UniMix: Towards Domain Adaptive and Generalizable LiDAR Semantic Segmentation in Adverse Weather

Haimei Zhao[1*]   Jing Zhang[1*]   Zhuo Chen[2]   Shanshan Zhao[1]   Dacheng Tao[3]

[1]The University of Sydney   [2]Tsinghua University   [3] Nanyang Technological University

{hzha7798, jing.zhang1}@sydney.edu.au, z-chen17@mails.tsinghua.edu.cn,

{sshan.zhao00, dacheng.tao}@gmail.com

Abstract

LiDAR semantic segmentation (LSS) is a critical task in autonomous driving and has achieved promising progress. However, prior LSS methods are conventionally investigated and evaluated on datasets within the same domain in clear weather. The robustness of LSS models in unseen scenes and all weather conditions is crucial for ensuring safety and reliability in real applications. To this end, we propose UniMix, a universal method that enhances the adaptability and generalizability of LSS models. UniMix first leverages physically valid adverse weather simulation to construct a Bridge Domain, which serves to bridge the domain gap between the clear weather scenes and the adverse weather scenes. Then, a Universal Mixing operator is defined regarding spatial, intensity, and semantic distributions to create the intermediate domain with mixed samples from given domains. Integrating the proposed two techniques into a teacher-student framework, UniMix efficiently mitigates the domain gap and enables LSS models to learn weather-robust and domain-invariant representations. We devote UniMix to two main setups: 1) unsupervised domain adaption, adapting the model from the clear weather source domain to the adverse weather target domain; 2) domain generalization, learning a model that generalizes well to unseen scenes in adverse weather. Extensive experiments validate the effectiveness of UniMix across different tasks and datasets, all achieving superior performance over state-of-the-art methods. The code will be released.

1. Introduction

LiDAR semantic segmentation (LSS) is a fundamental task of 3D scene understanding for autonomous driving, aiming to assign a semantic label for each 3D point. The progress in deep learning techniques, coupled with the availability of large-scale datasets, has propelled LSS methods to de-
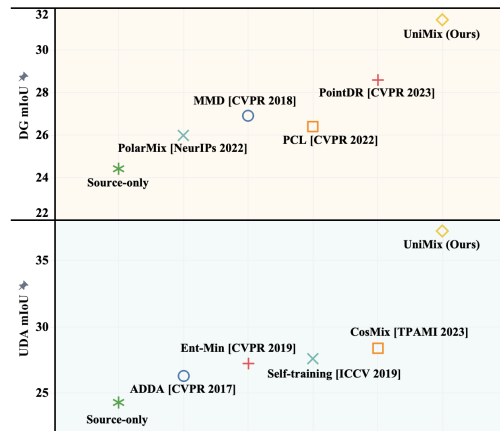

Figure 1. UniMix outperforms SOTA methods in both UDA and DG tasks, using SemanticKITTI [2] as the source and Semantic-STF [49] in all four weather conditions as the target.

liver promising results. However, traditional LSS models are typically trained and evaluated using data collected in clear weather conditions, assuming a consistent domain and lacking domain adaptive and generalizable ability. This limitation poses challenges in real-world applications where autonomous driving systems encounter diverse scenes and weather conditions, each characterized by distinct data distributions. Adverse weather, in particular, introduces variations in the spatial positions, intensity values, and semantic distributions of LiDAR point clouds [9, 29]. Models trained in ideal conditions often perform inadequately in adverse weather scenarios [49]. Therefore, the robust handling of unseen scenes in diverse weather is essential for ensuring the reliability and safety of autonomous driving.

While domain adaptation and generalization techniques have proven effective in 2D semantic segmentation [4, 12, 27], their application in LSS has primarily focused on adapting or generalizing between synthetic and real scenes or across different real-world scenes [35, 36]. This leaves a gap in the exploration for handling unseen scenes in adverse weather. To be specific, the domain discrepancy arises from

---

significant semantic distribution variations between datasets captured in different scenes, such as structural layout differences in various cities. Furthermore, adverse weather introduces spatial noise and impacts point cloud reflection intensity, resulting in shifts in data distribution—another form of domain discrepancy. The coexistence of these dual sources of domain discrepancy presents formidable challenges for adapting and generalizing existing LSS models.

To address these challenges, we introduce UniMix, a universal approach for learning weather-robust and domain-invariant representations, enabling the adaptation and generalization of LSS models from clear to adverse weather scenes. UniMix is structured as a two-stage framework, consisting of Source-to-Bridge and Bridge-to-Target stages, grounded in the well-established teacher-student framework [22, 35, 36]. First, we construct a Bridge Domain through physically realistic weather simulation on source domain data, generating data with the same scene characteristics but in diverse adverse weather conditions. We then introduce a Universal Mixing operator that blends point clouds from two given domains through spatial, intensity, and semantic mixing. These two techniques significantly mitigate the domain gap and enhance scene diversity during model training, thereby improving the model's adaptability and generalizability. Specifically, in the first Source-to-Bridge stage, even without access to target data, the model learns weather-robust representations from the mixed source and Bridge Domain data via Universal Mixing, enabling effective generalization to unseen adverse weather conditions. In the second Bridge-to-Target stage, with access to target data, the model further adapts to the target domain and learns domain-invariant representations from the mixed Bridge Domain and target domain data. UniMix thus offers a versatile solution and can be devoted to two tasks: 1) unsupervised domain adaptation (UDA) and 2) domain generalization (DG), boosting the performance of LSS models across diverse weather conditions and scenes.

We conduct a thorough evaluation of UniMix on both UDA and DG tasks. We employ large-scale public LiDAR segmentation benchmarks, utilizing SemanticKITTI [2] and SynLiDAR [48] as the source datasets and SemanticSTF [49] as the target dataset. As shown in Fig. 1, experimental results for both tasks demonstrate UniMix's efficacy in mitigating the complex domain gap arising from diverse scenes and adverse weather conditions, surpassing the performance of state-of-the-art (SOTA) methods. Furthermore, we present comprehensive analysis and ablation studies on its key components to validate their effectiveness.

In summary, our contributions are as follows:
- We propose UniMix, a universal method that enhances the adaptability and generalizability of LSS models to unseen adverse-weather scenes.
- We construct a Bridge Domain through physically realis-

tic weather simulation, bridging the gap between the clear weather scenes and the adverse weather scenes.
- We present Universal Mixing, which generates diverse intermediate point clouds from given domains to mitigate domain gaps caused by different scenes and weather.
- UniMix shows its effectiveness and achieves state-of-the-art performance in both UDA and DG tasks.

2. Related Work

LiDAR Semantic Segmentation refers to predicting semantic labels for each 3D point in LiDAR point clouds. It has been extensively investigated in existing approaches with various network architectures and from different perspectives. Approaches can be categorized into point-based (directly handling irregular point cloud data using MLPs [31, 32], convolution networks [25, 26, 39, 41], and transformers [10, 15]), 2D projection-based (saving computation costs by projecting points to range-view images [7, 28, 44, 45] or bird's-eye-view [54]), and voxel-based (representing point clouds as 3D voxels for employing dense 3D convolutions [30, 46] or sparse convolutions [6, 13]).

Unsupervised Domain Adaptation aims to transfer knowledge to address domain shift between a labeled source domain and an unlabeled target domain. In the field of LSS, some approaches [21, 34, 45, 56] project point clouds to 2D images, leveraging established UDA techniques for 2D semantic segmentation. Alternatively, some methods perform direct domain transfer in 3D space through canonical domain construction [51], domain mapping [23, 48], or domain mixing [36]. Notably, CoSMix [36] utilizes point cloud mixing by assembling semantic patches from different domains, demonstrating the efficacy of domain mixing. While these UDA methods make strides in real-to-real and synthetic-to-real adaptation tasks, none specifically addresses clear-to-adverse weather adaptation, which is crucial for mitigating adverse weather effects in real-world applications. Our method is tailored for domain adaptation and generalization to unseen scenes in adverse weather.

