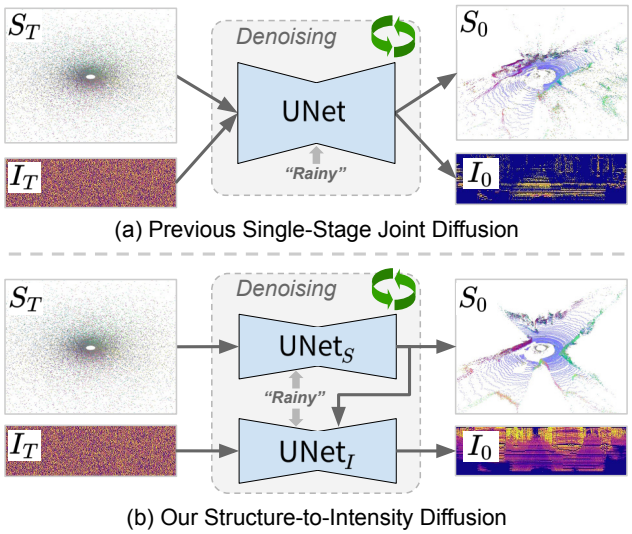


Figure 1. Conceptual comparison of generative architectures for adverse-weather LiDAR. (a) Previous state-of-the-art [45] jointly denoises structure S and intensity I via a unified network. Under complex degradations and limited real-world training data, geometric cues may dominate, making it challenging to capture reflectance details. (b) Our Structure-to-Intensity Diffusion framework explicitly decouples the denoising, resolving the modeling ambiguity, and yields more realistic adverse-weather scans.

Recent generative methods, especially diffusion-based models [8, 11, 17, 35, 37, 49, 52, 54], demonstrate strong performance in producing realistic LiDAR point clouds. By leveraging denoising diffusion processes to recover fine-grained spatial structures, they set new standards for LiDAR generation. Yet, most are designed under the clear-weather assumption and ignore the degradations under adverse weather.

To extend LiDAR generation beyond clear-weather assumptions, recent studies have investigated generation under adverse environments. Simulation-based approaches [12, 13, 20, 30] apply explicit physical models to mimic weather effects on LiDAR returns. While these models offer interpretability, they often rely on oversimplified assumptions about laser propagation and environmental

interactions, leading to limited realism and diversity. More recent learning-based methods, such as WeatherGen [45], combine simulation with data-driven learning to enhance fidelity. However, these approaches typically model geometry and reflectance jointly, without accounting for their fundamentally different degradation patterns [51]. This increases the optimization difficulty, as the model must handle distinct degradation behaviors. Consequently, the training process can be implicitly dominated by the more stable geometric cues, while the reflectance signals are underrepresented, as illustrated in Fig. 1(a).

To better address the imbalance between geometry and reflectance modeling, we revisit their causal relationship in LiDAR sensing. Geometry encodes the spatial structure of a scene, while reflectance intensity captures how laser pulses interact with surfaces, shaped by factors such as range, surface orientation, material properties, and atmospheric conditions. As a result, while not entirely determined by geometry, reflectance is strongly influenced by it [5, 19, 23]. In contrast, inferring geometry from intensity is fundamentally ill-posed, especially under adverse weather, where degradation introduces significant noise and ambiguity [3, 38]. This asymmetric dependency provides a valuable inductive bias that can simplify the generative modeling process. Instead of directly learning the joint distribution of geometry and intensity, we formulate the task as a conditional generation problem, explicitly modeling reflectance conditioned on geometry. We instantiate this idea through *Structure-to-Intensity Diffusion* (SiD) as shown in Fig. 1(b), a factorized diffusion framework in which each denoising step first reconstructs the geometric structure, then denoises the reflectance intensity conditioned on the recovered geometry. This design mirrors the physical formation of LiDAR measurements, reduces ambiguity and enables better learning of the distinct degradation behaviors of each aspect.

Another significant challenge lies in the scarcity of diverse and high-quality training data under real-world adverse weather. Existing datasets provide limited coverage or rely on coarse synthetic degradation, which hampers generalization. To overcome this, we propose *Real-Prior Weather Simulation* (RPWS), a parametric degradation module that synthesizes adverse-weather LiDAR from clear-weather scans by leveraging empirical sensor statistics derived from real-world adverse-weather observations. By capturing depth-dependent attenuation and structured occlusion patterns, along with reflectance distortions, RPWS enables physically grounded and scalable data augmentation that aligns with real sensing characteristics. This pipeline allows SiD framework to generalize effectively across diverse weather conditions in a unified framework.

In addition, existing evaluation metrics for LiDAR generation are often trained or calibrated on clear-weather data, making them less reliable under severe degradation. To address this, we introduce the *Fréchet Minkowski Distance* (FMD), a voxel-level metric built upon features extracted with a Minkowski UNet trained via the pretraining framework SLidR [41] on adverse-weather data. By combining real adverse-weather LiDAR grounding with rich semantic priors from large-scale pretrained representations [4], FMD yields a more robust and perceptually aligned evaluation of generation quality. In summary, our main contributions are:

- We analyze the asymmetric structure-intensity dependency in LiDAR sensing and introduce a factorized diffusion framework that decouples the generation.
- We propose *Structure-to-Intensity Diffusion* (SiD), a factorized diffusion model with structure-aware conditioning, combined with *Real-Prior Weather Simulation* (RPWS), which uses real-world scan statistics to synthesize realistic adverse-weather data.
- We conduct extensive experiments, and the results demonstrate that our method outperforms prior work and remains effective even with limited real training data.

2. Related Works

Generative models have recently emerged as a powerful paradigm for synthesizing realistic LiDAR point clouds. Early efforts relied on generative adversarial learning frameworks [40] and neural radiance field extensions [28], as demonstrated by methods like DUSty [31, 33]. With the advancement of diffusion models [16, 34, 42], diffusion-based approaches have achieved state-of-the-art performance in generating high-fidelity and semantically coherent LiDAR data [8, 11, 32, 35, 37, 44, 49, 52, 54]. LiDARGen [54] formulates generation as a stochastic denoising process over range images, while R2DM [32] performs sparse-to-dense completion in range–reflectance space using DDPM. Text2LiDAR [44] conditions the generation process on language inputs via an equirectangular transformer, and LiDM [37] leverages latent diffusion with curve-wise encoding for scalable synthesis. Beyond diffusion, UltraLiDAR [48] introduces a discrete codebook and an autoregressive transformer for sequence-style point cloud generation, while HoloDrive [46] scales up to street-level multi-modal scene synthesis with temporal modeling. However, these methods are predominantly developed and evaluated on clear-weather datasets [2, 25, 44], thus ignoring the complex degradations induced by adverse weather.

