

## Article

# LiDAR Point Cloud Augmentation for Adverse Conditions Using Conditional Generative Model

Yuxiao Zhang <sup>1</sup>, Ming Ding <sup>2</sup>, Hanting Yang <sup>1</sup>, Yingjie Niu <sup>1</sup>, Maoning Ge <sup>1</sup>, Kento Ohtani <sup>1</sup>, Chi Zhang <sup>3,\*</sup>  
and Kazuya Takeda <sup>1,4</sup>

<sup>1</sup> Graduate School of Informatics, Nagoya University, Furo-cho, Chikusa-Ward, Nagoya 464-8601, Japan; yuxiao.zhang@g.sp.m.is.nagoya-u.ac.jp (Y.Z.)

<sup>2</sup> Zhejiang Fubang Technology Inc., Ningbo R&D Campus Block A, Ningbo 315048, China

<sup>3</sup> School of Materials Science and Engineering, Liaoning University of Technology, No. 169, Shiyong St., Guta District, Jinzhou 121001, China

<sup>4</sup> Tier IV Inc., Nagoya University Open Innovation Center, 1-3, Mei-eki 1-chome, Nakamura-Ward, Nagoya 450-6610, Japan

\* Correspondence: clzhangchi@lnut.edu.cn

**Abstract:** The perception systems of autonomous vehicles face significant challenges under adverse conditions, with issues such as obscured objects and false detections due to environmental noise. Traditional approaches, which typically focus on noise removal, often fall short in such scenarios. Addressing the lack of diverse adverse weather data in existing automotive datasets, we propose a novel data augmentation method that integrates realistically simulated adverse weather effects into clear condition datasets. This method not only addresses the scarcity of data but also effectively bridges domain gaps between different driving environments. Our approach centers on a conditional generative model that uses segmentation maps as a guiding mechanism to ensure the authentic generation of adverse effects, which greatly enhances the robustness of perception and object detection systems in autonomous vehicles, operating under varied and challenging conditions. Besides the capability of accurately and naturally recreating over 90% of the adverse effects, we demonstrate that this model significantly improves the performance and accuracy of deep learning algorithms for autonomous driving, particularly in adverse weather scenarios. In the experiments employing our augmented approach, we achieved a 2.46% raise in the 3D average precision, a marked enhancement in detection accuracy and system reliability, substantiating the model's efficacy with quantifiable improvements in 3D object detection compared to models without augmentation. This work not only serves as an enhancement of autonomous vehicle perception systems under adverse conditions but also marked an advancement in deep learning models in adverse condition research.

**Keywords:** point cloud processing; data augmentation; adverse conditions; conditional generative model



**Citation:** Zhang, Y.; Ding, M.; Yang, H.; Niu, Y.; Ge, M.; Ohtani, K.; Zhang, C.; Takeda, K. LiDAR Point Cloud Augmentation for Adverse Conditions Using Conditional Generative Model. *Remote Sens.* **2024**, *16*, 2247. <https://doi.org/10.3390/rs16122247>

Academic Editors: Zhenwei Shi, Teresa Pamuła and Wiesław Pamuła

Received: 31 May 2024

Revised: 17 June 2024

Accepted: 19 June 2024

Published: 20 June 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Perception systems in autonomous driving technologies face significant challenges in adverse weather conditions, such as rain, fog, snow, intense light, and sensor contamination. Recent advancements in machine learning have led to the development of weather condition networks using deep learning techniques [1]. Adverse conditions such as snow present unique difficulties in autonomous driving scenarios and represent one of the most significant challenges to sensor fidelity, including the occlusion to pertinent objects and fake detection due to the accumulation of noise clusters [2] in LiDAR (light detection and ranging) point clouds. Adverse conditions present tangible threats to sensors, yet research on this topic is relatively sparse. Recent studies on the impact of adverse conditions on

point clouds have primarily focused on k-d-tree-based neighbor-searching outlier filters [3], with the de-noising performance nearly reaching its peak [4].

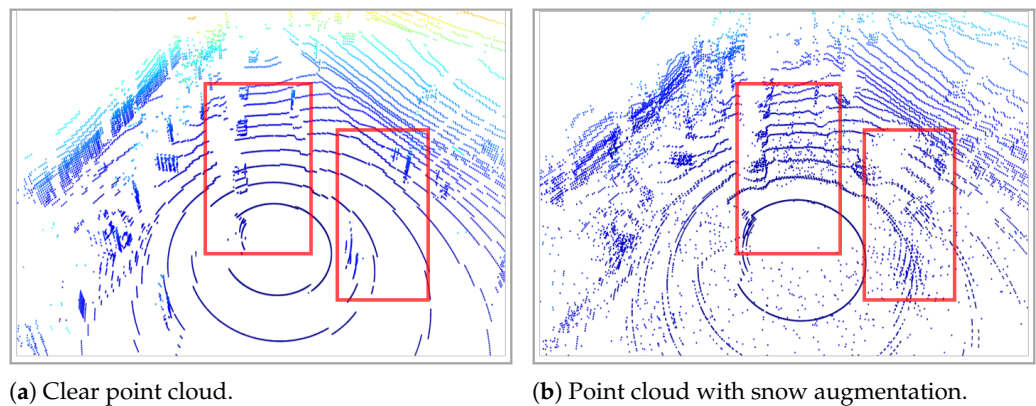
On the other hand, deep learning models, unlike filters, offer a certain level of explainability and hold the potential to discern both surface-level and hidden features of adverse effects in specific driving scenes, enabling them to perform snow synthesis alongside snow de-noising. LiDAR transformation models can only help the enhancement of perception to a certain limit because the removal of noise points only reduces the possibility of adverse effects being recognized as fake objects but does not deepen the understanding of adverse effects in point clouds. This is largely due to the inherent difficulty and unreliability of annotating or labeling adverse conditions [5] and the lack of samples featuring adverse conditions within current common driving datasets [1]. Furthermore, datasets specifically targeted at collecting adverse conditions in certain areas, like snowy regions, often face significant domain gap issues when being input into perception models trained with general data [6]. As a result, augmenting clear condition datasets with simulated adverse conditions has become one of the most effective and popular approaches to enhance the capability of deep-learning-based models in addressing adverse challenges.

Compared to methods based on classification and removal, those employing adverse-conditions data augmentation demonstrate superior effectiveness and accuracy in terms of perception enhancement due to the enabling of more adverse data used for model training [7]. The presence of weather phenomena such as water mist, agglomerated fog, and snow swirls sometimes manifest as solid “ghost” objects in the point cloud, while scattered noise point distribution is the result formed from ego vehicle motion, vehicle interactions, environmental layouts, and wind, instead of a fixed pattern or predictable mathematical distribution. As a result, we need a designated augmentation method to better adapt to the features of the adverse conditions with a deep understanding of not only the pattern of the adverse effect but the entire driving scenario, thereby enabling the generation of paired adverse data. The integration of paired datasets that include both adverse effects and corresponding classifications is proven to significantly improve the detectors’ proficiency in identifying smaller entities, like pedestrians and cyclists, in challenging driving conditions [8].

In this paper, we propose a LiDAR point cloud augmentation model with conditional guides to realize the expansion of adverse condition data. The proposed model generates segmentation maps with unique labels that identify the presence of adverse effects in point clouds, serving as conditional guides. The proposed approach involves developing efficient fusion techniques to integrate these guides with raw data into the generative model. This integration facilitates the classification and synthesis of adverse effects. The proposed method effectively incorporates realistic adverse conditions into point clouds, taking into account scene context and striving to preserve structural integrity, which is crucial for creating reliable paired datasets. Additionally, by leveraging the conditional guide, the aim is to bridge the domain gap related to traffic patterns and environmental layouts, which often pose challenges in realistically augmenting adverse data. This approach enables the creation of quasi-natural adverse effects in clear datasets, thereby enhancing the robustness and diversity of adverse data collection. Figure 1 shows an example of the expected adverse effect augmentation task.

The main contributions of this work are as follows:

1. Segmentation maps of adverse effects are produced via a designated 3D clustering algorithm, serving as conditional guides for generative models.
2. A novel early data fusion approach was developed to integrate raw and segmentation data, exhibiting remarkable capabilities in directing the creation of adverse effects.
3. High-level robustness in the generation of paired adverse data is presented through the generation of quasi-natural adverse effects across huge domain gaps in terms of traffic layouts and environments. A notable enhancement in detection performance with the proposed data augmentation scheme is validated through experiments.



**Figure 1.** Point cloud in clear driving scenario (a) and corresponding snow-augmented results (b) expected from augmentation models. Red boxes denote locations where snow effects were generated. Color encoded by height.

## 2. Related Works

Object augmentation primarily serves the purpose of completing occluded objects in indoor environments, while scene augmentation focuses on transforming entire scenes, commonly involving operations such as jittering, rotation, and shearing. These augmentation techniques aim to improve the accuracy and robustness of downstream tasks, primarily semantic segmentation and object detection. However, addressing adverse conditions in driving scenes requires a combination of both object and scene augmentation, which can be approached from two perspectives: adverse effect synthesis and existing adverse effect enrichment.

### 2.1. Adverse Effect Synthesis

Early attempts at adverse effect synthesis focused on replicating the optical returns of signals measured under weather conditions to analyze the influence of adverse conditions on LiDAR sensors [9,10]. While this laid the foundation for signal waveform processing in addressing adverse condition challenges, it highlighted the essence of LiDAR data augmentation: synthesizing perception results that mimic the effects of real adverse conditions on LiDAR sensors. However, the realization of the adverse effects in point clouds still depends on the collections in weather chambers [8] because of the requirement for paired data. Apart from the low domain similarity between chambers and real roads, common experimental facilities with controllable precipitation rates worldwide struggle to simulate complex adverse conditions such as dynamic snowfall [1].