Domain Generalization seeks to train a model using only source domain data that can generalize well to unseen target domains. In the 2D vision domain, it has been extensively investigated in previous research [1, 5, 24, 55, 57]. To address domain shift in 3D scene understanding within LSS, some DG methods have been proposed recently. Kim et al. [20] present a single-domain generalization method that enforces consistency between the source and augmented domain, generalizing from a source domain to a target domain with distinct LiDAR configuration and scenes. LiDOG [37] augments the learning network to be robust to sensor placement shifts and resolutions by incorporating an auxiliary BEV segmentation branch. While these methods prove effective in synthetic-to-real and real-to-real generalization, they focus on clear weather conditions. Xiao et al. [49] ex-

plore adverse weather generalization by creating additional augmented point cloud views to train the model for learning perturbation-invariant representations. Our method also targets learning weather-robust and domain-invariant representations and universally addresses domain adaptation and generalization to adverse weather scenes.

3D Perception in Adverse Conditions has gained significant attention among both academic and industrial researchers [8, 9, 58] due to its practical significance. Synthetic data exploration in adverse weather conditions, such as fog and snowfall, has been conducted by Kilic et al. [19] and Hahner et al. [16, 17], who propose robust 3D object detection methods. For real data investigations, Bijelic et al. [3] introduce the STF dataset, a large multimodal dataset for 3D object detection in adverse weather, along with a fusion-based model. SemanticSTF [49] extends STF [3] with additional semantic labels, facilitating the study of LiDAR semantic segmentation in adverse conditions. In this study, we adopt SemanticSTF as the target domain for the domain adaptation and generalization experiments.

3. Method

UniMix offers a universal approach to robust 3D representation learning for LiDAR semantic segmentation. It addresses domain discrepancies between clear and adverse weather conditions, serving both domain adaptation and generalization tasks. Specifically, two types of domain gaps between source and target domains are considered: scene-level and weather-level. To this end, we first introduce a Bridge Domain through physically valid weather condition simulation on source data (Sec. 3.1). Then, we present a Universal Mixing operator to comprehensively blend points across domains, enriching point cloud data in different spatial, intensity, and semantic distributions (Sec. 3.2). Utilizing a teacher-student framework [35], UniMix proves effective for both UDA and DG tasks. The overall pipeline is outlined in Fig. 2 and Algo. 1 (Sec. 3.3).

3.1. Bridge Domain Generation

Effect of Adverse Weather. Previous studies have shown that adverse weather significantly impacts active pulsed systems like LiDAR and hinders the performance of traditional perception models [8, 58]. In adverse weather such as fog, rain, and snow, airborne particles interact with the laser beam, absorbing, reflecting, or refracting its photons [9, 16, 17]. This interaction notably affects the intensity values, point count, spatial allocations, and semantic distributions of point clouds. To address this, we propose creating a Bridge Domain by weather simulation to help train adaptive and generalizable LSS models for adverse weather.
Bridge Domain. According to prior studies [16, 17, 33], the LiDAR system can be represented using a linear model for the received signal power. For non-elastic scattering, the
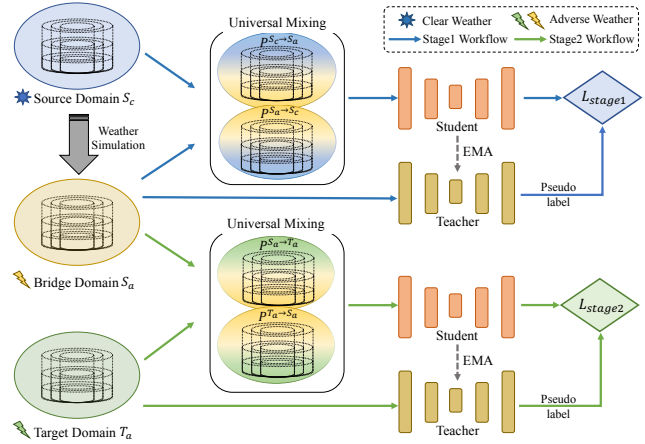

Figure 2. The overall pipeline of UniMix. In the first stage (top part), the clear-weather source domain S_c and the simulated adverse-weather Bridge Domain B_a are taken as input to generate the intermediate domain $\{P^{S_c \to B_a}, P^{B_a \to S_c}\}$ via Universal Mixing, and the student network is trained under the supervision of L_{stage1}. In the second stage (bottom part), the Bridge Domain and adverse-weather target domain T_a are utilized to generate the intermediate domain $\{P^{B_a \to T_a}, P^{T_a \to B_a}\}$ via Universal Mixing, and the student network is trained under the supervision of L_{stage2}. The teacher is leveraged to produce pseudo labels and is updated via EMA [40] of the student's weights.

range-dependent received power P_R is modeled as the timewise convolution of the time-dependent transmitted signal power P_T and the optical system's impulse response H:

$$P_R(R) = C_A \int_0^{2R/c} P_T(t) \, H\left(R - \frac{ct}{2}\right) dt, \quad (1)$$

where R is the object distance, c is the speed of light and C_A is a system constant independent of time t and range.

In challenging weather, adverse conditions affect laser pulse H due to scattering caused by rain, fog, and snowflakes, each with distinct properties [14, 33], *e.g.*, rain and fog consist of small water droplets and snowflakes contain ice crystals with varying densities. These scatterers, modeled as a statistical distribution of reflective spherical objects with varying diameters, attenuate the original power reflectance from solid objects. In this study, we adopt physically valid weather simulation methods [16, 17] to simulate scatterers in the air and the wet ground effect. Specifically, we consider four frequent weather conditions in real life — dense fog, light fog, snow, and rain. Simulating these conditions for each scene in the source domain constructs the Bridge Domain, incorporating both the scene characteristics of the source domain and the weather features of the target domain. This process can be formulated as:

$$B_a = (P^{B_a}, Y^{B_a}) = \Omega(P^{S_c}, Y^{S_c}). \quad (2)$$

Here, we use Ω to denote the overall weather simulation. The subscripts c and a denote clear weather and adverse
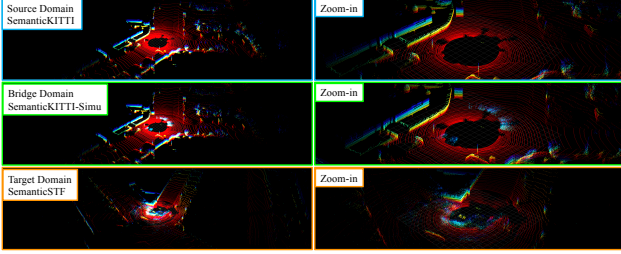
Figure 3. Visualization of point clouds in the source domain, Bridge Domain, and target domain. The Bridge and target domain are in light fog weather. Better viewed with zoom-in and in color.

weather, respectively. S_c denotes the source domain while B_a denotes the generated Bridge Domain. P and Y denote the point clouds and corresponding labels, respectively. We visualize the comparison of point clouds in the source domain, Bridge Domain, and target domain in Fig. 3. The Bridge Domain keeps the scene characteristic of the source domain while integrating the weather disturbance (fog particles in blue) encountered in the target domain. More weather simulation effects can be found in the Appendix.