Adverse-weather synthesis has been approached via physics-based and learning-based methods. Simulation pipelines like FSRL [12], LSS [13], and LiSA [20] apply atmospheric optics (*e.g.*, Beer–Lambert law, Mie scattering) to simulate fog, snow, and rain. While interpretable, these models are typically tailored to specific conditions and fail to generalize to the diversity seen in real data. Learning-based alternatives, such as WeatherGen [45], combine physics-aware simulation with diffusion models to capture

adverse-weather variations. However, most existing methods jointly model geometry and reflectance, which may limit learning effectiveness under limited real training data.

3. Methodology

In this section, we introduce the Structure-to-Intensity Diffusion (SiD) framework for generating LiDAR scans under adverse weather, as illustrated in Fig. 2. We begin in Sec. 3.1 with a brief overview of denoising diffusion probabilistic models (DDPMs) [16, 27]. Next, Sec. 3.2 presents our Real-Prior Weather Simulation (RPWS) module, which enriches training data by synthesizing realistic degradations over clear LiDAR scans with real-world statistics. Finally, Sec. 3.3 details the SiD framework, which factorizes the generative process by first denoising geometric structure and then conditionally generating reflectance intensity, alleviating the optimization complexity and improving generation quality for adverse conditions.

3.1. Preliminaries on Diffusion Models

Denoising diffusion probabilistic models (DDPMs) [16, 27] are a class of generative models that learn to synthesize data by reversing a gradual noising process. The forward diffusion process progressively corrupts a clean data sample $\mathbf{x}_0$ into pure noise $\mathbf{x}_T$ over T timesteps by adding Gaussian noise:

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

where $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$ is determined by the noise schedule.

To generate new samples, DDPMs learn a reverse process that denoises $\mathbf{x}_T$ step by step back to $\mathbf{x}_0$. Rather than modeling the full transition distribution $q(\mathbf{x}_{t-1} \mid \mathbf{x}_t)$, the model is trained to predict the noise $\boldsymbol{\epsilon}$ added at each step, by minimizing $\mathcal{L} = \mathbb{E}_{t,\mathbf{x}_0,\boldsymbol{\epsilon}}\left[\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\|_2^2\right]$, where $\boldsymbol{\epsilon}_\theta(\cdot)$ is a neural network and t is the diffusion timestep.

This noise-prediction formulation enables effective training and high-fidelity synthesis. We build on it for adverse-weather LiDAR by adopting a factorized reverse process that first denoises geometry structure and then reflectance intensity, improving learning efficiency and robustness under limited training data.

3.2. Real-Prior Weather Simulation

A major challenge in training LiDAR generation models for adverse weather lies in the scarcity of real-world data. Capturing LiDAR scans under fog, rain, or snow is logistically demanding and costly [1, 20], making large-scale collection difficult. Existing simulation approaches [12, 13, 20, 45] often rely on simplified physics or handcrafted noise models that may fail to capture the statistical complexity of real-world degradations. To bridge this gap, we propose *Real-Prior Weather Simulation* (RPWS), a parametric augmentation module that synthesizes adverse-weather LiDAR from clean scans using degradation statistics derived from real data. RPWS separately models geometric and reflectance distortions, guided by empirical distributions extracted from real adverse-weather observations.

Geometric Structure Degradation. Adverse weather affects LiDAR geometry mainly through two mechanisms: (1) *range-dependent attenuation*, which reduces point density with distance, and (2) *structured occlusions* aligned with the scan pattern, caused by large atmospheric particles (*e.g.*, fog droplets, raindrops, snowflakes) that block laser returns. While scattering and partial reflections may introduce random geometric noise, RPWS focuses on the dominant attenuation and occlusion effects observed in real data.

Given a clear-weather training scan, we randomly sample an adverse scan from the training set and compute a depth-wise decay profile $P_{\text{dec}}(r)$ from this pair. The range dimension is discretized into bins $[r_k, r_{k+1})$, and the decay rate for a point x in the clean scan is computed as:

$$P_{\text{dec}}(x) = \frac{N_{\text{adver}}(r(x))}{\max(N_{\text{adver}}(r(x)), N_{\text{clear}}(r(x)))}, \tag{2}$$

where $r(x)$ denotes the range bin of x, and N_{clear}, N_{adver} are the counts of valid returns in that bin for the clear and adverse scans, respectively. To capture structured occlusions, we modulate this decay with a spatial mask $S(x) \in [0, 1]$ derived from the sampled adverse scan's range-view representation, encoding empirical point-wise visibility. The final survival probability for each point in the clean scan is given by:

$$z(x) \sim \text{Bernoulli}\big(P_{\text{dec}}(x) \cdot S(x)\big). \tag{3}$$

This formulation transfers the adverse scan's attenuation and occlusion statistics to the clean scan, producing stochastic yet physically grounded geometric degradations.

Reflectance Intensity Degradation. Adverse weather also distorts LiDAR reflectance through complex, signal-dependent effects that are poorly captured by additive noise. Given the same clear/adverse pair, we learn a transformation $T : \mathbb{R} \to \mathbb{R}$ that maps clean reflectance values to an adverse-weather distribution. We compute empirical reflectance histograms P_{clear} and P_{adver}, and estimate an optimal transport map T^* by minimizing the Wasserstein distance:

$$T^* = \arg\min_T D\big(P_{\text{clear}} \circ T^{-1},\, P_{\text{adver}}\big), \tag{4}$$

subject to the push-forward constraint $T_{\#}P_{\text{clear}} = P_{\text{adver}}$. The resulting map T^* captures weather-specific shifts in the reflectance distribution and is used to transform the clean scan's reflectance channel, producing radiometrically consistent adverse-weather simulations.