To address this limitation, recent works have explored treating adverse effects like rain, fog, and snow as noise points with specific distributions (e.g., uniform or normal) and translating entire scenes accordingly [11,12]. While systematic, these approaches struggle to capture the variability and nuances in the distribution of adverse effects in real-world environments. Consequently, there is a need to develop methods for realizing adverse data augmentation in LiDAR point clouds that are similar to image-based approaches, thereby improving the robustness of LiDAR perception in autonomous driving scenarios.

With the development of Generative Adversarial Networks (GANs), researchers have leveraged their ability to create photorealistic images without paired examples, focusing on converting clear images to hazy ones [13–15]. Extending this approach to LiDAR point clouds has shown promising results, with initial attempts at translating simulated driving scenes into artificial point clouds [16,17]. Subsequent works have explored point cloud translations between sunny and challenging weather conditions like rain and fog [18]. While these methods realize adverse effect augmentation to some extent, they face challenges when dealing with different domains, such as translating from

fog chamber scenes to real roads. Artificial precipitation produced by sprinklers is often misinterpreted as vertical cylinders by LiDARs [19], potentially compromising the interpretation of weather reflection features in point clouds and diminishing the quality of the augmented data.

In summary, while existing methods for adverse effect synthesis in LiDAR point clouds have made progress, they still face limitations in capturing the variability and nuances of real-world adverse conditions. There is a need for more advanced techniques that can better synthesize realistic adverse effects while accounting for the complexities of different driving scenarios and environmental factors.

## 2.2. Existing Adverse Effects Enrichment

Adverse effects enrichment was initially explored in the domain of image processing, providing valuable insights and principles that can be applied to LiDAR point clouds. While augmentation techniques for rain and fog conditions are relatively more mature, snow augmentation remains less developed due to the scarcity of snow presence in datasets. Attempts at the generation of snowflakes and realistic rendering across the entire driving scene have been made [20] in images, and synthetic snow scenes have also been used for perception models' evaluation purposes [21]. However, the translation of snow scenes and the embodiment of different snow levels are still ongoing challenges.

Beyond precipitation conditions, sensor contamination, such as mud or water stains obscuring camera lenses, is another critical factor affecting perception. Acquiring datasets capturing these effects is extremely difficult, leading to the development of models capable of generating soiling effects on cameras and augmenting contamination datasets [22]. These augmented effects are remarkably genuine, and downstream removal models heavily rely on such data for training purposes.

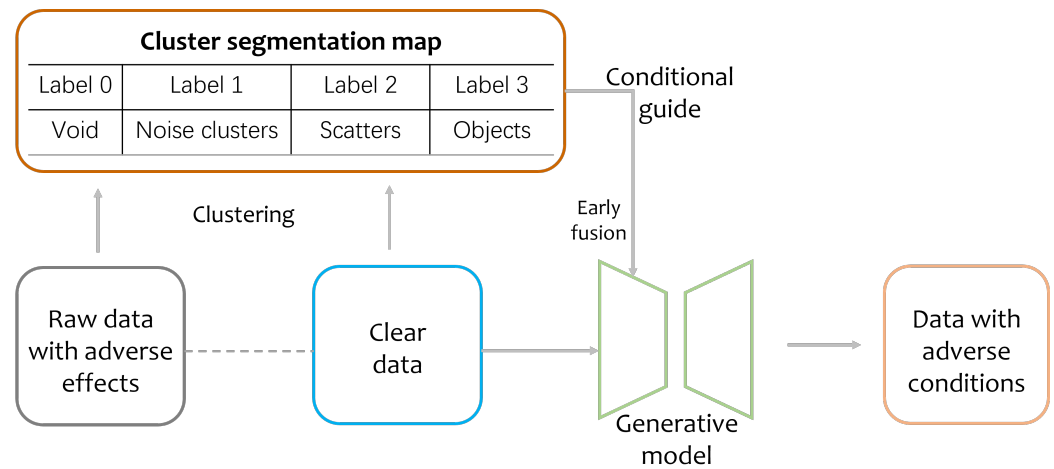
However, a significant challenge in using synthetic or augmented adverse effects is the potential domain gap between the synthetic data and real-world conditions [23]. To address this, researchers have developed models derived from pre-existing adverse effects within datasets, aiming to preserve the authenticity of the effects [7]. For instance, Hahner et al. [7] developed a method that simulates snow particles in a 2D space corresponding to each LiDAR line and adjusts each LiDAR beam's measurements based on the resulting geometry. They also incorporated ground wetness, a common occurrence during snowfall, into their LiDAR point clouds to supplement the augmentation. While their approach successfully simulated snowfall, it predominantly focuses on light snowfall conditions under the rate of 2.5 mm/h [24], wherein the prevalent snow effects in LiDAR point clouds manifest as dispersed noise points rather than snow clusters. The notable enhancement in the performance of 3D object detection subsequent to training with the addition of the semi-synthetic snowy data substantiates the successful simulation of snowfall and underscores the importance of adverse effect augmentation.

In summary, existing methods for adverse effects enrichment in LiDAR point clouds have made notable progress, particularly in simulating light precipitation conditions and sensor contamination. However, challenges remain in realistically translating and embodying different adverse effect levels, especially for heavier snowfall conditions that may result in more complex snow cluster formations. Additionally, addressing the domain gap between synthetic and real-world data is crucial for improving the generalization and robustness of perception models trained on augmented data.

## 3. Methods

Figure 2 shows the workflow of the condition-guided augmentation model. Our methodology begins with raw data that include adverse conditions, which are then processed using a 3D clustering algorithm to create a segmentation map. This segmentation map is then utilized as a conditional guide in our generative model, alongside clear data. For optimal training outcomes, it is advantageous for this clear dataset to have some correlation with the raw data. This correlation facilitates a more intuitive understanding of

the underlying logic in adverse effect classification. Therefore, we employ filtered raw data [25] in this context. In the following subsections, we will introduce the construction of the segmentation map and the architecture of the generative model.



**Figure 2.** Workflow of the condition-guided adverse effects augmentation model based on novel segmentation maps production and early data fusion techniques. The clear data input is obtained from filtered raw adverse data to establish intrinsic correlation for optimal training. The cluster segmentation map serves as a conditional guide, which can be input into the generative model through early data fusion. Data with adverse conditions are generated under the guidance of the segmentation map.

### 3.1. Clusters Classification and Segmentation Map

In our task of adverse effects augmentation, we use the CADC (Canadian Adverse Driving Conditions) dataset [26] for its excellent representation of snow clusters. To establish a solid baseline for classifying snow swirl clusters, we meticulously chose 2691 samples of LiDAR point cloud frames that prominently feature snow clusters from the entire dataset. These samples consist of 120 different sequences, corresponding to 120 unique driving scenarios. These scenarios comprehensively cover aspects such as the ego vehicle’s movement, interactions with other traffic participants, and wind influence.

For each of these scenarios, we select one representative frame for manual annotation at the cluster level. We manually identified and labeled all snow clusters that are discernible to the human eye. A 3D clustering algorithm based on OPTICS (ordering points to identify the clustering structure) [6,27] is then employed to aggregate and analyze their spatial and clustering characteristics. An adaptive cluster amount determination step based on DBSCAN [28] was added before the OPTICS algorithm to improve the clustering efficiency. With this mechanism, the possibility of one object being divided into several clusters or several objects being partitioned into one cluster is greatly minimized, leading to the result of the most appropriate number of cluster distributions. The 3D clustering algorithm produces seven metrics that could help the quantitative evaluation, as employed in our previous work [6]. Among these, reachability distance and cluster size are particularly useful for producing cluster-based segmentation maps as shown in Figure 1. The details of these metrics are shown below:

- Noise number: Points without any neighbor points within a designated range (solitary points) are considered noise points, mostly snowflakes. A decrease in noise number is one of the most direct indicators of a low adverse effect presence. The count  $N$  of the points that have fewer than  $P_{min}$  neighbors within a given radius  $\epsilon$  is shown as

$$N = |\{x \in X \mid P(x, \epsilon) < P_{min}\}| \quad (1)$$

where  $X$  is the set of all points and  $P(x,\epsilon)$  returns the count of points within the  $\epsilon$ -radius of  $x$  [29].

- Cluster number: A main output of the algorithm, representing groups of data points that are closely related based on their reachability. The cluster number can be simply denoted as  $C$  [30].
- Reachability distance: The smallest distance required to connect point  $A$  to point  $B$  via a path of points that satisfy the density criteria. Normally, the average reachability distance would rise along with larger cluster numbers.

For points  $A$  and  $B$ , the reachability distance  $R(A, B)$  could be defined as

$$R(A, B) = \max(\text{CD}(A), d(A, B)) \quad (2)$$

where  $\text{CD}(A)$  (the core distance) is the minimum distance required to separate  $A$  from its neighbors, and  $d(A, B)$  is the Euclidean distance between  $A$  and  $B$  [29].

- Inter-cluster distances (ICDs): The concept here involves identifying the centroid, or the average point, of each cluster, and subsequently computing the distance between every possible pair of centroids. Should there be a decrease in the average of these distances, it would suggest a rise in the number of clusters and a more concentrated cluster distribution. In the context of this study, such a pattern could be interpreted as an effect of high adverse effects appearance.

For clusters  $i$  and  $j$  with centroids  $C_i$  and  $C_j$ :

$$\text{ICD} = \frac{1}{\binom{C}{2}} \sum_{i \neq j} d(C_i, C_j) \quad (3)$$

where  $\binom{C}{2}$  denotes the number of unique pairs of clusters [31].

- Size of clusters: This is essentially determined by the average number of points each cluster holds. Under conditions dominated by scattered snow, the snow noise points tend to form numerous small-scale clusters. Their presence, consequently, leads to a diminution in the average size of the clusters [30].