3.2. Universal Mixing

Since point cloud mixing techniques have been proven effective in various LSS settings before, including CoSMix (synthetic-to-real adaptation), PolarMix (data augmentation), and LaserMix (Semi-supervised LSS), we further dig into point cloud sample mixing according to the impact from diverse weather conditions. In our specific clear-to-adverse scenario, we conduct point cloud mixing regarding spatial, intensity, and semantic distributions to enrich the diversity of mixed intermediate point clouds and help models learn weather-robust and domain-invariant representations.

Let $P \in \mathbb{R}^{N \times 4}$ denote a LiDAR point cloud with N points whose label is represented by Y. Each point p_i can be expressed as (x_i, y_i, z_i, I_i), where (x_i, y_i, z_i) is the 3D Cartesian coordinate of the point and I_i is its intensity value received from the laser scanner. Given a source domain $S = (P^S, Y^S)$ and a target domain $T = (P^T, Y^T)$, we define a Universal Mixing operator Ψ to create an intermediate domain by blending the samples from the two domains:

$$\Psi(S, T) = (S \odot (\mathbf{1} - \mathcal{M}_S)) \copyright (T \odot \mathcal{M}_T), \quad (3)$$

where $\mathcal{M}_S \in \{\mathcal{M}_S^{\circled{S}} \mid \circled{S} = \{\circled{Sp}, \circled{In}, \circled{Se}\}\}$ and $\mathcal{M}_T \in \{\mathcal{M}_T^{\circled{S}} \mid \circled{S} = \{\circled{Sp}, \circled{In}, \circled{Se}\}\}$ represent binary masks indicating where to select from source and target point cloud. $\circled{Sp}, \circled{In}, \circled{Se}$ denote different sampling operators, *i.e.*, spatial mixing, intensity mixing, and semantic mixing. $\mathbf{1}$ is a binary mask filled with ones, $\odot$ denotes Hadamard product, and $\copyright$ represents point cloud concatenation. After bidirectionally applying the Universal Mixing operator (denoted as $\overrightarrow{\overleftarrow{\Psi}}$) to the source data and target data, we can obtain the
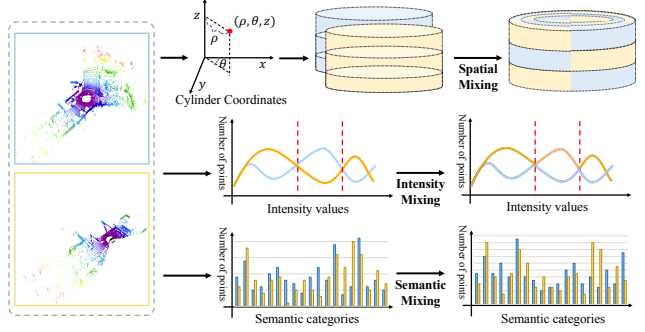

Figure 4. Illustration of different sample mixing methods, including spatial mixing, intensity mixing, and semantic mixing.

mixed intermediate domain data and labels as follows:

$$
\begin{aligned}
P^{S \to T}, P^{T \to S} &= \overrightarrow{\overleftarrow{\Psi}}(P^S, P^T), \\
Y^{S \to T}, Y^{T \to S} &= \overrightarrow{\overleftarrow{\Psi}}(Y^S, Y^T).
\end{aligned}
\quad (4)
$$

Spatial Mixing. For spatial mixing, we adopt a Cylinder coordinate-based partition to both source and target LiDAR point clouds. We first transform the Cartesian coordinates (x, y, z) of the 3D points to the Cylinder coordinate (ρ, θ, z), where the radius $\rho = \sqrt{(x^2 + y^2)}$ is the radial distance from the point to the z-axis and the azimuth $\theta = tan^{-1}(y/x)$ is the angle from x-axis to y-axis indicating the rotation angle in the sensor horizontal plane. Then, the point clouds can be split into a series of non-overlapping spaces along the three dimensions, as shown in Fig. 4 (top). For the convenience of exchanging points, we normalize the Cylinder coordinates of two LiDAR scans to obtain the same partition spaces. The spatial mixing mask can be obtained by dividing the intervals on three coordinate axes: $\mathcal{M}^{\circled{Sp}} = \{(\rho_h < \rho < (\rho_h + \Delta\rho)) \cap (\theta_i < \theta < (\theta_i + \Delta\theta)) \cap (z_j < z < (z_j + \Delta z))\}$. Here, $\Delta\rho$, $\Delta\theta$, and Δz are preset intervals along three dimensions. ρ_h, θ_i, and z_j are randomly picked coordinate values within the corresponding valid ranges. Points covered by the masks will be cut out from the source point cloud and supplemented with the cut from the target points, and vice versa. The labels of the point clouds are sampled and mixed in the same way.

Intensity Mixing. For intensity mixing, we conduct the point cloud partition according to the intensity values of 3D points. We first normalize the intensities to the range of $[0, 1]$. Then, the intensity mixing mask is obtained by dividing the intensity values into different partitions: $\mathcal{M}^{\circled{In}} = \{I_m < I_i < I_m + \Delta I\}$. Here, ΔI is a preset intensity interval and I_m is a randomly picked intensity value within the valid range. As shown in Fig. 4 (middle), the points covered by the masks in both the source and target point clouds (along with their labels) will be cut out and exchanged, alleviating the domain shift caused by intensity variations.

Semantic Mixing. Following CosMix [36], we also conduct semantic mixing which randomly selects a subset of

1: **Input:** Shuffled labeled source data (P^{S_c}, Y^{S_c}), Shuffled unlabeled target data P^{T_a}, weather simulation generator Ω, Universal Mixing operator $\overset{\rightleftharpoons}{\Psi}$. Student net Φ_1^s, Φ_2^s and Teacher net Φ_1^t, Φ_2^t, and training epochs E_1, E_2, for training stage1 and stage2, respectively.
2: **for** $e = 1$ **to** E_1 **do**
3:    $B_a = (P^{B_a}, Y^{B_a}) = \Omega(P^{S_c}, Y^{S_c})$ ▷ *Bridge domain*
   $P^{S_c \to B_a}, P^{B_a \to S_c} = \overset{\rightleftharpoons}{\Psi}(P^{S_c}, P^{B_a})$ ▷ *Data mixing*
4:    $\hat{Y}^{B_a} = \Phi_1^t(P^{B_a})$ ▷ *Pseudo-label gen.*
   $Y^{S_c \to B_a}, Y^{B_a \to S_c} = \overset{\rightleftharpoons}{\Psi}(Y^{S_c}, \hat{Y}^{B_a})$ ▷ *Label mixing*
5:    $\bar{Y}^{S_c \to B_a} = \Phi_1^s(P^{S_c \to B_a})$ ▷ *Student training stage1*
6:    $\bar{Y}^{B_a \to S_c} = \Phi_1^s(P^{B_a \to S_c})$ ▷ *Student training stage1*
7:    $L^{S_c \to B_a} = DiceLoss(\bar{Y}^{S_c \to B_a}, Y^{S_c \to B_a})$ ▷ *Stage1 loss*
8:    $L^{B_a \to S_c} = DiceLoss(\bar{Y}^{B_a \to S_c}, Y^{B_a \to S_c})$ ▷ *Stage1 loss*
9:    $L_{stage1} = L^{S_c \to B_a} + L^{B_a \to S_c}$ ▷ *Stage1 overall loss*
10:    Backward(L_{stage1}), Update(Φ_1^s), UpdateEMA(Φ_1^t)
11: **end for**
12: **if** DG **then**
13:    Pred = $\Phi_1^s(T_a)$ ▷ *Evaluation of DG*
14: **else if** UDA **then**
15:    **for** $e = 1$ **to** E_2 **do**
16:      $P^{B_a \to T_a}, P^{T_a \to B_a} = \overset{\rightleftharpoons}{\Psi}(P^{S_a}, P^{T_a})$ ▷ *Data mixing*
17:      $\hat{Y}^{T_a} = \Phi_2^t(P^{T_a})$ ▷ *Pseudo-label gen.*
     $Y^{B_a \to T_a}, Y^{T_a \to B_a} = \overset{\rightleftharpoons}{\Psi}(Y^{S_a}, \hat{Y}^{T_a})$ ▷ *Label mixing*
18:      $\bar{Y}^{B_a \to T_a} = \Phi_2^s(P^{B_a \to T_a})$ ▷ *Student training stage2*
19:      $\bar{Y}^{T_a \to B_a} = \Phi_2^s(P^{T_a \to B_a})$ ▷ *Student training stage2*
20:      $L^{B_a \to T_a} = DiceLoss(\bar{Y}^{B_a \to T_a}, Y^{B_a \to T_a})$ ▷ *Stage2 loss*
21:      $L^{T_a \to B_a} = DiceLoss(\bar{Y}^{T_a \to B_a}, Y^{T_a \to B_a})$ ▷ *Stage2 loss*
22:      $L_{stage2} = L^{B_a \to T_a} + L^{T_a \to B_a}$ ▷ *Stage2 overall loss*
23:      Backward(L_{stage2}), Update(Φ_2^s), UpdateEMA(Φ_2^t)
24:    **end for**
25:    Pred = $\Phi_2^s(T_a)$ ▷ *Evaluation of UDA*
26: **end if**