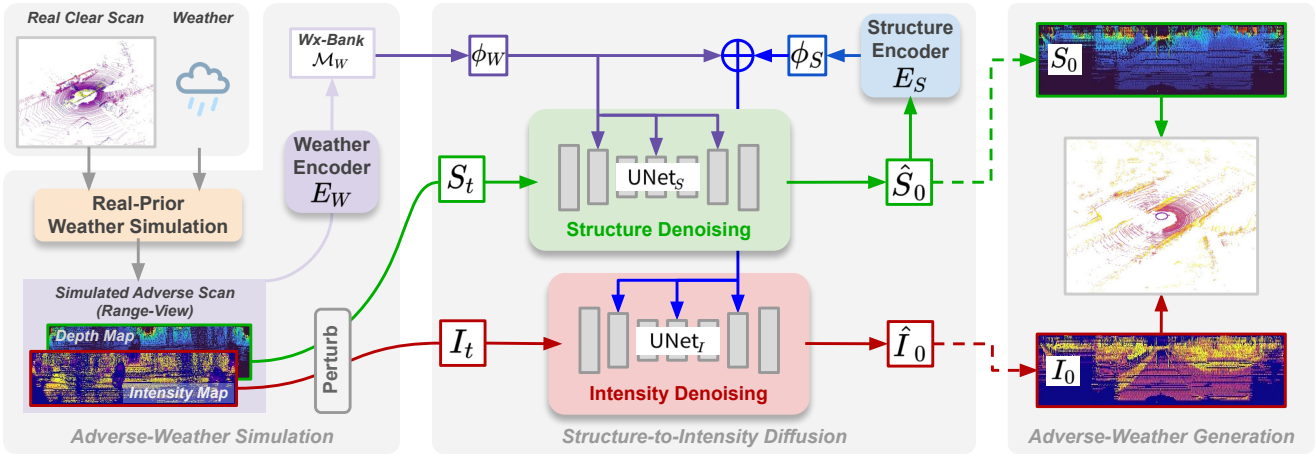


Figure 2. Overview of the proposed Structure-to-Intensity Diffusion (SiD) framework for adverse-weather LiDAR generation. We begin by enriching training data using the Real-Prior Weather Simulation (RPWS) module, which synthesizes realistic adverse-weather degradations from clear LiDAR scans based on real-world priors. The resulting scan is processed by a weather encoder to extract a compact embedding ϕ_W, which is added into a weather bank $\mathcal{M}_W$ for future sampling. During **SiD training**, raw training scans are projected into range images and corrupted into noisy states (S_t, I_t). Conditioned on the weather embedding ϕ_W, a structure denoiser UNet_S reconstructs the geometry $\hat{S}_0$ and encodes it into a structure embedding ϕ_S. An intensity denoiser UNet_I then generates reflectance conditioned on both ϕ_W and ϕ_S, following the dependency order $S \rightarrow I$. During **SiD sampling**, the embedding ϕ_W for target weather is retrieved from the weather bank. Starting from Gaussian noise (S_T, I_T), the model first denoises the geometry structure and then the reflectance intensity at each timestep, ultimately yielding the final adverse-weather LiDAR scan.

In summary, during each training iteration, RPWS takes a clean scan and randomly samples an adverse scan to extract structure- and reflectance-level degradation statistics. These are then applied directly to the clean scan, generating diverse and physically faithful adverse-weather augmentations without introducing additional learnable parameters.

3.3. Structure-to-Intensity Diffusion

The RPWS module in the previous subsection underscores a key observation: geometric structure and reflectance intensity respond differently to adverse weather, exhibiting distinct degradation patterns. This disparity suggests that directly modeling their joint distribution $\mathcal{P}(S, I \mid c)$, where S and I represent geometry and reflectance, and c denotes the conditioning weather, is inherently difficult, especially under limited real-world supervision. In practice, the stronger geometric cues often dominate the learning process, while the weaker and noisier reflectance signals tend to be underfitted or physically inconsistent [18, 19]. To mitigate this imbalance, we introduce a structure-to-intensity generation framework that explicitly decouples the learning of S and I by exploiting their asymmetric dependency.

We build on the key insight that LiDAR sensing exhibits an asymmetric causal relationship between geometry and intensity. Intensity I is a radiometric response determined by the underlying geometry S and contextual factors (*e.g.*, range, incidence angle, surface material, and atmospheric attenuation). While I is influenced by other factors and remains ambiguous, it is nonetheless conditioned on the spatial structure S, establishing an asymmetric dependency. Based on this observation, we formulate the generative process using a causal factorization:

$$\mathcal{P}(S, I \mid c) = \mathcal{P}(S \mid c)\, \mathcal{P}(I \mid S, c), \tag{5}$$

where S and I denote the geometry structure and reflectance intensity, and c represents the conditioning context (*e.g.*, weather tag). This decomposition mirrors the underlying dependency in LiDAR sensing and simplifies the learning objective by separating two conditionally related distributions.

To implement this factorization, we adopt a denoising diffusion probabilistic model (DDPM) as our generative backbone. We design a factorized reverse process that first denoises the geometry structure and then generates the reflectance intensity conditioned on the recovered structure. In particular, at each reverse timestep t, the transition distribution can be formulated as:

$$\begin{aligned} &p_\theta(S_{t-1}, I_{t-1} \mid S_t, I_t, c) \\ &\quad = p_\theta(S_{t-1} \mid S_t, c) \cdot p_\theta(I_{t-1} \mid I_t, S_t, c) \\ &\quad = p_\theta(S_{t-1} \mid S_t, c) \cdot p_\theta(I_{t-1} \mid I_t, \hat{S}_0, c). \end{aligned} \tag{6}$$

The second line reflects a substitution in which the conditioning on S_t in the intensity branch is replaced by its denoised estimate $\hat{S}_0$, which is deterministically predicted from S_t at each step. This is justified by the causal structure of LiDAR sensing: reflectance is primarily a function of the underlying geometry. Since $\hat{S}_0$ better approximates

the clean geometry than the noisy S_t, it offers a more stable and physically meaningful reference for guiding the denoising of intensity. This substitution retains the Markov property of the diffusion chain while improving generation quality through cleaner conditioning.

3.3.1. Model Architecture

We present our approach as the *Structure-to-Intensity Diffusion* (SiD) framework, as illustrated in Fig. 2 middle. LiDAR scans are represented in range view as $x \in \mathbb{R}^{2\times H\times W}$, where the two channels correspond to geometric structure S (*e.g.*, depth) and reflectance intensity I. The proposed method operates directly on this 2D range-view representation and consists of four main components: a structure denoising branch with UNet_S, an intensity denoising branch with UNet_I, a weather encoder E_W, and a structure encoder E_S.

In the structure denoising phase, UNet_S takes a noisy geometric structure map S_t and predicts the noise component $\hat{\epsilon}_t^S$, resulting in the estimated cleaned output $\hat{S}_0$, which is then passed to the structure encoder E_S and aggregated into a latent embedding $\phi_S \in \mathbb{R}^D$ capturing geometric context. This embedding is used to condition UNet_I in the intensity denoising step, which denoises the noisy intensity map I_t into $\hat{I}_0$. Both branches also receive a shared weather-specific embedding $\phi_W \in \mathbb{R}^D$ that encodes environmental information, such as the desired weather tag.