For cluster  $i$  with  $n_i$  points, the average size  $S$  could be

$$S = \frac{1}{C} \sum_{i=1}^C n_i \quad (4)$$

- Silhouette score: Measures the cohesion within clusters and the separation between clusters. A silhouette score close to 1 indicates a good clustering quality, while a score close to  $-1$  indicates poor clustering. A lower silhouette score is commonly observed in adverse and snowy conditions due to the more overlap between clusters.

For a point  $x$  in cluster  $A$ , the silhouette score  $s(x)$  is calculated as

$$s(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}} \quad (5)$$

where  $a(x)$  is the average distance from  $x$  to the other points in the same cluster  $A$ , and  $b(x)$  is the smallest average distance from  $x$  to points in a different cluster, minimized over clusters. The overall silhouette score is then the average  $s(x)$  over all points [32].

- Davies–Bouldin index (DBI): Measures the ratio of within-cluster scatter to between-cluster separation and assesses the quality of the overall cluster separation. A lower Davies–Bouldin index indicates better clustering, with zero being the ideal value. Adverse conditions with many noise points or swirl clusters exhibit higher values of the DBI.

The Davies–Bouldin index DBI is calculated as

$$DBI = \frac{1}{C} \sum_{i=1}^C \max_{i \neq j} \left( \frac{s_i + s_j}{d(C_i, C_j)} \right) \quad (6)$$

where  $s_i$  is the average distance of all points in cluster  $i$  to centroid  $C_i$  [33].

With the establishment of clusters, we developed an automatic tool for classifying snow effects. This tool operates according to Algorithm 1. In Algorithm 1, X means the horizontal axis in BEV (Bird Eye View) and Y means the vertical axis while Z denotes the height axis (from ground to sky). The threshold values for the spatial XYZ coordinates are established through a detailed evaluation of the geometric properties of each relevant cluster by the proposed algorithm. These determined ranges effectively reduce the likelihood of objects being mistakenly identified as snow clusters to the minimum level. Notably, the XYZ coordinates represent the central point of each cluster, thus almost eliminating the risk of misidentifying clusters on the margins.

---

**Algorithm 1** Point cloud segmentation through a 3D clustering algorithm based on OPTICS

---

**Require:** Cluster size, Center X, Y, Z coordinates, and average reachability distance for each cluster.

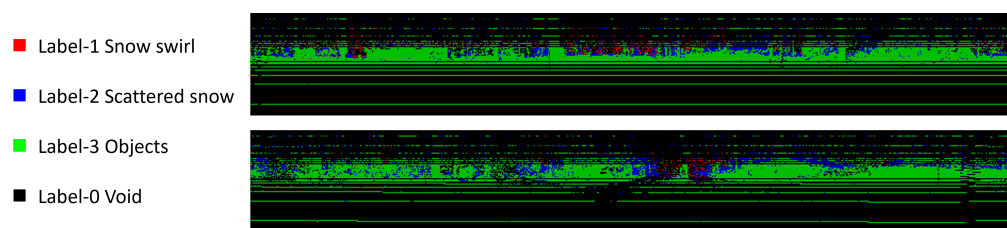
**Ensure:** Labels for each cluster representing different snow conditions.

- 1: **for** each cluster in the dataset **do**
  - 2:   **if** Cluster size  $\in [2, 500]$  **then**
  - 3:     **if** Center X  $\in [x_1, x_2]$  **and** Center Y  $\in [y_1, y_2]$  **and** Center Z  $\in [z_1, z_2]$  **and** avg. reachability distance  $\in [0.5, 2]$  **then**
  - 4:       Label all points within this cluster as 1 (swirl clusters).
  - 5:     **else**
  - 6:       Label all points within this cluster as 3 (objects).
  - 7:     **end if**
  - 8:   **else if** Cluster size  $\in [1, 1]$  **then**
  - 9:     Label all points within this cluster as 2 (scattered noise points).
  - 10: **else**
  - 11:    Label all points within this cluster as 3 (objects).
  - 12: **end if**
  - 13: **end for**
  - 14: Examine each pixel in the depth images for detectable reflectance values.
  - 15: **if** pixels lack reflectance **then**
  - 16:    Label these pixels as 0 (void areas).
  - 17: **end if**
- 

The CADC dataset was collected in Canada where vehicles drive on the right side of the road and oncoming vehicles will approach the ego vehicle from the left. This justifies the slightly closer threshold for Center X on the left ( $-7.7646$ ) compared to the right ( $8.4354$ ), relative to the ego vehicle. Additionally, the movement of the ego vehicle tends to disturb snow on the ground, creating whirls of snow clusters in its wake based on previous observations [2]. Consequently, the threshold for Center Y at the rear ( $-10.4947$ ) is found further from the vehicle compared to the front ( $7.5053$ ). Outside of the interval of  $[0.5165, 1.9999]$  on the height direction, no clusters fit the criteria of snow within the given X and Y range, hence the Z thresholds.

The decision to select a cluster size interval ranging from 2 to 500 is based on our analysis of each cluster, which includes every non-solitary noise point. Observations across all 120 scenarios indicate that there are no snow clusters exceeding 500 in terms of cluster size. Similarly, the interval for reachability distance is determined based on thorough inspections of the maximum and minimum limits. To ensure absolute accuracy in cluster identification, all the thresholds including the XYZ coordinates are further corroborated by a secondary layer of human verification.

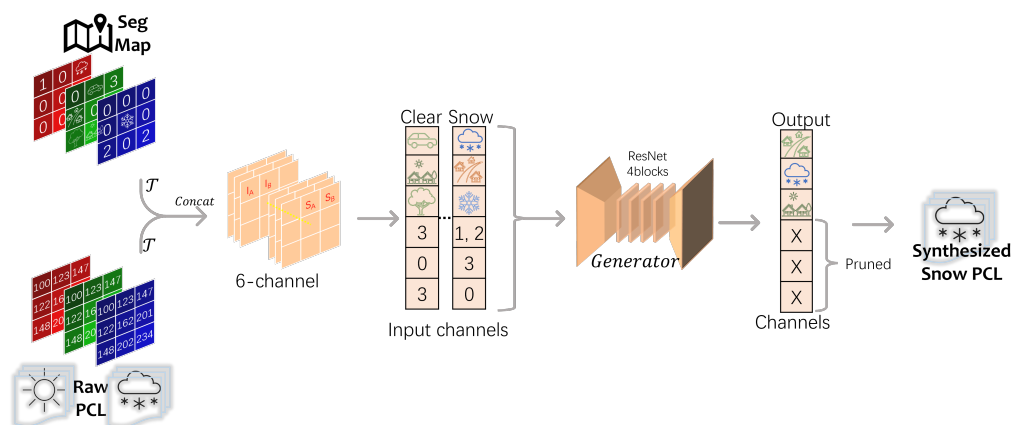
This classification equates to a segmentation map with four distinct labels, effectively illuminating the presence of various adverse conditions. We adopt the same RGB depth image format in our previous work [6,34], which represents the perpendicular projection of point clouds while keeping the depth information (i.e., the distances from points to the ego vehicle) within its pixel value, while the RGB channels now store labels in the segmentation map datasets. In this map, Labels 1 and 2 are assigned to noise clusters and individual noise points, stored in the red and blue channels, respectively. Label 3 encompasses all other elements, such as objects and structures, essentially anything not classified as an adverse effect, stored in the green channel. Label 0 represents void areas, characterized by the absence of reflective signals in the point clouds, stored in all three channels. This self-adaptive approach is designed to accurately capture all snow clusters, theoretically eliminating the possibility of errors. The outcomes from our classification tool are depicted in Figure 3. This segmentation map not only provides a fundamental insight into the characteristics of adverse effects but also aids in the generation of such conditions in data augmentation.



**Figure 3.** Examples of segmentation maps of CADC dataset in a depth image format for visualization. Images are rendered with pixel values multiplied by 64 under the OpenCV BGR environment for better illustration purposes. Red points denote snow clusters, blue denotes scattered snow points, green denotes all the objects, and black means void (no signal).

### 3.2. Conditional Guide Data Fusion

The above segmentation map is utilized as a conditional guide in our generative model, through an early fusion technique alongside clear data. The clear data would go through the same segmentation-map-producing procedure in order to provide compatible conditional guide data for the generative models. The core principle of early fusion with the segmentation map is to activate the effectiveness of the conditional guide right from the beginning. This ensures that the entire generation process is monitored and constrained within the parameters set by the guide, as shown in Figure 4.



**Figure 4.** Diagram of the early fusion process for conditional augmentation in point clouds.

The fusion process is initiated by setting a random transforming seed,  $\mathcal{T}$ , a standardized protocol to ensure that both the point cloud inputs from domains  $A$  and  $B$  —  $I_A, I_B$  — and their corresponding segmentation maps  $S_A, S_B$  undergo an identical transforming process prior to being fed into the generator. This consistent transforming process is vital

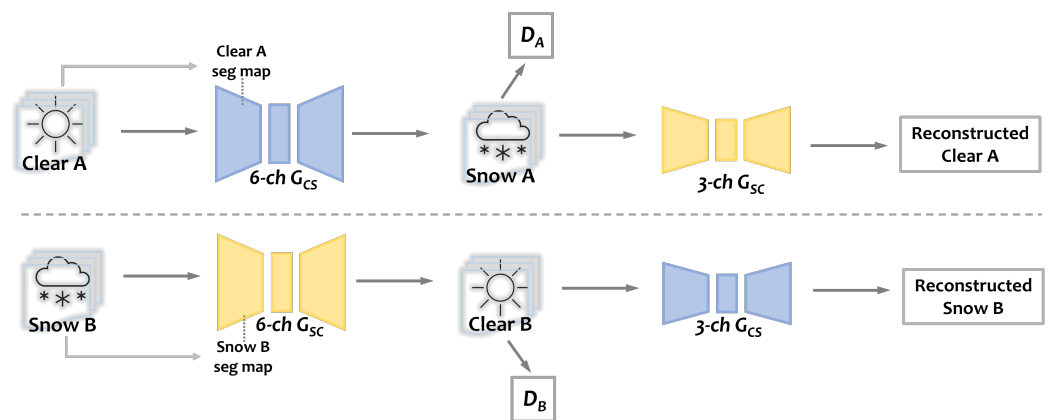
for aligning and rendering the datasets compatible for concatenation. Following this, a 6-channel dataset is constructed for each domain, with the first three channels comprising the original image and the latter three filled with the labels from the segmentation maps. This process is uniformly applied to both the clear and snowy datasets.