semantic patches and their labels from the source and target point clouds and exchanges them with each other. Formally, $\mathcal{M}^{\circledS} = \{y_i \in \{C\}\}$. Here, y_i is the ground truth (or pseudo) semantic label of point p_i, and C is the randomly selected semantic category set. By selecting points from various categories, the mixed point clouds have different semantic distributions with diverse contexts of both domain samples, as shown in Fig. 4 (bottom).

We employ Universal Mixing in both stages to generate the intermediate domains: $\{P^{S_c \to B_a}, P^{B_a \to S_c}\}$ and $\{P^{B_a \to T_a}, P^{T_a \to B_a}\}$, as shown in Fig. 2. The intermediate domains comprehensively blend point clouds across domains, which can be utilized to mitigate the domain shift for training domain adaptive and generalizable models.

3.3. Overall Pipeline

The overall architecture is illustrated in Fig. 2 and the details are summarized in Algo. 1. We adopt the teacher-student framework and form a two-stage training pipeline for handling the UDA and DG tasks universally. In each stage, the teacher network is used for producing pseudo-labels for the target data before label mixing and is updated

via the exponential moving average (EMA) of the student's weights. As shown in Fig. 2, the intermediate domain data $\{P^{S_c \to B_a}, P^{B_a \to S_c}\}$ and their labels $\{Y^{S_c \to B_a}, Y^{B_a \to S_c}\}$ mixed from the clear-weather source domain (P^{S_c}, Y^{S_c}) and the generated Bridge Domain with predicted pseudo-label $(P^{S_a}, \hat{Y}^{S_a})$ are fed into the student network Φ_1^s to obtain the predictions $\{\bar{Y}^{S_c \to B_a}, \bar{Y}^{B_a \to S_c}\}$. The Dice segmentation loss [18] is adopted to update the learnable weights of the student network Φ_1^s:

$$L^{S_c \to B_a} = DiceLoss(\bar{Y}^{S_c \to B_a}, Y^{S_c \to B_a}),$$
$$L^{B_a \to S_c} = DiceLoss(\bar{Y}^{B_a \to S_c}, Y^{B_a \to S_c}), \quad (5)$$
$$L_{stage1} = L^{S_c \to B_a} + L^{B_a \to S_c}.$$

Here, L_{stage1} is the overall loss for the first stage of training. Since the data in the adverse weather target domain is inaccessible in the first stage, the student network after the first stage of training can be evaluated on the DG task.

For handling the UDA task, the framework subsequently forwards to the second stage *i.e.*, Bridge-to-Target adaptation, where the Bridge Domain B_a serves the source domain to adapt to the adverse-weather target domain T_a. Similar to the first stage, the teacher network is utilized to produce the pseudo-label for the target domain. The mixed intermediate domain data $\{P^{B_a \to T_a}, P^{T_a \to B_a}\}$ and labels $\{Y^{B_a \to T_a}, Y^{T_a \to B_a}\}$ are obtained through Universal Mixing. The student network Φ_2^s generates the predictions $\{\bar{Y}^{B_a \to T_a}, \bar{Y}^{T_a \to B_a}\}$ and is trained by minimizing the loss $L_{stage2} = L^{B_a \to T_a} + L^{T_a \to B_a}$. L_{stage2} is similar as L_{stage1}, changing the Source-to-Bridge to Bridge-to-Target domain adaptation, as shown in Algo. 1. Thus, the student network can be used for evaluation on the UDA task after training. Implementation details can be found in Appendix.

4. Experiments

We evaluate the adaptive and generalizable ability of UniMix for adverse weather conditions in two tasks, *i.e.*, UDA and DG. For each task, we adopt two benchmark settings, *i.e.*, SemantiKITTI→SemanticSTF and SynLiDAR→SemanticSTF, where the clear weather real dataset SemanticKITTI [2] and synthetic dataset SynLiDAR [48] are used as the source domain to adapt or generalize to the adverse-weather target domain SemanticSTF [49], respectively. To further demonstrate the adaptation ability for the synthetic-to-real UDA scenario in clear weather, we also evaluate UniMix in the SynLiDAR→SemantiKITTI setting. In addition, we employ Universal Mixing for data augmentation to evaluate its impact on increasing the performance of supervised LSS models on the source domain.

4.1. Datasets and Metrics

Datasets. SemanticKITTI [2] is a large-scale LiDAR semantic segmentation dataset, extended from the KITTI Vi-

Method	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	oth-g.	build.	fence	veget.	trunk	terra.	pole	traf.	mIoU
Oracle	89.4	42.1	0.0	59.9	61.2	69.6	39.0	0.0	82.2	21.5	58.2	45.6	86.1	63.6	80.2	52.0	77.6	50.1	61.7	54.7
								SemanticKITTI→SemanticSTF												
Source-only	55.9	0.0	0.2	1.9	10.9	10.3	6.0	0.0	61.2	10.9	32.0	0.0	67.9	41.6	49.8	27.9	40.8	29.6	17.5	24.4
ADDA [42]	65.6	0.0	0.0	21.0	1.3	2.8	1.3	16.7	64.7	1.2	35.4	0.0	66.5	41.8	**57.2**	32.6	42.2	23.3	26.4	26.3
Ent-Min [43]	69.2	0.0	10.1	31.0	5.3	2.8	2.6	0.0	65.9	2.6	35.7	0.0	72.5	42.8	52.4	32.5	**44.7**	24.7	21.1	27.2
Self-training [59]	71.5	0.0	10.3	33.1	7.4	5.9	1.3	0.0	65.1	6.5	36.6	0.0	67.8	41.3	51.7	32.9	42.9	25.1	25.0	27.6
CoSMix [36]	65.0	**1.7**	22.1	25.2	7.7	33.2	0.0	0.0	64.7	11.5	31.1	0.9	62.5	37.8	44.6	30.5	41.1	**30.9**	28.6	28.4
UniMix	75.3	0.9	44.9	11.7	**13.6**	38.2	50.3	31.9	71.1	15.0	46.4	6.5	74.3	51.0	49.8	36.8	34.4	25.5	28.9	**37.2**
								SynLiDAR→SemanticSTF												
Source-only	27.1	**3.0**	0.6	15.8	0.1	25.2	1.8	5.6	23.9	0.3	14.6	0.6	36.3	19.9	37.9	17.9	**41.8**	9.5	2.3	15.0
ADDA [42]	55.8	0.0	3.6	26.1	1.3	25.2	7.5	9.9	17.2	**23.4**	4.4	0.9	43.9	18.4	45.2	21.8	33.6	28.0	19.7	20.3
Ent-Min [43]	48.3	0.1	5.6	**28.7**	0.1	23.3	2.5	19.8	19.3	6.7	22.6	1.4	46.9	20.7	43.2	25.2	34.1	26.0	22.2	20.9
Self-training [59]	50.6	0.0	6.1	31.0	0.5	26.0	4.8	12.0	20.7	4.6	23.5	1.5	45.3	19.5	44.6	25.0	35.1	**29.2**	20.8	21.1
CoSMix [36]	51.5	0.2	5.0	28.1	0.0	26.5	**17.0**	9.9	20.2	3.6	24.6	2.2	52.6	20.6	47.5	24.3	34.6	28.2	24.1	22.1
UniMix	73.6	0.0	7.9	26.9	2.9	29.1	13.7	21.8	38.0	8.0	26.3	3.4	56.0	21.2	56.1	29.6	38.0	28.2	26.5	26.7