Structure and Weather Encoders. The structure encoder E_S extracts high-level geometric cues $\hat{S}_0$ estimated by the structure denoising branch UNet_S at step t. This encoder is designed to extract spatial and geometric features into a compact embedding $\phi_S \in \mathbb{R}^D$, which serves as a conditioning signal for the intensity denoising branch. Specifically, E_S employs a sequence of layers to progressively aggregate features into the embedding space:

$$\phi_S = \mathrm{E}_S(\hat{S}_0). \tag{7}$$

This embedding captures the geometric context necessary for guiding intensity denoising, ensuring that intensity predictions remain coherent with the underlying structure.

Similarly, the weather-conditioning embedding ϕ_W is obtained via a weather encoder E_W, which shares the same architectural design as E_S. The weather encoder takes as input the augmented clean sample $[S_a, I_a]$ with the adverse-weather simulation, and produces a weather-specific embedding:

$$\phi_W = \mathrm{E}_W\left(S_a, I_a\right). \tag{8}$$

This embedding encodes the environmental degradation characteristic of the target weather condition, enabling the model to adapt its generation process accordingly.

Weather Embedding Bank. To enable weather conditioning in generation, we construct a weather bank $\mathcal{M}_W$ that stores various weather embeddings for different adverse weather conditions. Specifically, we apply RPWS to clear LiDAR scans to simulate realistic degradations under various weather tags. Each resulting scan is passed through the weather encoder E_W to extract an embedding $\phi_W^{(i)}$ via Eq. 8, with the associated weather condition c_i. The embedding $\phi_W^{(i)}$ is stored in the memory bank along with its label, yielding:

$$\mathcal{M}_W = \{(c_i, \phi_W^{(i)})\}_{i=1}^M.$$

At generation time, given a target condition c, we retrieve a weather-specific embedding ϕ_W by randomly sampling from the subset $\{\phi_W^{(i)} \mid c_i = c\}$ in $\mathcal{M}_W$. This approach provides a lightweight, data-driven conditioning mechanism that avoids runtime simulation and enables diverse weather-aware generation.

Denoiser Backbone. Both UNet_S and UNet_I follow a standard encoder-decoder architecture with skip connections [26, 39, 53], incorporating Adaptive Group Normalization (AdaGN) [6] at each layer to integrate the conditioning embeddings ϕ_W and ϕ_S. At each denoising step t, the structure branch UNet_S predicts the ϵ_t from the noisy input S_t, conditioned solely on the weather embedding ϕ_W and the timestep t:

$$\hat{\epsilon}_\theta^S = \mathrm{UNet}_S(S_t, t, \phi_W), \tag{9}$$

which is then used to compute an estimated clean structure map $\hat{S}_0 = \frac{S_t - \sqrt{1-\bar{\alpha}_t}\hat{\epsilon}_\theta^S}{\sqrt{\bar{\alpha}_t}}$. The $\hat{S}_0$ is subsequently fed into the structure encoder E_S to produce ϕ_S, which, together with ϕ_W, conditions the intensity branch UNet_I:

$$\hat{\epsilon}_\theta^I = \mathrm{UNet}_I(I_t, t, \phi_W + \phi_S). \tag{10}$$

In practice, we allow the denoiser to access both I_t and S_t to provide richer structural context and improve stability during training. Gradients from the intensity branch are detached with respect to S_t to preserve the intended dependency direction.

To effectively incorporate these conditioning signals ϕ into the UNets, for each intermediate feature map h in the network, the conditioning embeddings are first transformed via internal MLPs of AdaGN to produce scale and shift parameters:

$$\gamma, \beta = \mathrm{MLP}(\phi). \tag{11}$$

These parameters are then combined and applied to modulate the normalized features:

$$\mathrm{AdaGN}(h, \phi) = \mathrm{GN}(h) \odot (1 + \gamma) + \beta, \tag{12}$$

where $\mathrm{GN}(\cdot)$ denotes standard group normalization and $\odot$ represents element-wise multiplication. This modulation mechanism allows the network to dynamically adjust its internal representations based on both the weather-specific

context encoded in ϕ_W and the geometric structure encoded in ϕ_S, thereby enabling fine-grained control over the generated intensity patterns.

3.3.2. Training and Sampling

Training. At each training step, SiD receives an adverse-weather LiDAR scan in range view (S_0, I_0) and its corresponding weather label c. Each sample is either from real data or synthesized via RPWS. The associated weather embedding ϕ_W^c is computed using Eq. 8. To train the diffusion models, we apply the standard forward noising process:

$$S_t = \sqrt{\bar{\alpha}_t} S_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_S, \quad I_t = \sqrt{\bar{\alpha}_t} I_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_I, \tag{13}$$

where $\epsilon_S, \epsilon_I \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, with $\sqrt{\bar{\alpha}_t}$ and $\sqrt{1-\bar{\alpha}_t}$ being the signal and noise scales.

The structure denoiser UNet_S predicts $\hat{\epsilon}_\theta^S$ and reconstructs a clean estimate $\hat{S}_0$, which is then encoded into ϕ_S via E_S to condition the intensity denoiser UNet_I to predict $\hat{\epsilon}_\theta^I$. The training objective is a sum of denoising losses:

$$\mathcal{L} = \mathbb{E}_{t,\epsilon}\left[w(t)\|\epsilon_S - \hat{\epsilon}_\theta^S\|_2^2 + w(t)\|\epsilon_I - \hat{\epsilon}_\theta^I\|_2^2\right], \tag{14}$$

where $w(t)$ denotes the Min-SNR time-dependent weighting [14]. To enforce the causal factorization $S \to I$, gradients from the intensity branch are detached with respect to $\hat{S}_0$, ensuring that UNet_S is updated solely via its own loss.

Sampling. During inference, SiD follows the same sequential generation scheme as during training. Given a target weather label c, the corresponding weather embedding ϕ_W^c is randomly sampled from the weather bank $\mathcal{M}_W$. Starting from Gaussian noise $(S_T, I_T) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, geometry and reflectance are denoised in order. The structure branch produces $\hat{S}_0$, which is encoded into ϕ_S and used to condition the intensity denoiser. The reverse update for intensity $I_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(I_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\hat{\epsilon}_\theta^I\right) + \sigma_t z$, where $z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and S_{t-1} is updated in a similar way. This ensures reflectance aligns with the recovered geometry and target weather. The full sampling process is presented in Algorithm 1.