Incorporating a 4-ResNet-block structure based on the GAN structure [6,13], the input channels of the generator  $G$  are adjusted to accommodate the six-channel input. This modification facilitates the integration of the concatenated data. After the encoding–decoding phase within the generator, the output comprises snow-augmented point clouds,  $O_A$  and  $O_B$ , for each domain. These outputs retain the format of the original datasets, with the exception that the latter three channels, dedicated to labels, are subsequently pruned to revert the data back to the standard RGB depth image format. This results in the generation of the synthesized snow point cloud.

Throughout this entire fusion process, the segmentation map remains integrally connected with the raw data, imparting directional guidance at every stage of the conditional generative model. This method ensures optimal guidance with minimal supervision, effectively circumventing potential issues of deviation or overfitting, which could arise from the relatively lower data weight of the labels compared to the pixels. The efficiency of this early fusion approach is thus underscored.

### 3.3. Architecture and Loss Functions

Figure 5 denotes the overall architecture of the whole conditional generative model based on the CycleGAN [13] backbone. Clear A and Snow B along with their segmentation maps are the input data. The condition-guided conversions are conducted by 6-channel generators while the reconstructions are completed by 3-channel generators.  $D_A$  and  $D_B$  are discriminators that oversee the conversion between clear and snow, respectively.



**Figure 5.** Architecture of the condition-guided adverse effects augmentation model based on CycleGAN [13]. Clear A and Snow B along with their segmentation maps are the input data. The condition-guided conversions are conducted by 6-channel generators while the reconstructions are completed by 3-channel generators.  $D_A$  and  $D_B$  are discriminators.

#### 3.3.1. Custom Loss

We employed the same customized loss functions designated to LiDAR depth image transformation in our previous work [6], including a depth loss [35,36] and an SSIM loss [37], as shown in Equations (7)–(9):

$$\mathcal{L}_{depth} = \frac{1}{n} \sum_i (\hat{d}_i - d_i)^2 - \frac{\lambda_{depth}}{n^2} \left( \sum_i (\hat{d}_i - d_i) \right)^2 \quad (7)$$

where  $\hat{d}_i$  and  $d_i$  are the reconstructed and initial depth, respectively, while the hyperparameter  $\lambda_{depth}$  governs the scale invariance.  $\lambda_{depth}$  was assigned to 1 to achieve the full invariance:

$$SSIM(N, \hat{N}) = \frac{(2\mu_N\mu_{\hat{N}} + c_1)(2\sigma_{N\hat{N}} + c_2)}{(\mu_N^2 + \mu_{\hat{N}}^2 + c_1)(\sigma_N^2 + \sigma_{\hat{N}}^2 + c_2)} \quad (8)$$

$$\mathcal{L}_{ssim} = 1 - SSIM(\hat{N}, N) \quad (9)$$

where  $N$  is the normalized image tensor of the original real image,  $\hat{N}$  is the normalized image tensor of the reconstructed image,  $\mu_{\hat{N}}$  is the average of  $\hat{N}$ ,  $\mu_N$  is the average of  $N$ ,  $\sigma_{\hat{N}}^2$  is the variance of  $\hat{N}$ ,  $\sigma_N^2$  is the variance of  $N$ ,  $\sigma_{\hat{N}N}$  is the covariance of  $\hat{N}$  and  $N$ , and  $c_1$  and  $c_2$  are two variables to stabilize the division with a weak denominator.

The depth loss regulates the scales of any generated point cloud while the SSIM loss ensures that the overall structure of the point cloud remains unchanged. Together, these components constitute the custom loss, as specified in Equation (10):

$$\mathcal{L}_{custom} = \lambda_d \mathcal{L}_{depth} + \lambda_s \mathcal{L}_{ssim} \quad (10)$$

where  $\lambda_d$  was set to 10, the same as the default settings in GAN loss, and  $\lambda_s$  was set to 0.1 because too high an SSIM loss might lock the model into overfitting.

### 3.3.2. Identity Loss

In our model's endeavor to augment snow effects onto clear point clouds, a key focus is placed on preserving the intrinsic structural integrity of the input data. To achieve this, besides the custom loss functions, we incorporated an approach similar to identity loss, a mechanism that ensures the augmented output retains the fundamental characteristics of the original point cloud. This aspect is crucial for maintaining realism and accuracy in the synthesized snowy scenes. The identity loss function is shown below:

$$\mathcal{L}_{id} = \| G_{SC}(I_C) - I_C \| + \| G_{CS}(I_S) - I_S \| \quad (11)$$

Here,  $G_{SC}$  represents a generator transforming clear to snowy, and  $G_{CS}$  represents a generator transforming snowy to clear.  $I_C$  and  $I_S$  are input point clouds from the clear and snowy domains, respectively.

Central to this methodology is the strategic input of the target dataset into the model. By feeding the clear point clouds as inputs in their original form, the model learns a mapping that minimizes alterations to the inherent structure of these point clouds. The identity loss function acts as a regulatory mechanism, guiding the model to respect and preserve the original data's topology and spatial configuration.

During the training phase, the model processes the unaltered clear point clouds alongside the primary task of generating snow-augmented point clouds. This dual processing enables the model to compare the output against the original input, ensuring that the augmentation process does not compromise the essential structural elements. The identity loss quantifies the deviation of the augmented output from the original structure, pushing the model to generate outputs where the addition of snow effects is seamless and natural, without distorting the underlying point cloud architecture.

The integration of identity loss in our model serves to ensure structural fidelity, as it maintains the geometric and spatial characteristics of the original point cloud. It enhances the realism of the augmentation by constraining the model to respect the original data's structure, aligning the snow effects with real-world dynamics. Furthermore, this approach enhances the model's robustness, reducing the risk of overfitting to specific snow patterns or anomalies in the training data and thereby improving its generality.

### 3.3.3. Overall Loss Function

Upon integrating the conventional GAN adversarial losses between the clear and snowy data and the cycle consistency loss [13], we arrive at the comprehensive objective loss:

$$\begin{aligned} \mathcal{L}(G_{SC}, G_{CS}, D_A, D_B) = & \lambda_g \mathcal{L}_{GAN}(G_{SC}, D_A, I_S, I_C) \\ & + \lambda_g \mathcal{L}_{GAN}(G_{CS}, D_B, I_C, I_S) \\ & + \lambda_{cyc} \mathcal{L}_{cyc} + \lambda_{cus} \mathcal{L}_{custom} + \lambda_{id} \mathcal{L}_{id} \end{aligned} \quad (12)$$

where  $\lambda_g$ ,  $\lambda_{cyc}$ ,  $\lambda_{cus}$ , and  $\lambda_{id}$  denote the weight coefficients of adversarial losses, cycle consistency loss, custom loss, and identity loss, respectively. The higher the weight, the larger the influence the corresponding loss function has on the model. During the training process, the parameters were set as follows to achieve the best performance:  $\lambda_g = 10$ ,  $\lambda_{cyc} = 40$ ,  $\lambda_{cus} = 1$ , and  $\lambda_{id} = 0.5$ .

### 3.4. Violations and Solutions in LiDAR Data Augmentation

Violations concerning LiDAR sensor properties are a frequent occurrence in data augmentation, and various solutions have been proposed to address them. LiDAR sensors perceive the world in a polar coordinate system, with the sensor at the center, as opposed to a Cartesian coordinate system. Initially, this difference was not adequately considered in scene transformations, because added noise was uniformly distributed throughout the scene [11]. However, as demands for accuracy increased, the physical characteristics of noise points began to be incorporated. For instance, a normal distribution was used for arranging synthetic snow points to more accurately reflect their physical properties [11], but the violation remains. A more refined approach involves resampling the entire dataset in a resolution that corresponds to the sensor's horizontal turning rate and the number of vertical channels. This method, which aligns with the depth image approach used in our research [6,34], has been validated for its effectiveness [12]. Such resampling ensures that the augmented data more accurately mirror the way LiDAR sensors capture and interpret the world, leading to more realistic and useful augmentation outcomes.

Occlusion poses a potential risk in point cloud data augmentation, where the addition of extra points could obscure the trace of an original signal, rendering the original point's presence illogical. We offer a perspective on spherical coordinate filtering to address this problem. By converting the LiDAR point cloud's Cartesian coordinates ( $x, y, z$ ) into spherical coordinates ( $\theta, \phi, r$ ), it becomes feasible to identify points sharing the same angular coordinates ( $\theta, \phi$ ) but differing in the radial distance ( $r$ ). Points with identical angular coordinates but larger radial distances would be recognized as occluded by the artificially added points. This conversion and subsequent filtering ensure an accurate representation of the occlusion effects caused by augmented data.

## 4. Experiments and Results

Training, testing, and data processing were executed using the Pytorch framework. Our model incorporates four ResNet residual blocks within the generator, optimized for training on two NVIDIA RTX 3090Ti graphics cards. We set a batch size of four, a decision influenced by the complexity inherent in the conditional generative models. A linearly declining learning rate schedule starting from 0.02 was employed until the model reached convergence.