Table 1. Comparison of SOTA domain adaptation methods on SemanticKITTI→SemanticSTF and SynLiDAR→SemanticSTF. SemanticKITTI and SynLiDAR serve as the source domain while SemanticSTF with all four weather conditions serves as the target domain. Our UniMix achieves the best gain of 12.8 and 11.7 mIoU over the source-only model in two benchmarks.

sual Odometry benchmark [11]. It covers a wide range of real-world urban scenes captured in German and provides point-wise annotations over 19 semantic categories. We use sequences 00-07 and 09-10 for training and sequence 08 as validation following the official protocol. **SynLi-DAR** [48] is a large-scale synthetic LiDAR semantic segmentation dataset, containing virtual urban cities, suburban towns, neighborhoods, and harbor scenes. It provides point-wise annotations over 32 semantic categories. Following the command practice [35, 36, 48, 49], we use 19,840 point clouds for training and 1,976 for validation. **SemanticSTF** [49] is an adverse-weather point cloud semantic segmentation dataset that provides point-wise annotations over 21 semantic categories. It contains 2,076 scans from the STF dataset [3] that cover various adverse weather conditions including 637 dense fog, 631 light fog, and 114 rain and 694 snow scenes. We follow the official protocol using 1,326 scans for training, and 250 for validation.

Metrics. We adopt the evaluation metrics of Intersection over the Union (IoU) for each segmentation class and the mean IoU (mIoU) over all classes. Note that since the label set of SemanticSTF is the same as SemanticKITTI, we map the labels in all the datasets to the 19 common classes of SemanticKITTI following the common practice.

4.2. UniMix for Unsupervised Domain Adaptation

Setup. In the UDA task, we evaluate UniMix on two benchmarks: SemantiKITTI→SemanticSTF and SynLiDAR→SemanticSTF. We compare UniMix with four representative UDA approaches including ADDA [42], entropy minimization (Ent-Min) [43], Self-training [59], and CoSMix [36]. For a fair comparison, all these methods are trained using the MinkowskiNet [6] as the 3D segmentation backbone.

Performance. As shown in Table 1, we take the whole SemanticSTF dataset (including all adverse weather scenes) as the target domain to show the performance of UDA methods on the validation set, adapting the model from the SemanticKITTI source (top part) and SynLiDAR source (bottom part), respectively. The results of the *Oracle* model and the *Source-only* model are the upper bound and lower bound for each adaptation scenario. All UDA methods achieve superior performance than the source-only model in both benchmarks. Our UniMix achieves the best gain of 12.8 mIoU for SemantiKITTI→SemanticSTF adaptation and 11.7 mIoU for SynLiDAR→SemanticSTF adaption, with an improvement of 8.8 mIoU and 4.6 mIoU over the SOTA synthetic-to-real adaptation method CoSMix [36]. Additionally, the results indicate that the adaptation from a synthetic dataset is more difficult for all UDA methods than that from a real dataset, which may be attributed to the additional domain shift between synthetic and real-world data.

4.3. UniMix for Domain Generalization

Setup. Different from the UDA setting, the target domain is supposed to be inaccessible during training for DG. Following PointDR [49], we compare UniMix with three augmentation-based methods (*i.e.*, Dropout [38], Noise perturbation, PolarMix [47]) and two representative 2D DG methods (*i.e.*, MMD [24] and PCL [50]). We evaluate the methods on two benchmarks: SemantiKITTI→SemanticSTF and SynLiDAR→SemanticSTF.

Performance. In Table 2, we present the generalization results of different methods on the two benchmarks. Source-only model is trained on the source SemanticKITTI or SynLiDAR dataset without applying any generalization techniques. It performs very poorly in the adverse-weather target domain, due to the large domain gap. The augmentation-based methods and 2D DG methods improve the performance over the Source-only baseline but the gains are very limited in the range of 1.3-2.5 mIoU for

Method	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	oth-g.	build.	fence	veget.	trunk	terra.	pole	traf.	Dense-fog	Light-fog	Rain	Snow	mIoU
Oracle	89.4	42.1	0.0	59.9	61.2	69.6	39.0	0.0	82.2	21.5	58.2	45.6	86.1	63.6	80.2	52.0	77.6	50.1	61.7	51.9	54.6	57.9	53.7	54.7
SemanticKITTI→SemanticSTF																								
Source-only	55.9	0.0	0.2	1.9	10.9	10.3	6.0	0.0	61.2	10.9	32.0	0.0	67.9	41.6	49.8	27.9	40.8	29.6	17.5	29.5	26.0	28.4	21.4	24.4
Dropout [38]	62.1	0.0	**15.5**	3.0	11.5	5.4	2.0	0.0	58.4	12.8	26.7	1.1	72.1	43.6	52.9	**34.2**	43.5	28.4	15.5	29.3	25.6	29.4	24.8	25.7
Perturbation [49]	74.4	0.0	0.0	**23.3**	0.6	19.7	0.0	0.0	60.3	10.8	33.9	0.7	72.0	45.2	58.7	17.5	42.4	22.1	9.7	26.3	27.8	30.0	24.5	25.9
PolarMix [47]	57.8	1.8	3.8	16.7	3.7	26.5	0.0	2.0	65.7	2.9	32.5	0.3	71.0	**48.7**	53.8	20.5	45.4	25.9	15.8	29.7	25.0	28.6	25.6	26.0
MMD [24]	63.6	0.0	2.6	0.1	11.4	28.1	0.0	0.0	**67.0**	**14.1**	37.9	0.3	67.3	41.2	57.1	27.4	47.9	28.2	16.2	30.4	28.1	32.8	25.2	26.9
PCL [50]	65.9	0.0	0.0	17.7	0.4	8.4	0.0	0.0	59.6	12.0	35.0	1.6	**74.0**	47.5	**60.7**	15.8	**48.9**	26.1	27.5	28.9	27.6	30.1	24.6	26.4
PointDR [49]	67.3	0.0	4.5	19.6	9.0	18.8	2.7	0.0	62.6	12.9	**38.1**	0.6	73.3	43.8	56.4	32.2	45.7	28.7	27.4	31.3	29.7	31.9	26.2	28.6
UniMix	**82.7**	**6.6**	8.6	4.5	**15.1**	**35.5**	**15.5**	**37.7**	55.8	10.2	36.2	1.3	72.8	40.1	49.1	33.4	34.9	23.5	**33.5**	**34.8**	**30.2**	**34.9**	**30.9**	**31.4**
SynLiDAR→SemanticSTF																								
Source-only	27.1	**3.0**	0.6	15.8	0.1	25.2	1.8	5.6	23.9	0.3	14.6	0.6	36.3	19.9	37.9	17.9	41.8	9.5	2.3	16.9	17.2	17.2	11.9	15.0
Dropout [38]	28.0	**3.0**	1.4	9.6	0.0	17.1	0.8	0.7	34.2	6.8	19.1	0.1	35.5	19.1	42.3	17.6	36.0	14.0	2.8	15.3	16.6	20.4	14.0	15.2
Perturbation [49]	27.1	2.3	2.3	16.0	0.1	23.7	1.2	4.0	27.0	3.6	16.2	0.8	29.2	16.7	35.3	**22.7**	38.3	17.9	5.1	16.3	16.7	19.3	13.4	15.2
PolarMix [47]	39.2	1.1	1.2	8.3	1.5	17.8	0.8	0.7	23.3	1.3	17.5	0.4	45.2	24.8	46.2	20.1	38.7	7.6	1.9	16.1	15.5	19.2	15.6	15.7
MMD [24]	25.5	2.3	2.1	13.2	0.7	22.1	1.4	7.5	30.8	0.4	17.6	0.2	30.9	19.7	37.6	19.3	**43.5**	9.9	2.6	17.3	16.3	20.0	12.7	15.1
PCL [50]	30.9	0.8	1.4	10.0	0.4	23.3	**4.0**	7.9	28.5	1.3	17.7	1.2	39.4	18.5	40.0	16.0	38.6	12.1	2.3	17.8	16.7	19.3	14.1	15.5
PointDR [49]	37.8	2.5	2.4	**23.6**	0.1	26.3	2.2	3.3	27.9	**7.7**	17.5	0.5	47.6	25.3	45.7	21.0	37.5	17.9	5.5	19.5	19.9	21.1	16.9	18.5
UniMix	**65.4**	0.1	**3.9**	16.9	**5.3**	**32.3**	2.0	**19.3**	**52.1**	5.0	**27.3**	**3.0**	**49.4**	20.3	**58.5**	22.7	23.2	**26.9**	10.4	**24.3**	**22.9**	**26.1**	**20.9**	**23.4**