Algorithm 1 SiD Sampling Procedure

Require: Weather type $c \in \{\text{foggy, rainy, snowy}\}$, steps T
Require: Weather embedding bank $\mathcal{M}_W$
Require: Encoders E_S, E_W, Denoisers $\mathrm{UNet}_S, \mathrm{UNet}_I$
Sample weather embedding: $\phi_W^c \leftarrow \mathcal{M}_W(c)$
Initialize noise: $S_T, I_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, T-1, \ldots, 1$ **do**
  $\hat{\epsilon}_\theta^S \leftarrow \mathrm{UNet}_S(S_t, t, \phi_W^c)$
  $\hat{S}_0 \leftarrow \mathrm{Recon}(S_t, \hat{\epsilon}_\theta^S)$
  $\phi_S \leftarrow E_S(\hat{S}_0)$
  $\hat{\epsilon}_\theta^I \leftarrow \mathrm{UNet}_I([I_t; S_t], t, \phi_W^c + \phi_S)$
  $\hat{I}_0 \leftarrow \mathrm{Recon}(I_t, \hat{\epsilon}_\theta^I)$
  Update S_{t-1}, I_{t-1} via reverse diffusion
end for
return Final output: $\hat{x}_0 = [\hat{S}_0, \hat{I}_0]$

4. Experiments

4.1. Datasets and Evaluation

We conduct experiments with standard datasets including KITTI-360 [25] and STF [1], representing clear and adverse weather domains, respectively. KITTI-360 provides 76,715 clean LiDAR scans from a 64-beam Velodyne sensor, while STF includes approximately 7,000 real-world scans under fog, rain, and snow. Each weather condition spans hundreds to thousands of sequences. Following WeatherGen [45], we generate 10,000 clear-weather and 200 adverse-weather scans, and compare their distributions to testing sets from KITTI-360 and STF. We adopt standard evaluation metrics from prior work [32, 37, 44, 45, 54] to assess generation fidelity across multiple views: point clouds, range images, and bird's-eye views (BEV). Specifically, Fréchet Point Distance (FPD) and Fréchet Range Distance (FRD) measure distributional similarity using PointNet++ [36] and RangeNet++ [29] embeddings, while Jensen–Shannon Divergence (JSD) and Maximum Mean Discrepancy (MMD) evaluate occupancy and statistical gaps in BEV space.

To overcome the limitations of clear-weather-trained metrics, we further introduce the **Fréchet Minkowski Distance (FMD)**, which measures distributional alignment in a voxel-based feature space. This space is extracted using a Minkowski UNet trained via the SLidR framework [41] on the STF dataset. SLidR distills dense semantic features from large-scale pretrained image encoders into 3D point clouds through pixel-to-point supervision. As a result, the learned voxel features inherit the robustness and generalization ability, while being grounded in real-world adverse-weather LiDAR data. This makes FMD more resilient to weather-induced noise and better aligned with both geometric and reflectance structures, offering a more reliable evaluation under challenging conditions.

4.2. Implementation Details

Our model consists of two denoising UNets and two condition encoders. The UNets follow the same four-stage encoder-decoder design, with three residual blocks per stage for enhanced feature learning. The condition encoders use three convolutional layers, interleaved with two Vision-Mamba refinement layers [43], and two fully connected layers to produce compact embeddings. We adopt a two-stage training strategy: first, training on RPWS-augmented data to learn general weather priors, followed by fine-tuning on STF data to adapt to real-world conditions. Training is conducted on four NVIDIA L40S GPUs. We use learning rates

Table 1. Quantitative comparison under different adverse weather conditions on STF [1].

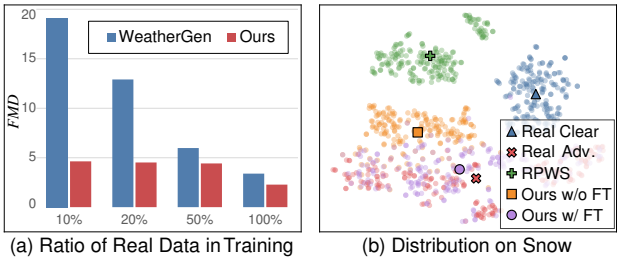
	Method	Point Cloud		Image	Bird's-Eye View	
		FMD ↓	FPD ↓	FRD ↓	MMD $\times 10^4$ ↓	JSD $\times 10^1$ ↓
Fog	FSRL [12]	-	319.32	2105.10	8.56	3.69
	WxGen [45]	3.29	314.14	1968.90	8.08	2.66
	Ours	**2.11**	**119.29**	**1370.00**	**3.62**	**1.47**
Rain	LISA [20]	-	301.11	1452.30	4.30	3.23
	WxGen [45]	4.65	86.40	1270.60	4.15	0.93
	Ours	**0.42**	**63.81**	**1140.21**	**1.11**	**0.93**
Snow	LSS [13]	-	106.37	1421.70	3.59	2.11
	WxGen [45]	3.39	59.28	1241.70	1.71	0.77
	Ours	**2.28**	**32.38**	**1080.68**	**1.28**	**0.65**

of 4×10^{-4} and 1×10^{-4} for pretraining and fine-tuning, respectively, with a batch size of 8 and 75,000 iterations. A cosine annealing schedule with warm-up is used, along with an Exponential Moving Average (decay 0.995) to stabilize training.

4.3. Weather-Conditioned LiDAR Generation

Adverse-Weather Generation. Tab. 1 reports the performance across fog, rain, and snow conditions on the STF benchmark [1]. Our method consistently achieves the best results across all metrics and weather types, showcasing its strong ability to model both geometric degradation and reflectance distortion. In fog, our method significantly improves over WeatherGen [45], reducing FPD from 314.14 to 119.29 and FRD from 1968.90 to 1370.00. It also achieves the lowest BEV discrepancies, with MMD dropping to 3.62 and JSD to 1.47. For rain, our method also leads in all metrics—most notably lowering FRD to 1140.21 and FPD to 63.81, while also achieving the best FMD score (0.42). Under snow, it continues to outperform baselines, reducing FPD to 32.38 and FRD to 1080.68, along with the best BEV alignment (MMD 1.28, JSD 0.65). In addition, as shown in Fig. 3(a), our approach is highly efficient in terms of real adverse data. It maintains strong performance even with just 10% of real data, while WeatherGen's quality degrades more sharply under limited supervision. This emphasizes the advantage of our structure-to-intensity strategy in jointly optimizing geometry and reflectance via synthetic priors, which effectively reduces dependence on costly real-world data. These consistent improvements highlight the benefits of our geometry-to-intensity generation strategy. By conditioning reflectance on structure, the proposed method produces physically coherent outputs and better captures how different weather conditions affect LiDAR signals. Quantitative results are also presented in Fig. 4.

Clear-Weather Generation. While our method is designed for adverse conditions, the Structure-to-Intensity strategy also holds in clear weather, where geometric and radiometric signals remain informative yet distinct. To evaluate its generality, we evaluate our work on the KITTI-360 dataset [25], which features high-quality LiDAR in clean environments. As shown in Tab. 2, our method outperforms the LiDAR generation baselines in point cloud fidelity (FMD 1.75), and achieves state-of-the-art level distributional alignment (FRD 135.13), and BEV consistency (JSD 0.20). These results affirm that decoupling geometry and intensity benefits generation quality even in clear settings.