### 4.1. Reproduction of Real Adverse Conditions

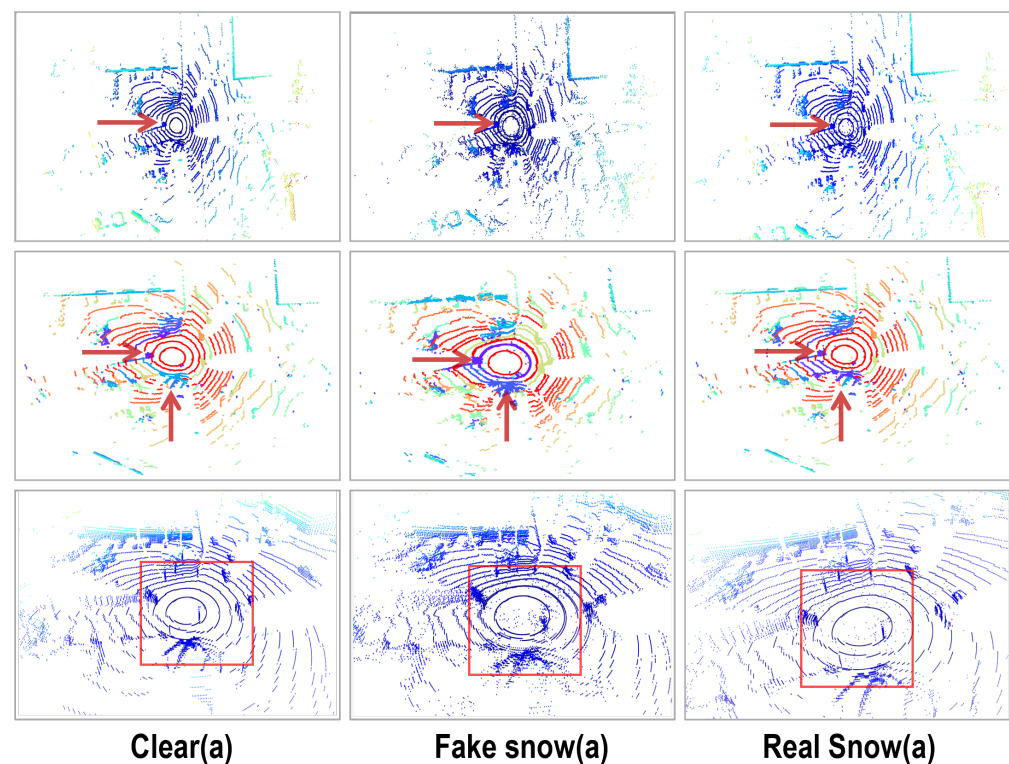
We carried out experiments using our trained model on the CADC dataset, which was split in a 6:1:3 ratio for training, validation, and testing purposes. With the available ground truth provided by the CADC dataset, our goal was to replicate the adverse effects within this dataset and assess the model's generation capabilities.

#### 4.1.1. Qualitative Results

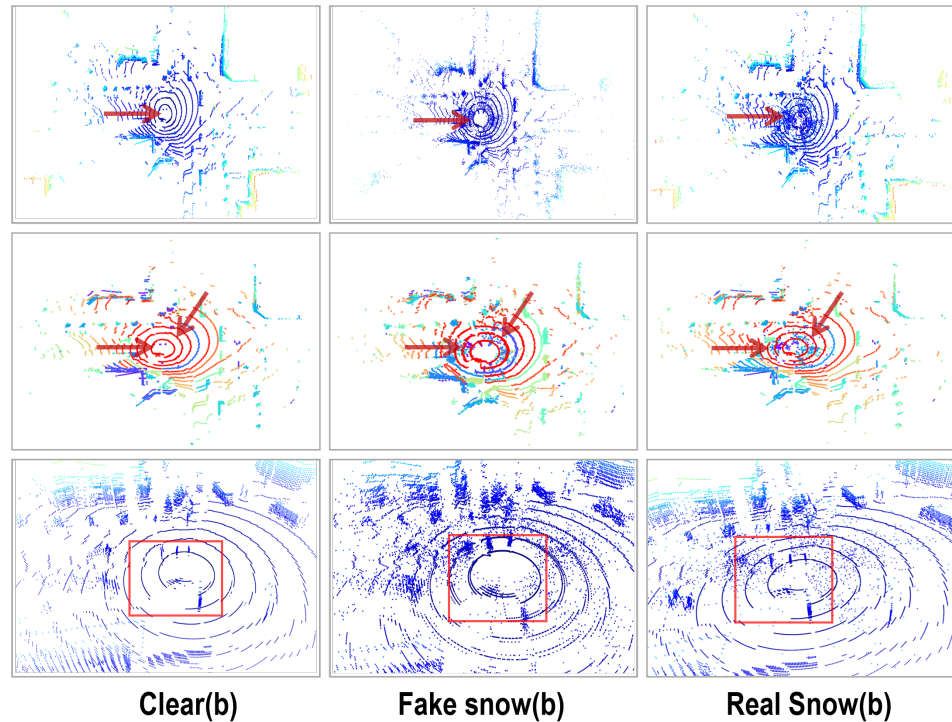
Figures 6 and 7 present the adverse effect reproduction of the CADC dataset based on the early fusion—conditional generative model. For sets (a) and (b), the clear data, fake snow generation from our model, and the real snow from CADC are placed at the

left, middle, and right columns, respectively. Each of the scenarios features an overall BEV in the top row; the clustered results produced by the 3D clustering algorithm show the changes in snow clusters in the middle row; and the bottom rows show magnified third-person views of the point cloud's central region, where the ego vehicle is situated.

Observations from the areas highlighted by red arrows and enclosed within red boxes indicate a notable reproduction of scattered snow features. Snow clusters, particularly those in purple and dark blue around the ego vehicle, appear enhanced. A detailed examination of the clustering in Figure 6 reveals that the synthesized snow clusters are denser and more extensive than in the actual scenarios. However, in Figure 7, the synthetic clusters do not extend across the central region as broadly as in the real case. This outcome is attributed to the early fusion method's characteristic of capturing both snow clusters and objects via segmentation maps. As a result, while the model efficiently replicates snow clusters, it simultaneously accentuates relevant objects with minimal interference from surrounding minor snow clusters. This effect is evident in the bottom row of Figure 7, where an area with multiple pedestrians is shown to be enlarged. Here, snow clusters are generated without significantly dispersing the signal on pedestrians, but rather enhancing it. In summary, the model effectively replicates adverse weather effects in the CADC dataset while ensuring that pertinent objects remain clearly identifiable, maintaining the integrity of their structure.



**Figure 6.** Set (a) augmentation results in the Canadian driving scenario. First row—BEV scenes, colored by height; middle row—clustered results, colored by cluster groups; bottom row—enlarged third-person view center part around the ego vehicle, colored by height. Red boxes and arrows—locations where snow's effects are reproduced.



**Figure 7.** Set (b) augmentation results in the Canadian driving scenario. First row—BEV scenes, colored by height; middle row—clustered results, colored by cluster groups; bottom row—enlarged third-person view center part around the ego vehicle, colored by height. Red boxes and arrows—locations where snow’s effects are reproduced.

#### 4.1.2. Quantitative Results and Ablation Study

To assess the effectiveness of the proposed conditional generative model in synthesizing adverse effects, particularly in comparison with other methods, a precision–recall analysis was conducted along with an ablation study on the model without the conditional guide, as outlined in Equations (13) and (14). Similar to the previous approach, the point clouds were analyzed on a cluster basis. This analysis, focusing on replicating snow effects within the CADC dataset, bases its calculations on the CADC ground truth to ensure reliability in the generative model performance evaluation.

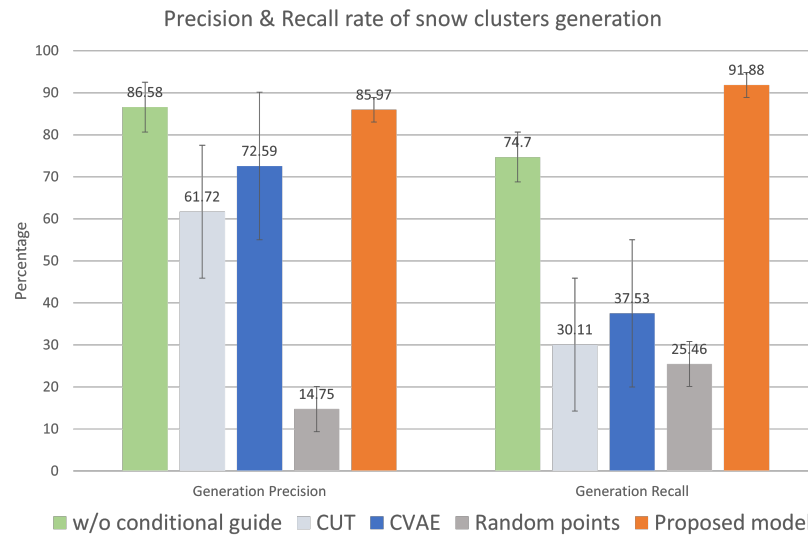
For this evaluation, a subset of samples from the test dataset, which had not been included in training, was selected without bias. Each cluster in these samples underwent manual labeling to identify the presence of snow clusters. The parameters for all comparative methods were uniformly set, ensuring equal treatment in parameter application across all methods under evaluation. This consistent parameter setting allows for a fair comparison of each method’s ability to generate realistic adverse effects while preserving the integrity of the driving scene:

$$\text{Precision} = \frac{\text{generated clusters} \cap \text{labeled as snow}}{\text{total generated clusters}} \quad (13)$$

$$\text{Recall} = \frac{\text{generated clusters} \cap \text{labeled as snow}}{\text{total labeled snow clusters in CADC}} \quad (14)$$

Looking at Figure 8, it is evident that both the model without a conditional guide and the Contrastive Unpaired Translation (CUT) [38] model struggle to accurately reproduce snow clusters found in the original dataset. The F1 score of CUT is only 0.405, and this challenge stems from their limited understanding of adverse effects. Intriguingly, the precision of the model without a conditional guide slightly exceeds the proposed model by less than 1%. This minor advantage primarily arises from its handling of occluded objects, a residual aspect in this model. As the proposed model endeavors to regenerate snow

clusters, it inadvertently reconstructs or augments some partially composed non-snow objects, leading to a precision just shy of 90%, but an apparent improvement in recall rate, which can be further affirmed by the higher F1 score (0.888) compared to that of the model without a conditional guide (0.802).