Table 2. Comparison of SOTA domain generalization methods on SemanticKITTI→SemanticSTF and SynLiDAR→SemanticSTF.

Bridge Domain	UDA mIoU					DG mIoU				
	D-fog	L-fog	Rain	Snow	All	D-fog	L-fog	Rain	Snow	All
None	39.8	31.6	35.3	29.7	31.6	29.5	26.0	28.4	21.4	24.4
L-fog	38.2	32.0	37.9	31.0	34.3	32.6	28.1	29.6	22.1	27.1
D-fog	40.0	33.0	38.4	31.3	34.7	33.7	26.4	29.4	22.5	27.5
Rain	40.2	31.9	39.3	31.5	34.9	34.5	25.9	32.9	23.2	28.1
Snow	39.5	32.5	37.3	31.6	35.5	31.8	28.7	32.9	24.5	28.4
Mixed-fog	40.3	33.3	38.7	32.2	36.3	35.7	26.8	32.5	23.5	29.7
All	40.2	34.0	37.5	33.2	37.2	34.8	30.2	34.9	30.9	31.4

Table 3. Ablation study of Bridge Domain compositions used in UniMix for UDA and DG on SemanticKITTI→SemanticSTF. "D-fog" and "L-fog" denote dense fog and light fog, respectively.

Mixing Method	UDA mIoU					DG mIoU				
	D-fog	L-fog	Rain	Snow	All	D-fog	L-fog	Rain	Snow	All
Source-only	26.9	25.2	27.7	23.5	24.3	29.5	26.0	28.4	21.4	24.4
Spatial	41.6	34.6	35.7	27.8	35.2	33.0	28.3	27.9	23.3	26.4
Intensity	41.2	34.5	37.4	28.8	34.9	33.8	25.6	30.0	18.9	25.8
Semantic	41.7	34.3	29.6	36.3	35.4	37.4	26.7	27.8	22.3	26.7
All	40.2	34.0	37.5	33.2	37.2	34.8	30.2	34.9	30.9	31.4

Table 4. Ablation study of mixing operators for UDA and DG on SemanticKITTI→SemanticSTF.

Method	UDA mIoU	DG mIoU
Source-only	24.4	24.4
CosMix [36]	35.4	26.7
PolarMix [47]	33.0	26.0
LaserMix [22]	33.3	26.9
Universal Mixing	**37.2**	**31.4**

Table 5. Comparison of mixing methods in our UniMix pipeline for UDA and DG on SemanticKITTI→SemanticSTF.

SemanticKITTI→SemanticSTF generalization and 0.1-0.7 mIoU for SynLiDAR→SemanticSTF generalization. The state-of-the-art method PointDR achieves a gain of 4.2 mIoU and 3.5 mIoU but still lags behind our UniMix, which achieves a gain of 7.0 and 8.4 mIoU over the Source-only baseline. We also show the generalization results for each weather in the right part of Table 2, where UniMix also achieves the best performance among all methods.

4.4. Ablation Study

We conduct ablation studies of UniMix on SemanticKITTI →SemanticSTF for both UDA and DG tasks, focusing on three key components of UniMix: the Bridge Domain, Universal Mixing, and the two-stage training architecture.

Bridge Domain Compositions. Our method produces four frequent weather conditions in real life, *i.e.*, dense fog, light fog, rain, and snow to form the Bridge Domain. As shown in Table 3, we ablate the impact of the simulated weather conditions in generating the Bridge Domain, including four single-weather settings and a mixed fog setting, where 50% dense fog data and 50% light fog data are generated. We show the results evaluated on the whole validation set with

all four weather conditions as well as on the subsets of each weather condition. As can be seen, UniMix with any of the choices for generating the Bridge Domain achieves a gain of 2.7-5.6 mIoU (2.7-7.0 mIoU) for UDA (DG) over the vanilla UniMix baseline without using the Bridge Domain. Among all the choices, the mixed-fog brings the largest improvement in both tasks, while the light-fog produces the smallest improvement. This can be attributed to the small perturbations brought by the light-fog simulation, while the mixed-fog domain integrates diverse disturbance patterns via both dense-fog and light-fog simulation. We also observe a relatively high performance when the simulated weather in the Bridge Domain is the same as the weather of the validation subset, which is reasonable and implies the necessity of simulating all weather conditions in the Bridge Domain as validated by the last row.

Method	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	park.	sidew.	oth-g.	build.	fence	veget.	trunk	terra.	pole	traf.	mIoU
Source-only	42.0	5.0	4.8	0.4	2.5	12.4	43.3	1.8	48.7	4.5	31.0	0.0	18.6	11.5	60.2	30.0	48.3	19.3	3.0	20.4
ADDA [42]	52.5	4.5	11.9	0.3	3.9	9.4	27.9	0.5	52.8	4.9	27.4	0.0	61.0	17.0	57.4	34.5	42.9	23.2	4.5	23.0
Ent-Min [43]	58.3	5.1	14.3	0.3	1.8	14.3	**44.5**	0.5	50.4	4.3	34.8	0.0	48.3	19.7	67.5	34.8	**52.0**	33.0	6.1	25.8
ST [59]	62.0	5.0	12.4	1.3	9.2	16.7	44.2	0.4	53.0	2.5	28.4	0.0	57.1	18.7	69.8	35.0	48.7	32.5	6.9	26.5
PCT [48]	53.4	5.4	7.4	0.8	10.9	12.0	43.2	0.3	50.8	3.7	29.4	0.0	48.0	10.4	68.2	33.1	40.0	29.5	6.9	23.9
ST-PCT [48]	70.8	7.3	13.1	1.9	8.4	12.6	44.0	0.6	56.4	4.5	31.8	0.0	66.7	**23.7**	**73.3**	34.6	48.4	39.4	11.7	28.9
CosMix [36]	75.1	6.8	29.4	27.1	**11.1**	22.1	25.0	**24.7**	**79.3**	14.9	**46.7**	**0.1**	53.4	13.0	67.7	31.4	32.1	37.9	13.4	32.2
PolarMix [47]	76.3	**8.4**	17.8	3.9	6.0	**26.6**	40.8	15.9	70.3	0.0	44.4	0.0	**68.4**	14.7	69.6	**38.1**	37.1	**40.6**	10.6	31.0
UniMix	**80.3**	6.1	**32.5**	**29.2**	10.6	23.7	30.0	24.5	62.2	**15.9**	46.5	**0.1**	60.9	15.8	70.7	34.9	41.2	38.6	**17.1**	**33.7**

Table 6. Comparison of SOTA domain adaptation methods on SynLiDAR→SemanticKITTI adaptation scenario. The superior performance demonstrates the potential of UniMix to perform as a general UDA method for LSS.