(a) Ratio of Real Data in Training (b) Distribution on Snow

Figure 3. (a) Performance under snow condition with varying amounts of real data used in training. (b) t-SNE visualizations of real-world data, RPWS-simulated data, and our generated data with and without fine-tuning for snow condition.

Table 2. Quantitative comparison with state-of-the-art methods on the KITTI-360 dataset [25].

Method	Point Cloud		Image	Bird's-Eye View	
	FMD ↓	FPD ↓	FRD ↓	MMD $\times 10^4$ ↓	JSD $\times 10^1$ ↓
LiDARGen [54]	-	90.29	579.39	7.39	0.73
LiDM [37]	-	34.36	334.55	1.07	0.47
R2DM [32]	2.00	6.24	149.66	1.91	0.30
Tx2LDR [44]	2.72	**4.81**	164.16	0.49	**0.20**
WxGen [45]	1.86	6.15	138.62	**0.39**	**0.20**
Ours	**1.75**	5.62	**135.13**	0.41	**0.20**

4.4. Method Analysis

Component Ablation. We analyze the individual contributions of our simulation module (RPWS) and diffusion architecture (SiD) by comparing them with the WeatherGen [45] baseline, which employs Map-based Data Producer (MDP) for simulation and Spider Mamba Generator (SMG) as the generation backbone, as shown in Tab. 3. Replacing MDP with our RPWS yields substantial improvements across all metrics (*e.g.*, FPD drops from 271.12 to 109.42), demonstrating the effectiveness of using real-weather statistics for physically grounded simulation. Incorporating our SiD backbone further reduces FPD to 91.28, highlighting the advantage of modeling the asymmetric LiDAR sensing process via structure-to-intensity factorization. These benefits become more pronounced after fine-tuning on real STF data. Our full model (RPWS + SiD + FT) achieves the best performance, reducing FPD from 59.28 (WeatherGen + FT)

Table 3. Component ablation study conducted on the STF [1] snow test set. MDP and SMG are from WeatherGen [45], while RPWS and SiD are ours. Subscripts S, I denote structure and reflectance simulation in RPWS. "FT" indicates fine-tuning on real STF data.

Aug.	Arc.	FT	FMD ↓	FPD ↓	FRD ↓	MMD $\times 10^4$ ↓	JSD $\times 10^1$ ↓
MDP	SMG	-	6.11	271.12	2231.16	10.53	1.80
MDP	SMG	✓	3.39	59.28	1241.70	1.71	0.77
$\text{RPWS}_{S,I}$	SMG	-	3.94	109.42	1680.11	6.83	1.90
$\text{RPWS}_{S,I}$	SMG	✓	2.91	47.36	1282.60	1.78	0.96
RPWS_{S}	SiD	-	8.25	105.53	2074.08	7.50	1.88
RPWS_{I}	SiD	-	3.91	475.11	2066.29	12.03	2.60
$\text{RPWS}_{S,I}$	SiD	-	3.67	91.28	1523.65	4.41	1.59
$\text{RPWS}_{S,I}$	SiD	✓	**2.28**	**32.38**	**1080.68**	**1.28**	**0.65**

Table 4. Ablation study on the conditioning strategy in Structure-to-Intensity Diffusion (SiD). Metrics are evaluated on the STF [1] snow test set.

Conditioning Strategy	FMD ↓	FPD ↓	FRD ↓	MMD $\times 10^4$ ↓	JSD $\times 10^1$ ↓
Independent	4.92	55.79	1122.67	1.57	0.83
$\text{I} \rightarrow \text{S}$	4.42	66.20	1177.72	1.61	0.93
$\text{S} \rightarrow \text{I}$	**2.28**	**32.38**	**1080.68**	**1.28**	**0.65**

to 32.38 and FRD from 1241.70 to 1080.68. The baseline shows moderate improvements after fine-tuning (*e.g.*, FMD drops from 6.11 to 3.39), but still lags behind our model's performance. This gap suggests that without disentangling geometry and reflectance, SMG struggles to fully capture the complex, modality-specific degradation patterns inherent in adverse weather, particularly under limited real supervision. In contrast, our SiD-based design, with its structured conditioning, more effectively leverages real data, achieving stronger alignment with real-world distributions. In Fig. 3(b), we also visualize the distribution of different stages. As shown, our generated samples especially after fine-tuning are significantly closer to the real adverse-weather distributions across fog, rain, and snow.

Analysis of RPWS. RPWS is crucial for generating realistic and diverse training data under adverse weather conditions. As shown in Tab. 3 (rows 5–7), compared to the full design $\text{RPWS}_{S,I}$, removing either geometry degradation (denoted as RPWS_{I}) or reflectance degradation (denoted as RPWS_{S}) degrades generation quality, but their impacts differ. Without reflectance degradation, FMD and FRD worsen notably (FMD rises to 8.25), indicating poor alignment in intensity distributions. Conversely, excluding geometry degradation results in a large FPD (475.11), reflecting compromised structural fidelity. The full model combining both achieves the best overall performance across all metrics. These results suggest that weather-induced distortions impact geometry and reflectance in distinct ways, validating our design in SiD.

Analysis of SiD. The factorized denoising SiD improves generation fidelity through its sequential, physically

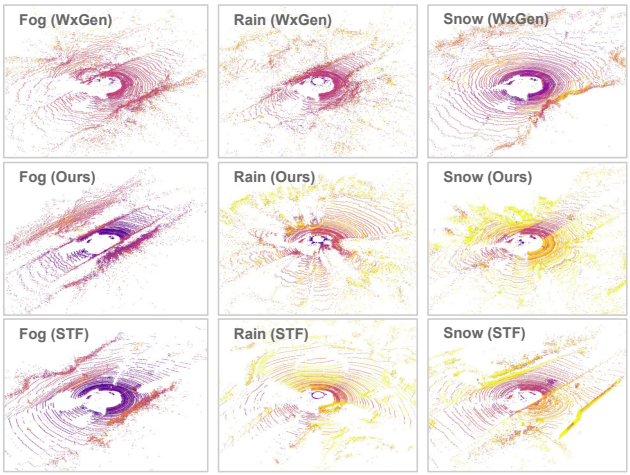


Figure 4. Visual results of generated LiDAR scans under adverse weather. Top to bottom: samples from WeatherGen [45], our model, and the STF dataset [1]. Compared to WeatherGen, our method generates more realistic and coherent point clouds, exhibiting structural details and reflectance patterns that align more closely with real-world characteristics.

informed design. As shown in Tab. 4, conditioning the intensity branch on geometry ($S \rightarrow I$) achieves the best performance across all metrics. Removing the conditioning entirely or reversing the dependency ($I \rightarrow S$) leads to noticeable degradation, particularly in FMD and FRD. These results highlight the importance of aligning the generation process with the causal structure of LiDAR formation, where reflectance depends on the underlying geometry. By disentangling the two modalities, SiD reduces modeling ambiguity, stabilizes optimization, and enhances both structural coherence and reflectance realism.