**Figure 8.** Precision and recall rates comparisons of adverse effects generation based on snow clusters.

Conversely, the CVAE model, despite similar conditional guidance, fails to effectively address driving scenarios. Its focus is predominantly on generating snow clusters around the vehicle's immediate vicinity, neglecting the broader scene. This results in a significant loss of the original point structure and tiny snow clusters in other areas, contributing to its notably low recall rate with an F1 score of 0.495. As for randomly generated points with a normal distribution, the 3D clustering algorithm has a hard time identifying distinct clusters, leading to a substantial increase in noise with a corresponding decrease in cluster identification, as reflected in the 14.75% precision rate and 0.187 F1 score.

In addition, the proposed model, with the additional segmentation map, shows a modest increase in inference time per frame (137.6 ms) compared to 118.1 ms for the basic model (without a conditional guide) and 74.1 ms for CUT, due to the increased input channels and added computational complexity. Despite this increase, the performance remains within acceptable limits for real-time applications in autonomous driving, where a sub-150 ms inference time per frame is considered practical [39].

In summary, the proposed condition-guided generative model stands out with the highest recall in reproducing snow effects from the CADC dataset, achieving this with satisfactory precision. This underscores the efficacy of the proposed model in generating realistic adverse weather conditions for LiDAR point clouds.

#### 4.1.3. Detection Rate Improvement

In order to evaluate the improvements of proposed data augmentation in perception enhancement, we utilize the 3D-object-detection metrics from the KITTI evaluation framework [40]. In line with [41], we report the average precision (AP) across 40 recall positions to ensure a balanced comparison. Our investigation focuses on the impact of our augmentation scheme for adverse conditions on two widely used 3D-object-detection methods [42,43]. We assess our approach against a baseline model without any augmentation, and we then consider the results of using augmentation data generated by our proposed model. In addition, we compare these results with those obtained from de-noising point clouds using the DROR method [3].

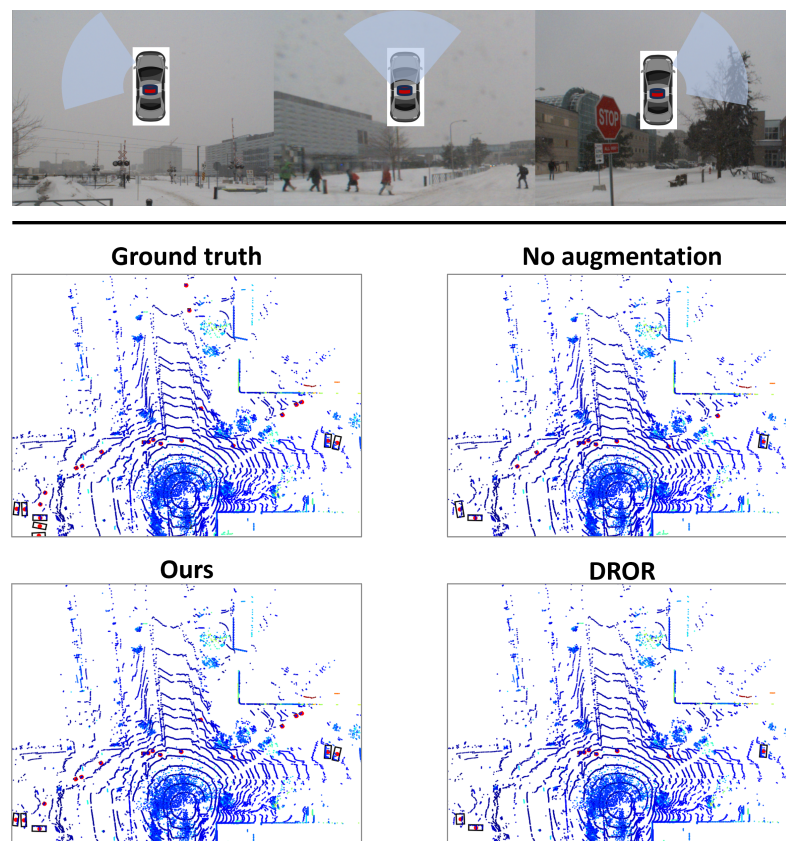
The detection rates are detailed in Table 1. The 3D average precision (AP) is reported on baseline (no augmentation) and our proposed augmentation model with 2691 samples.

Augmentation with DROR filters is also provided for reference. A key observation from this data is that our comprehensive augmentation, featuring both scattered and clustered adverse effects, markedly enhances performance in the most challenging test scenario, specifically during fierce snowfall. This enhancement is significant when compared to both the baseline approach and the de-noising filter. Here, our complete augmentation outperforms the baseline by a noteworthy 2.46% increase in AP.

**Table 1.** Comparison of augmentation methods for 3D object detection in snowfall on CADC. Bold numbers indicate highest results.

Detection Method	PV-RCNN [42]			SECOND [43]		
	None	DROR	Ours	None	DROR	Ours
Augmentation method						
3D average precision (AP)	43.11	38.69	<b>45.57</b>	37.08	35.31	<b>38.23</b>

An example of visual results showing the data augmentation scheme is presented in Figure 9. The top row displays a composite of RGB images from three cameras, oriented left-front, directly ahead, and right-front, collectively representing a 180° frontal view of the ego vehicle. The subsequent BEV point clouds compare detection outcomes using our augmentation method and the DROR filter against a baseline with no augmentation, alongside the ground truth. Pedestrians are denoted in red dots while cars (and trucks) are denoted in black bounding boxes with red dots in the center.



**Figure 9.** Qualitative comparison of detection results on samples from CADC containing fierce adverse conditions. The top row shows the corresponding forward 180° RGB images. The rest show the LiDAR point clouds with ground-truth boxes and predictions using the baseline (“no augmentation”), our augmentation, and DROR. Red dots denote pedestrians, and black boxes with red dots in the center denote cars (or trucks). Point cloud colors encoded in height.

From the RGB images and ground truth, we observe a scenario with many moving pedestrians and parked cars under heavy snow. The baseline model, lacking augmentation data, struggles to detect pedestrians and cars beyond a certain distance or those partially obscured, and often fails to accurately gauge the gesture and dimensions of detected objects. In contrast, our augmented data method enables a more precise detection of cars and pedestrians, particularly those nearer to the ego vehicle, with accurate parameters. As for the augmentation scheme with DROR filters, the performance turns out to be worse than the baseline. This is attributed to the removal of critical points necessary for object detection.

However, it is important to note that two undetected parked cars at the bottom left and two undetected pedestrians far ahead, as shown in the ground truth, are only observed through camera assistance due to their minimal LiDAR signal presence. This is a common issue in adverse condition datasets and partly explains the generally low average detection rate observed in Table 1.

There is a car near the rear of the ego vehicle in the ground truth, barely noticeable due to snow swirl occlusion, which none of the methods detect due to severe occlusion. This might hint at the limitation of current learning-based perception improvement methods and suggest the need for advancements in sensor hardware to further overcome such challenges.

#### 4.2. Synthetic Adverse Conditions

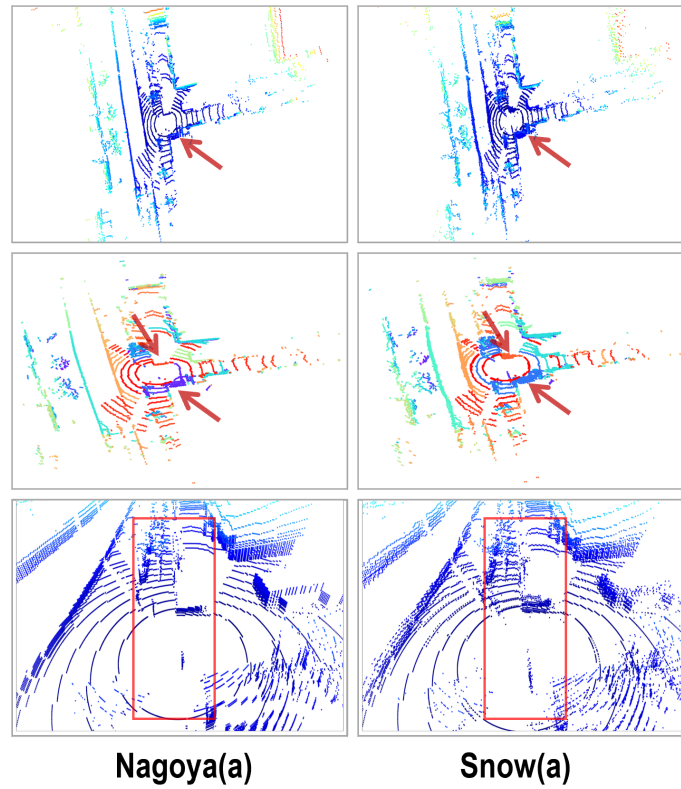
In the other experiment conducted on the LIBRE Nagoya dataset [19], our objective was to evaluate the model's performance in overcoming domain gaps. The pre-trained model based on CADC was tested on the 6000 frames of the Nagoya dataset, which was collected under clear conditions in the urban area of Nagoya, Japan. This selection is particularly representative due to the significant domain differences between Canada and Japan, especially regarding scenario layouts and traffic behaviors. The CADC dataset predominantly features suburban environments with sparse buildings and abundant vegetation, a contrast to the urban settings commonly found in the Nagoya dataset.