Dataset	Method	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	oth-g.	build.	fence	veget.	trunk	terra.	pole	traf.	mIoU
SemanticKITTI	MinkNet [6]	95.7	3.7	44.9	53.2	42.1	53.7	68.9	0.0	92.8	43.0	80.0	1.8	90.5	60.0	87.4	64.5	73.3	62.1	43.7	55.9
	+UM	90.2	50.5	74.4	82.3	70.2	62.7	67.5	0.1	73.1	56.9	68.5	21.4	86.5	43.2	84.6	70.1	51.5	71.3	43.5	61.5$_{+5.6}$
SynLiDAR	MinkNet [6]	77.5	12.0	83.6	87.5	88.6	70.6	42.7	94.1	80.7	76.7	69.5	93.7	92.6	80.3	85.2	86.1	65.4	69.2	75.5	75.3
	+UM	75.7	70.7	80.8	86.4	86.9	65.7	89.1	94.4	93.0	73.1	80.8	92.6	93.8	77.0	87.8	83.1	74.2	65.4	63.7	80.7$_{+5.4}$

Table 7. Universal Mixing (denoted as "UM") can also be used as a data augmentation technique to improve the performance of supervised semantic segmentation.

Sample Mixing Methods. Universal Mixing contains three kinds of mixing methods, *i.e.*, spatial, intensity, and semantic mixing. As shown in Table 4, we ablate the choice of mixing methods. According to the results, each of the three mixing methods enhances the performance of the Source-only baseline on both UDA and DG tasks. Semantic mixing delivers the largest improvement on the whole validation set while performing worse in the rain subset than other mixing methods. Only using intensity mixing brings the smallest improvement while performing the best in the rain subset. We attribute this to the distinct reflectance attenuation caused by different weather conditions. By incorporating all the mixing methods together, UniMix achieves the best performance on both tasks.

In addition, we compare our Universal Mixing with some representative point cloud mixing methods, including LaserMix [22] proposed for semi-supervised LSS, PolarMix [47] for point cloud data augmentation, and CosMix [36] for synthetic-to-real domain adaptation. We replace our Universal Mixing with these alternatives in our UniMix pipeline and evaluate them on both UDA and DG tasks. As shown in Table 5, the proposed Universal Mixing achieves the best performance, showing its superiority over others.

Training Stage. Multi-stage training is widely used in 2D UDA semantic segmentation methods [52, 53]. Our method adopts a two-stage training framework as well. We show the UDA results without the Bridge Domain and the first-stage training, denoted as "None" in the first row in Table 3. The performance is decreased but still superior to other SOTA methods. In the DG setting, this version is the Source-only model, without using any Bridge or target domain data. It is also noteworthy that our two-stage UniMix for UDA can generate a DG model after the first training stage, providing a versatile solution and holding practical significance.

4.5. UniMix in Clear Weather

Although UniMix is tailored for clear-to-adverse weather domain adaptation and generalization, we also examine UniMix in synthetic-to-real adaptation in clear weather. In Table 6, UniMix achieves 33.7 mIoU with a gain of 13.3 mIoU over the Source-only model. Its superior performance over the SOTA methods validates the effectiveness of UniMix for synthetic-to-real adaptation, implying that it could be used as a general method for LSS UDA tasks.

4.6. Universal Mixing as Data Augmentation

We also use Universal Mixing as a data augmentation technique in supervised LSS tasks, using the widely adopted MinkowskiNet [6] as our baseline. Results in Table 7 show significant performance enhancement with Universal Mixing on both datasets compared to the baseline model.

5. Conclusion

In this paper, we introduce UniMix, a universal method designed to enhance the adaptive and generalizable capabilities of LiDAR semantic segmentation models in adverse weather conditions. First, a Bridge Domain is constructed, capturing scene characteristics of the clear-weather source domain with adverse weather conditions to learn weather-robust representations. Universal Mixing is then proposed and employed in two stages to blend samples from the source and target domains, considering spatial, intensity, and semantic distribution, to learn domain invariant features. UniMix proves highly effective in UDA and DG tasks, significantly outperforming state-of-the-art methods.

References

[1] Yogesh Balaji, Swami Sankaranarayanan, and Rama Chellappa. Metareg: Towards domain generalization using meta-regularization. *Advances in Neural Information Processing Systems*, 31, 2018. 2

[2] Jens Behley, Martin Garbade, Andres Milioto, Jan Quenzel, Sven Behnke, Cyrill Stachniss, and Jurgen Gall. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9297–9307, 2019. 1, 2, 5

[3] 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 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020. 3, 6

[4] David Brüggemann, Christos Sakaridis, Prune Truong, and Luc Van Gool. Refign: Align and refine for adaptation of semantic segmentation to adverse conditions. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3174–3184, 2023. 1

[5] Sungha Choi, Sanghun Jung, Huiwon Yun, Joanne T Kim, Seungryong Kim, and Jaegul Choo. Robustnet: Improving domain generalization in urban-scene segmentation via instance selective whitening. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11580–11590, 2021. 2

[6] Christopher Choy, JunYoung Gwak, and Silvio Savarese. 4d spatio-temporal convnets: Minkowski convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3075–3084, 2019. 2, 6, 8

[7] Tiago Cortinhal, George Tzelepis, and Eren Erdal Aksoy. Salsanext: Fast, uncertainty-aware semantic segmentation of lidar point clouds. In *Advances in Visual Computing*, pages 207–222. Springer, 2020. 2

[8] Yinpeng Dong, Caixin Kang, Jinlai Zhang, Zijian Zhu, Yikai Wang, Xiao Yang, Hang Su, Xingxing Wei, and Jun Zhu. Benchmarking robustness of 3d object detection to common corruptions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1022–1032, 2023. 3

[9] Mariella Dreissig, Dominik Scheuble, Florian Piewak, and Joschka Boedecker. Survey on lidar perception in adverse weather conditions. *arXiv preprint arXiv:2304.06312*, 2023. 1, 3

[10] Lunhao Duan, Shanshan Zhao, Nan Xue, Mingming Gong, Gui-Song Xia, and Dacheng Tao. Condaformer: Disassembled transformer with local structure enhancement for 3d point cloud understanding. *Advances in Neural Information Processing Systems*, 36, 2024. 2

[11] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3354–3361, 2012. 6

[12] Ziyang Gong, Fuhao Li, Yupeng Deng, Wenjun Shen, Xianzheng Ma, Zhenming Ji, and Nan Xia. Train one, generalize to all: Generalizable semantic segmentation from single-scene to all adverse scenes. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 2275–2284, 2023. 1