Analysis of Efficiency. Despite adopting a two-branch architecture, we maintain competitive efficiency by designing each branch to be lightweight. Our model contains 38.06M parameters and requires 1.599 GFLOPs, similar to WeatherGen [45]'s 32.59M parameters and 1.398 GFLOPs. Yet our method delivers much better generation quality across all metrics as demonstrated in Sec. 4.3. This confirms that our optimized dual-branch design strikes an effective balance between computational cost and generation fidelity.

5. Conclusions

We present a diffusion-based framework for realistic LiDAR generation under adverse weather. By combining Real-Prior Weather Simulation (RPWS), which models empirical degradation patterns, with Structure-to-Intensity Diffusion (SiD), which enforces the causal dependency of reflectance on geometry, our method captures both physical realism and structural fidelity. Extensive experiments show that our method consistently surpasses existing approaches across diverse weather conditions, offering a scalable solution for robust perception in challenging environments.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFA1004100, and in part by the National Natural Science Foundation of China under Grants 62476048 and 62572093.

References

[1] Mario Bijelic, Tobias Gruber, Fahim Mannan, Florian Kraus, Werner Ritter, Klaus Dietmayer, and Felix Heide. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2020. 3, 6, 7, 8

[2] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuScenes: A multimodal dataset for autonomous driving. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, pages 11621–11631, 2020. 2

[3] Anatoli Chaikovsky, Oleg Dubovik, Brent Holben, Andrey Bril, Philippe Goloub, Didier Tanré, Gelsomina Pappalardo, Ulla Wandinger, Ludmila Chaikovskaya, Sergey Denisov, et al. LiDAR-radiometer inversion code (LIRIC) for the retrieval of vertical aerosol properties from combined LiDAR/radiometer data: Development and distribution in EARLINET. *Atmos. Meas. Tech.*, 9(3):1181–1205, 2016. 2

[4] Xinlei Chen, Haoqi Fan, Ross B. Girshick, and Kaiming He. Improved baselines with momentum contrastive learning. *arXiv preprint*, arXiv:2003.04297, 2020. 2

[5] Weichen Dai, Shenzhou Chen, Zhaoyang Huang, Yan Xu, and Da Kong. LiDAR intensity completion: Fully exploiting the message from LiDAR sensors. *Sensors*, 22(19):7533, 2022. 2

[6] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat GANs on image synthesis. In *Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 8780–8794, 2021. 5

[7] Mariella Dreissig, Dominik Scheuble, Florian Piewak, and Joschka Boedecker. Survey on LiDAR perception in adverse weather conditions. In *Proc. of the IEEE Intelligent Vehicles Symp. (IV)*, 2023. 1

[8] Ryan Faulkner, Luke Haub, Simon Ratcliffe, Anh-Dzung Doan, Ian D. Reid, and Tat-Jun Chin. Simultaneous diffusion sampling for conditional LiDAR generation. *arXiv preprint*, arXiv:2410.11628, 2024. 1, 2

[9] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, pages 3354–3361, 2012. 1

[10] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennamoun. Deep learning for 3D point clouds: A survey. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, 43(12):4338–4364, 2021. 1

[11] Hamed Haghighi, Amir Samadi, Mehrdad Dianati, Valentina Donzella, and Kurt Debattista. Taming transformers for realistic LiDAR point cloud generation. *arXiv preprint*, arXiv:2404.05505, 2024. 1, 2

[12] 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 *Proc. of the IEEE/CVF Int. Conf. on Comput. Vis. (ICCV)*, 2021. 1, 2, 3, 7

[13] Martin Hahner, Christos Sakaridis, Mario Bijelic, Felix Heide, Fisher Yu, Dengxin Dai, and Luc Van Gool. LiDAR snowfall simulation for robust 3D object detection. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2022. 1, 2, 3, 7

[14] Tiankai Hang, Shuyang Gu, Chen Li, Jianmin Bao, Dong Chen, Han Hu, Xin Geng, and Baining Guo. Efficient diffusion training via min-SNR weighting strategy. In *Proc. of the IEEE/CVF Int. Conf. on Comput. Vis. (ICCV)*, 2023. 6

[15] Robin Heinzler, Philipp Schindler, Jürgen Seekircher, Werner Ritter, and Wilhelm Stork. Weather influence and classification with automotive LiDAR sensors. In *Proc. of the IEEE Intelligent Vehicles Symp. (IV)*, 2019. 1

[16] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 6840–6851, 2020. 2, 3

[17] Qianjiang Hu, Zhimin Zhang, and Wei Hu. RangeLDM: Fast realistic LiDAR point cloud generation. In *Proc. of the Eur. Conf. on Comput. Vis. (ECCV)*, pages 115–135, 2024. 1

[18] Boris Jutzi and Hans Gross. Normalization of LiDAR intensity data based on range and surface incidence angle. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 38:213–218, 2009. 4

[19] Alireza G. Kashani, Michael J. Olsen, Christopher E. Parrish, and Nicholas Wilson. A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. *Sensors*, 15(11):28099–28128, 2015. 2, 4

[20] Velat Kilic, Deepti Hegde, A. Brinton Cooper, Vishal M. Patel, and Mark Foster. LiDAR light scattering augmentation (LISA): Physics-based simulation of adverse weather conditions for 3D object detection. In *Proc. of the IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP)*, 2025. 1, 2, 3, 7

[21] 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 *Proc. of the IEEE/CVF Int. Conf. on Comput. Vis. (ICCV)*, 2023. 1

[22] Lingdong Kong, Wesley Yang, Jianbiao Mei, Youquan Liu, Ao Liang, Dekai Zhu, Dongyue Lu, Wei Yin, Xiaotao Hu, Mingkai Jia, et al. 3D and 4D world modeling: A survey. *arXiv preprint*, arXiv:2509.07996, 2025. 1

[23] Antero Kukko, Sanna Kaasalainen, and Paula Litkey. Effect of incidence angle on laser scanner intensity and surface data. *Appl. Opt.*, 47(7):986–992, 2008. 2

[24] Haoyang Li, Xin Wang, Ziwei Zhang, and Wenwu Zhu. OOD-GNN: Out-of-distribution generalized graph neural network. *IEEE Trans. Knowl. Data Eng. (TKDE)*, 35(7): 7328–7340, 2023. 1