##### 4.2.1. Qualitative Results

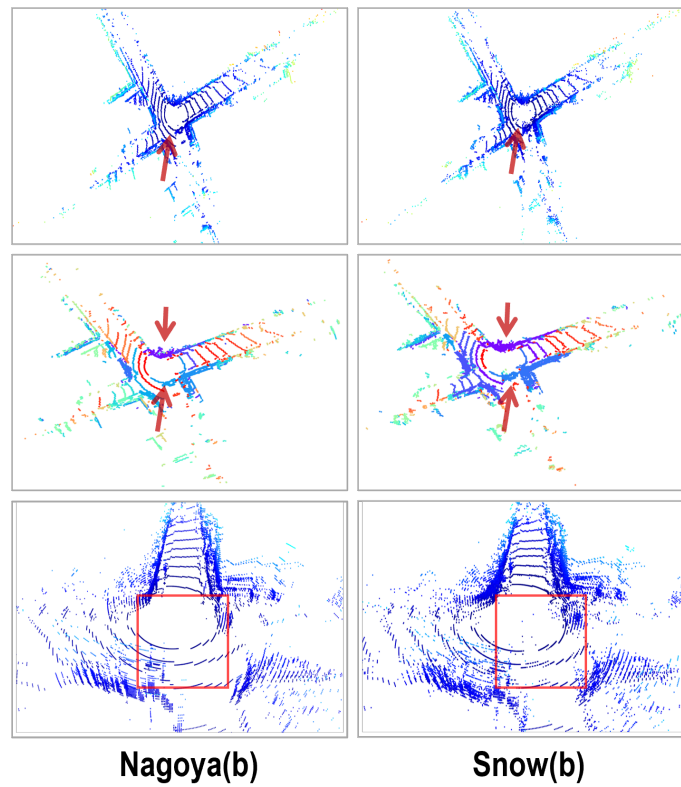
Figures 10 and 11 present the adverse effect synthesis on the Nagoya dataset where a domain gap exists. The clear data and the synthetic snow generation from our model are placed in the left and right columns, respectively. For sets (a) and (b), each of the scenarios features an overall BEV in the top row; the clustered results show the changes in snow clusters in the middle row; and the bottom rows show magnified third-person views of the point cloud's central region, where the ego vehicle is situated.

The red arrows in the figures highlight the areas where snow clusters were generated. Upon examining these indicated locations, it is evident that a number of snow clusters formed near the central area, resembling typical patterns observed in snow-affected driving scenarios. However, it is important to note that the level of generated snow is not an exact match for the snow conditions found in the CADC dataset, particularly in terms of the density of the purple and dark blue clusters depicted in the middle rows. This discrepancy partly arises from our deliberate decision to reduce the extent of snow generation to avoid generating artifacts. Additionally, the model operates across a significant domain gap, characterized by different environmental layouts and traffic patterns, which necessitates a more cautious approach.

As a result, and as illustrated by the red boxes, the snow generation in the vicinity of the ego vehicle and around the sidewalks predominantly features scattered snow points rather than extensive clusters. This cautious approach is particularly appropriate given the Nagoya dataset's inherent composition, which includes numerous small-sized clusters, often resulting from city greenbelts and occlusions. Therefore, generating snow in a more restrained manner is the most effective strategy to maintain the original structural integrity of the dataset while successfully synthesizing quasi-natural adverse effects.



**Figure 10.** Set (a) augmentation results in the Nagoya driving scenario. Red boxes and arrows—locations where adverse effects are synthesized.



**Figure 11.** Set (b) augmentation results in the Nagoya driving scenario. Red boxes and arrows—locations where adverse effects are synthesized.

#### 4.2.2. 3D Clustering Results for Nagoya Snow Synthesis

Since the Nagoya dataset lacks ground truth, the assessment will provide 3D clustering results instead of precision and recall. These evaluations were essential for a quantitative assessment of the model's effectiveness in generating synthetic snow on a clean dataset. The metrics reported in the subsequent results sections represent average values derived from these samples.

Table 2 offers a quantitative comparison between the original Nagoya dataset and its augmented version with synthesized snow, demonstrating the effects of the snow synthesis process. Below is a detailed analysis of each metric.

The synthesized snow dataset shows a notable increase in the number of noise points, rising from 1204.67 in the original Nagoya dataset to 2631.46. This increase is indicative of the additional complexity introduced by the synthetic snow, representing a more challenging scenario for processing and analysis. This statistic is one of the most direct indicators of a successful adverse effect synthesis.

**Table 2.** The 3D clustering metrics (avg.) comparisons between the original Nagoya dataset and synthesized snow in Nagoya.

Items	Nagoya	Synthesized Snow
Noise number	1204.67	2631.46
Cluster number	480.25	1073.52
Reachability distance	0.2470	0.3952
Inter-cluster distance	59.30	49.45
Size of clusters	28.2638	13.3046
Davies–Bouldin index	2.3653	4.4149
Silhouette score	−0.2170	−0.2927

The cluster number more than doubles in the synthesized dataset (1073.52) compared to the original dataset (480.25) and is a direct consequence of the snow synthesis process. The additional clusters likely represent the snowflakes or snow clusters, providing a more realistic representation of a snowy environment.

Reachability distances in the synthesized dataset show an increase to 0.3952 from 0.2470 in the original dataset. This change suggests that the introduction of snow creates a more scattered distribution of data points, emulating the dispersed nature of snowfall and its impact on the visibility and distinguishability of objects within the dataset.

The average inter-cluster distance decreases in the synthesized dataset (49.45) compared to the original dataset (59.30). This reduction might reflect the additional clusters formed due to snow, which are closer together, simulating the dense and overlapping nature of snow in the environment.

There is a notable decrease in the average size of clusters in the synthesized data (13.3046) compared to the original dataset (28.2638). This decrease can be attributed to the synthesized snow creating smaller, more numerous clusters, which is consistent with the physical characteristics of snow affecting spatial data.

The silhouette score drops slightly in the synthesized dataset (−0.2927) compared to the original (−0.2170), suggesting a decrease in the separation distance between neighboring clusters due to the added snow.

The Davies–Bouldin index increases in the synthesized dataset (4.4149) compared to the original (2.3653), indicating less compact but more separated clusters, a likely result of snow altering the spatial relationships within the data.

Overall, these metrics demonstrate the significant impact of synthesized snow on the Nagoya dataset, successfully adding quasi-natural adverse effects to the scenario, which is crucial for testing and improving algorithms in snowy conditions.

## 5. Conclusions

In this paper, we introduced an innovative approach for augmenting adverse weather condition data in autonomous driving applications. Our model leverages conditional guides to generate natural adverse effects in LiDAR point cloud data, providing a much-needed solution to the scarcity of diverse weather conditions in existing datasets. This scarcity often hampers the development of robust perception systems for autonomous vehicles.

Our method's key innovation lies in its ability to produce realistic adverse weather conditions within point cloud data by generating segmentation maps to guide data augmentation effectively. By replicating conditions like snow and rain with high fidelity, our model enables the training of perception systems to handle real-world adverse weather scenarios better.

The effectiveness of our approach was demonstrated through significant improvements in detection performance compared to existing baseline models. Our experiments highlighted that the model consistently outperforms state-of-the-art filters and augmentation methods in terms of accuracy and precision, particularly in challenging scenarios involving intense adverse weather conditions. Moreover, the proposed model exhibited remarkable capabilities, accurately and naturally recreating over 90% of the adverse effects. This facilitated a considerable boost in the performance and accuracy of deep learning algorithms for autonomous driving, especially under adverse weather scenarios. Notably, when employing our augmented approach, we observed a 2.46% increase in the 3D average precision metric, indicating a substantial enhancement in detection accuracy and system reliability. These quantitative improvements in 3D object detection, compared to models trained without augmentation, validate the efficacy of our proposed method.

However, it is important to acknowledge that our current work primarily derives its adverse condition data from a single dataset. While the CADC dataset has provided a robust foundation for our initial investigations, particularly with snow as a test subject, the diversity of adverse conditions experienced globally suggests a need for a broader data source. Moving forward, a key area of development will be to extend our data sources beyond the singular dataset.

**Author Contributions:** Conceptualization, Y.Z. (Yuxiao Zhang) and H.Y.; methodology, Y.Z. (Yuxiao Zhang) and Y.N.; formal analysis, M.D. and K.O.; manuscript preparation, Y.Z. (Yuxiao Zhang); review and editing, M.G. and C.Z.; supervision, K.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Automotive Lightweight Engineering Technology Center of the Liaoning University of Technology.

**Data Availability Statement:** The CADC dataset used in this research is publically available at <http://cadcd.uwaterloo.ca/> (accessed on 18 June 2024). Details and availability of the LIBRE Nagoya dataset used in this research can be found at <https://sites.google.com/g.sp.m.is.nagoya-u.ac.jp/libre-dataset> (accessed on 18 June 2024).