[25] Yiyi Liao, Jun Xie, and Andreas Geiger. KITTI-360: A novel dataset and benchmarks for urban scene understanding in 2D and 3D. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, 45(3):3292–3310, 2023. 2, 6, 7

[26] Yun Liu, Yu-Huan Wu, Guolei Sun, Le Zhang, Ajad Chhatkuli, and Luc Van Gool. Vision transformers with hierarchical attention. *Machine Intelligence Research*, 21(4): 670–683, 2024. 5

[27] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. DPM-Solver++: Fast solver for guided sampling of diffusion probabilistic models. *Machine Intelligence Research*, 22(4):730–751, 2025. 3

[28] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *Proc. of the Eur. Conf. on Comput. Vis. (ECCV)*, 2020. 2

[29] Andres Milioto, Ignacio Vizzo, Jens Behley, and Cyrill Stachniss. RangeNet++: Fast and accurate LiDAR semantic segmentation. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 4213–4220, 2019. 6

[30] Alhassan Mumuni, Fuseini Mumuni, and Nana Kobina Gerrar. A survey of synthetic data augmentation methods in machine vision. *Machine Intelligence Research*, 21(5):831–869, 2024. 1

[31] Kazuto Nakashima and Ryo Kurazume. Learning to drop points for LiDAR scan synthesis. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 222–229, 2021. 2

[32] Kazuto Nakashima and Ryo Kurazume. LiDAR data synthesis with denoising diffusion probabilistic models. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 14724–14731, 2024. 2, 6, 7

[33] Kazuto Nakashima, Yumi Iwashita, and Ryo Kurazume. Generative range imaging for learning scene priors of 3D LiDAR data. In *Proc. of the IEEE/CVF Winter Conf. on Applications of Comput. Vis. (WACV)*, pages 1256–1266, 2023. 2

[34] Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *Proc. of the Int. Conf. on Machine Learning (ICML)*, pages 8162–8171, 2021. 2

[35] Lucas Nunes, Rodrigo Marcuzzi, Benedikt Mersch, Jens Behley, and Cyrill Stachniss. Scaling diffusion models to real-world 3D LiDAR scene completion. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2024. 1, 2

[36] Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In *Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017. 6

[37] Haoxi Ran, Vitor Guizilini, and Yue Wang. Towards realistic scene generation with LiDAR diffusion models. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2024. 1, 2, 6, 7

[38] Ralph H. Rasshofer, Martin Spies, and Hans Spies. Influences of weather phenomena on automotive laser radar systems. *Adv. Radio Sci.*, 9:49–60, 2011. 2

[39] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Proc. of the Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 234–241, 2015. 5

[40] Axel Sauer, Kashyap Chitta, Jens Müller, and Andreas Geiger. Projected GANs converge faster. In *Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2021. 2

[41] Corentin Sautier, Gilles Puy, Spyros Gidaris, Alexandre Boulch, Andrei Bursuc, and Renaud Marlet. Image-to-LiDAR self-supervised distillation for autonomous driving data. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2022. 2, 6

[42] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *Proc. of the Int. Conf. on Learning Representations (ICLR)*, 2021. 2

[43] Renkai Wu, Yinghao Liu, Guochen Ning, Pengchen Liang, and Qing Chang. UltraLight VM-UNet: Parallel vision mamba significantly reduces parameters for skin lesion segmentation. *Patterns*, 6(11):101298, 2025. 6

[44] Yang Wu, Kaihua Zhang, Jianjun Qian, Jin Xie, and Jian Yang. Text2LiDAR: Text-guided LiDAR point cloud generation via equirectangular transformer. In *Proc. of the Eur. Conf. on Comput. Vis. (ECCV)*, pages 291–310, 2024. 2, 6, 7

[45] Yang Wu, Yun Zhu, Kaihua Zhang, Jianjun Qian, Jin Xie, and Jian Yang. WeatherGen: A unified diverse weather generator for LiDAR point clouds via spider mamba diffusion. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2025. 1, 2, 3, 6, 7, 8

[46] Zehuan Wu, Jingcheng Ni, Xiaodong Wang, Yuxin Guo, Rui Chen, Lewei Lu, Jifeng Dai, and Yuwen Xiong. HoloDrive: Holistic 2D–3D multi-modal street scene generation for autonomous driving. *arXiv preprint*, arXiv:2412.01407, 2024. 2

[47] Aoran Xiao, Jiaxing Huang, Weihao Xuan, Ruijie Ren, Kangcheng Liu, Dayan Guan, Abdulmotaleb El Saddik, Shijian Lu, and Eric P. Xing. 3D semantic segmentation in the wild: Learning generalized models for adverse-condition point clouds. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2023. 1

[48] Yuwen Xiong, Wei-Chiu Ma, Jingkang Wang, and Raquel Urtasun. UltraLiDAR: Learning compact representations for LiDAR completion and generation. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2023. 2

[49] Tianyi Yan, Junbo Yin, Xianpeng Lang, Ruigang Yang, Cheng-Zhong Xu, and Jianbing Shen. OLiDM: Object-aware LiDAR diffusion models for autonomous driving. In *Proc. of the AAAI Conf. on Artificial Intelligence (AAAI)*, 2025. 1, 2

[50] Longyu Yang, Ping Hu, Shangbo Yuan, Lu Zhang, Jun Liu, Hengtao Shen, and Xiaofeng Zhu. Towards explicit geometry-reflectance collaboration for generalized LiDAR

segmentation in adverse weather. In *Proc. of the IEEE/CVF Conf. on Comput. Vis. and Pattern Recog. (CVPR)*, 2025. 1

[51] Longyu Yang, Ping Hu, Lu Zhang, Jun Liu, Yap-Peng Tan, Heng Tao Shen, and Xiaofeng Zhu. Rethinking range-view LiDAR segmentation in adverse weather. *arXiv preprint*, arXiv:2506.08979, 2025. 2

[52] Shengyuan Zhang, An Zhao, Ling Yang, Zejian Li, Chenye Meng, Haoran Xu, Tianrun Chen, AnYang Wei, Perry Pengyun Gu, and Lingyun Sun. Distilling diffusion models to efficient 3D LiDAR scene completion. In *Proc. of the IEEE/CVF Int. Conf. on Comput. Vis. (ICCV)*, 2025. 1, 2

[53] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. UNet++: A nested U-Net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 3–11, 2018. 5

[54] Vlas Zyrianov, Xiyue Zhu, and Shenlong Wang. Learning to generate realistic LiDAR point clouds. In *Proc. of the Eur. Conf. on Comput. Vis. (ECCV)*, 2022. 1, 2, 6, 7