**Acknowledgments:** The authors thank Alexander Carballo from Gifu University for providing the LIBRE dataset.

**Conflicts of Interest:** Author Ming Ding is employed by the Zhejiang Fubang Technology Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

LiDAR	Light detection and ranging
GAN	Generative Adversarial Networks
CADC	Canadian Adverse Driving Conditions
OPTICS	Ordering points to identify the clustering structure

ICD	Inter-cluster distance
SSIM	Structural Similarity Index
DROR	Dynamic Radius Outlier Removal
DSOR	Dynamic Statistical Outlier Removal
PGM	Polar Grid Map
DBI	Davies–Bouldin index
BEV	Bird Eye View
CUT	Contrastive Unpaired Translation
AP	Average precision

## References

- Zhang, Y.; Carballo, A.; Yang, H.; Takeda, K. Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. *ISPRS J. Photogramm. Remote Sens.* **2023**, *196*, 146–177. [[CrossRef](#)]
- Jokela, M.; Kuttila, M.; Pyykönen, P. Testing and validation of automotive point-cloud sensors in adverse weather conditions. *Appl. Sci.* **2019**, *9*, 2341. [[CrossRef](#)]
- Charron, N.; Phillips, S.; Waslander, S.L. De-noising of Lidar point clouds corrupted by snowfall. In Proceedings of the Conference on Computer and Robot Vision (CRV), Toronto, ON, Canada, 9–11 May 2018; pp. 254–261.
- Le, M.H.; Cheng, C.H.; Liu, D.G. An Efficient Adaptive Noise Removal Filter on Range Images for LiDAR Point Clouds. *Electronics* **2023**, *12*, 2150. [[CrossRef](#)]
- Bergius, J. LiDAR Point Cloud De-Noising for Adverse Weather. Ph.D. Thesis, Halmstad University, Halmstad, Sweden, 2022.
- Zhang, Y.; Ding, M.; Yang, H.; Niu, Y.; Feng, Y.; Ohtani, K.; Takeda, K. L-DIG: A GAN-Based Method for LiDAR Point Cloud Processing under Snow Driving Conditions. *Sensors* **2023**, *23*, 8660. [[CrossRef](#)] [[PubMed](#)]
- Hahner, M.; Sakaridis, C.; Bijelic, M.; Heide, F.; Yu, F.; Dai, D.; Van Gool, L. Lidar snowfall simulation for robust 3d object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 16364–16374.
- Heinzler, R.; Piewak, F.; Schindler, P.; Stork, W. CNN-based lidar point cloud de-noising in adverse weather. *IEEE Robot. Autom. Lett.* **2020**, *5*, 2514–2521. [[CrossRef](#)]
- Rasshofer, R.H.; Spies, M.; Spies, H. Influences of weather phenomena on automotive laser radar systems. *Adv. Radio Sci.* **2011**, *9*, 49–60. [[CrossRef](#)]
- Wallace, A.M.; Halimi, A.; Buller, G.S. Full waveform lidar for adverse weather conditions. *IEEE Trans. Veh. Technol.* **2020**, *69*, 7064–7077. [[CrossRef](#)]
- Guo, A.; Feng, Y.; Chen, Z. LiRTTest: Augmenting LiDAR point clouds for automated testing of autonomous driving systems. In Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis, Virtual, 18–22 July 2022; pp. 480–492.
- Piroli, A.; Dallabetta, V.; Walessa, M.; Meissner, D.; Kopp, J.; Dietmayer, K. Robust 3D Object Detection in Cold Weather Conditions. In Proceedings of the 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 5–9 June 2022; pp. 287–294.
- Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
- Yang, H.; Carballo, A.; Takeda, K. Disentangled Bad Weather Removal GAN for Pedestrian Detection. In Proceedings of the 2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring), Helsinki, Finland, 19–22 June 2022; pp. 1–6.
- Jaw, D.W.; Huang, S.C.; Kuo, S.Y. DesnowGAN: An efficient single image snow removal framework using cross-resolution lateral connection and GANs. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *31*, 1342–1350. [[CrossRef](#)]
- Sallab, A.E.; Sobh, I.; Zahran, M.; Essam, N. LiDAR Sensor modeling and Data augmentation with GANs for Autonomous driving. *arXiv* **2019**, arXiv:1905.07290.
- Sobh, I.; Amin, L.; Abdelkarim, S.; Elmadawy, K.; Saeed, M.; Abdeltawab, O.; Gamal, M.; El Sallab, A. End-to-end multi-modal sensors fusion system for urban automated driving. In Proceedings of the NIPS Workshop on Machine Learning for Intelligent Transportation Systems, Montreal, QC, Canada, 3–8 December 2018.
- Lee, J.; Shiotsuka, D.; Nishimori, T.; Nakao, K.; Kamijo, S. GAN-Based LiDAR Translation between Sunny and Adverse Weather for Autonomous Driving and Driving Simulation. *Sensors* **2022**, *22*, 5287. [[CrossRef](#)] [[PubMed](#)]
- Carballo, A.; Lambert, J.; Monrroy, A.; Wong, D.; Narksri, P.; Kitsukawa, Y.; Takeuchi, E.; Kato, S.; Takeda, K. LIBRE: The multiple 3D LiDAR dataset. In Proceedings of the Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 1094–1101.
- Von Bernuth, A.; Volk, G.; Bringmann, O. Simulating photo-realistic snow and fog on existing images for enhanced CNN training and evaluation. In Proceedings of the Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 27–30 October 2019; pp. 41–46.
- Zhang, K.; Li, R.; Yu, Y.; Luo, W.; Li, C. Deep Dense Multi-Scale Network for Snow Removal Using Semantic and Depth Priors. *IEEE Trans. Image Process.* **2021**, *30*, 7419–7431. [[CrossRef](#)] [[PubMed](#)]

22. Uříčář, M.; Sistu, G.; Rashed, H.; Vobecky, A.; Kumar, V.R.; Krizek, P.; Burger, F.; Yogamani, S. Let's Get Dirty: GAN Based Data Augmentation for Camera Lens Soiling Detection in Autonomous Driving. In Proceedings of the Winter Conference on Applications of Computer Vision (WACV), Virtual, 5–9 January 2021; pp. 766–775.
23. Chen, Z.; Wang, Y.; Yang, Y.; Liu, D. PSD: Principled Synthetic-to-Real Dehazing Guided by Physical Priors. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 7180–7189.
24. Bijelic, M.; Gruber, T.; Mannan, F.; Kraus, F.; Ritter, W.; Dietmayer, K.; Heide, F. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 11682–11692.
25. Kurup, A.; Bos, J. DSOR: A Scalable Statistical Filter for Removing Falling Snow from LiDAR Point Clouds in Severe Winter Weather. *arXiv* **2021**, arXiv:2109.07078.
26. Pitropov, M.; Garcia, D.E.; Rebello, J.; Smart, M.; Wang, C.; Czarniecki, K.; Waslander, S. Canadian adverse driving conditions dataset. *Int. J. Robot. Res.* **2021**, *40*, 681–690. [[CrossRef](#)]
27. Ankerst, M.; Breunig, M.M.; Kriegel, H.P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod Rec.* **1999**, *28*, 49–60. [[CrossRef](#)]
28. El Yabroudi, M.; Awedat, K.; Chabaan, R.C.; Abudayyeh, O.; Abdel-Qader, I. Adaptive DBSCAN LiDAR Point Cloud Clustering For Autonomous Driving Applications. In Proceedings of the 2022 IEEE International Conference on Electro Information Technology (eIT), Mankato, MN, USA, 19–21 May 2022; pp. 221–224.
29. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd-Proceedings*; AAAI Press: Washington, DC, USA, 1996; Volume 96, pp. 226–231.
30. Schubert, E.; Sander, J.; Ester, M.; Kriegel, H.P.; Xu, X. DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Trans. Database Syst. (TODS)* **2017**, *42*, 19. [[CrossRef](#)]
31. Jain, A.K.; Murty, M.N.; Flynn, P.J. Data clustering: A review. *ACM Comput. Surv. (CSUR)* **1999**, *31*, 264–323. [[CrossRef](#)]
32. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
33. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *PAMI-1*, 224–227. [[CrossRef](#)]
34. Zhang, Y.; Ding, M.; Yang, H.; Niu, Y.; Feng, Y.; Ge, M.; Carballo, A.; Takeda, K. LiDAR Point Cloud Translation Between Snow and Clear Conditions Using Depth Images and GANs. In Proceedings of the 2023 IEEE Intelligent Vehicles Symposium (IV), Anchorage, AK, USA, 4–7 June 2023; pp. 1–7.
35. Mertan, A.; Duff, D.J.; Unal, G. Single image depth estimation: An overview. *Digit. Signal Process.* **2022**, *123*, 103441. [[CrossRef](#)]
36. Eigen, D.; Puhersch, C.; Fergus, R. Depth map prediction from a single image using a multi-scale deep network. In Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014), Montreal, QC, Canada, 8–13 December 2014; Volume 27.
37. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
38. Park, T.; Efros, A.A.; Zhang, R.; Zhu, J.Y. Contrastive Learning for Unpaired Image-to-Image Translation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
39. Betz, T.; Karle, P.; Werner, F.; Betz, J. An analysis of software latency for a high-speed autonomous race car—A case study in the indy autonomous challenge. *SAE Int. J. Connect. Autom. Veh.* **2023**, *6*, 283–296. [[CrossRef](#)]
40. Geiger, A.; Lenz, P.; Urtasun, R. Are we ready for autonomous driving? the kitti vision benchmark suite. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3354–3361.
41. Simonelli, A.; Buló, S.R.; Porzi, L.; López-Antequera, M.; Kotschieder, P. Disentangling monocular 3d object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1991–1999.
42. Shi, S.; Guo, C.; Jiang, L.; Wang, Z.; Shi, J.; Wang, X.; Li, H. PV-RCNN: Point-voxel feature set abstraction for 3d object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10529–10538.
43. Yan, Y.; Mao, Y.; Li, B. Second: Sparsely embedded convolutional detection. *Sensors* **2018**, *18*, 3337. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.