[13] Benjamin Graham, Martin Engelcke, and Laurens Van Der Maaten. 3d semantic segmentation with submanifold sparse convolutional networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9224–9232, 2018. 2

[14] KLS Gunn and JS Marshall. The distribution with size of aggregate snowflakes. *Journal of Atmospheric Sciences*, 15 (5):452–461, 1958. 3

[15] Meng-Hao Guo, Jun-Xiong Cai, Zheng-Ning Liu, Tai-Jiang Mu, Ralph R Martin, and Shi-Min Hu. Pct: Point cloud transformer. *Computational Visual Media*, 7:187–199, 2021. 2

[16] Martin Hahner and et al. Fog simulation on real lidar point clouds for 3d object detection in adverse weather. In *CVPR*, 2021. 3

[17] Martin Hahner and et al. Lidar snowfall simulation for robust 3d object detection. In *CVPR*, 2022. 3

[18] Shruti Jadon. A survey of loss functions for semantic segmentation. In *IEEE conference on Computational Intelligence in Bioinformatics and Computational Biology*, pages 1–7, 2020. 5

[19] Velat Kilic and et al. Lidar light scattering augmentation (lisa): Physics-based simulation of adverse weather conditions for 3d object detection. *arXiv preprint arXiv:2107.07004*, 2021. 3

[20] Hyeonseong Kim, Yoonsu Kang, Changgyoon Oh, and Kuk-Jin Yoon. Single domain generalization for lidar semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17587–17598, 2023. 2

[21] Lingdong Kong, Niamul Quader, and Venice Erin Liong. Conda: Unsupervised domain adaptation for lidar segmentation via regularized domain concatenation. In *IEEE International Conference on Robotics and Automation*, pages 9338–9345, 2023. 2

[22] Lingdong Kong, Jiawei Ren, Liang Pan, and Ziwei Liu. Lasermix for semi-supervised lidar semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21705–21715, 2023. 2, 7, 8

[23] Ferdinand Langer, Andres Milioto, Alexandre Haag, Jens Behley, and Cyrill Stachniss. Domain transfer for semantic segmentation of lidar data using deep neural networks. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 8263–8270, 2020. 2

[24] Haoliang Li, Sinno Jialin Pan, Shiqi Wang, and Alex C Kot. Domain generalization with adversarial feature learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5400–5409, 2018. 2, 6, 7

[25] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. Pointcnn: Convolution on x-transformed

points. *Advances in Neural Information Processing Systems*, 31, 2018. 2

[26] Zhijian Liu, Haotian Tang, Yujun Lin, and Song Han. Point-voxel cnn for efficient 3d deep learning. *Advances in Neural Information Processing Systems*, 32, 2019. 2

[27] Xianzheng Ma, Zhixiang Wang, Yacheng Zhan, Yinqiang Zheng, Zheng Wang, Dengxin Dai, and Chia-Wen Lin. Both style and fog matter: Cumulative domain adaptation for semantic foggy scene understanding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18922–18931, 2022. 1

[28] Andres Milioto, Ignacio Vizzo, Jens Behley, and Cyrill Stachniss. Rangenet++: Fast and accurate lidar semantic segmentation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4213–4220, 2019. 2

[29] Karl Montalban, Christophe Reymann, Dinesh Atchuthan, Paul-Edouard Dupouy, Nicolas Riviere, and Simon Lacroix. A quantitative analysis of point clouds from automotive lidars exposed to artificial rain and fog. *Atmosphere*, 12(6): 738, 2021. 1

[30] Charles R Qi, Hao Su, Matthias Nießner, Angela Dai, Mengyuan Yan, and Leonidas J Guibas. Volumetric and multi-view cnns for object classification on 3d data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5648–5656, 2016. 2

[31] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 652–660, 2017. 2

[32] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems*, 30, 2017. 2

[33] Ralph H Rasshofer, Martin Spies, and Hans Spies. Influences of weather phenomena on automotive laser radar systems. *Advances in radio science*, 9:49–60, 2011. 3

[34] Mrigank Rochan, Shubhra Aich, Eduardo R Corral-Soto, Amir Nabatchian, and Bingbing Liu. Unsupervised domain adaptation in lidar semantic segmentation with self-supervision and gated adapters. In *IEEE International Conference on Robotics and Automation*, pages 2649–2655, 2022. 2

[35] Cristiano Saltori, Fabio Galasso, Giuseppe Fiameni, Nicu Sebe, Elisa Ricci, and Fabio Poiesi. Cosmix: Compositional semantic mix for domain adaptation in 3d lidar segmentation. In *European Conference on Computer Vision*, pages 586–602. Springer, 2022. 1, 2, 3, 6

[36] Cristiano Saltori, Fabio Galasso, Giuseppe Fiameni, Nicu Sebe, Fabio Poiesi, and Elisa Ricci. Compositional semantic mix for domain adaptation in point cloud segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 1, 2, 4, 6, 7, 8

[37] Cristiano Saltori, Aljosa Osep, Elisa Ricci, and Laura Leal-Taixé. Walking your lidog: A journey through multiple domains for lidar semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 196–206, 2023. 2

[38] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014. 6, 7

[39] Hang Su, Varun Jampani, Deqing Sun, Subhransu Maji, Evangelos Kalogerakis, Ming-Hsuan Yang, and Jan Kautz. Splatnet: Sparse lattice networks for point cloud processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2530–2539, 2018. 2

[40] Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in Neural Information Processing Systems*, 30, 2017. 3

[41] Hugues Thomas, Charles R Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6411–6420, 2019. 2

[42] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7167–7176, 2017. 6, 8

[43] Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2517–2526, 2019. 6, 8

[44] Bichen Wu, Alvin Wan, Xiangyu Yue, and Kurt Keutzer. Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud. In *IEEE International Conference on Robotics and Automation*, pages 1887–1893, 2018. 2

[45] Bichen Wu, Xuanyu Zhou, Sicheng Zhao, Xiangyu Yue, and Kurt Keutzer. Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud. In *IEEE International Conference on Robotics and Automation*, pages 4376–4382, 2019. 2

[46] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1912–1920, 2015. 2

[47] Aoran Xiao, Jiaxing Huang, Dayan Guan, Kaiwen Cui, Shijian Lu, and Ling Shao. Polarmix: A general data augmentation technique for lidar point clouds. *Advances in Neural Information Processing Systems*, 35:11035–11048, 2022. 6, 7, 8

[48] Aoran Xiao, Jiaxing Huang, Dayan Guan, Fangneng Zhan, and Shijian Lu. Transfer learning from synthetic to real lidar point cloud for semantic segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2795–2803, 2022. 2, 5, 6, 8

[49] 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 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9382–9392, 2023. 1, 2, 3, 5, 6, 7

[50] Xufeng Yao, Yang Bai, Xinyun Zhang, Yuechen Zhang, Qi Sun, Ran Chen, Ruiyu Li, and Bei Yu. Pcl: Proxy-based contrastive learning for domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7097–7107, 2022. 6, 7

[51] Li Yi, Boqing Gong, and Thomas Funkhouser. Complete & label: A domain adaptation approach to semantic segmentation of lidar point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15363–15373, 2021. 2

[52] Pan Zhang, Bo Zhang, Ting Zhang, Dong Chen, Yong Wang, and Fang Wen. Prototypical pseudo label denoising and target structure learning for domain adaptive semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12414–12424, 2021. 8

[53] Qiming Zhang, Jing Zhang, Wei Liu, and Dacheng Tao. Category anchor-guided unsupervised domain adaptation for semantic segmentation. *Advances in Neural Information Processing Systems*, 32, 2019. 8

[54] Yang Zhang, Zixiang Zhou, Philip David, Xiangyu Yue, Zerong Xi, Boqing Gong, and Hassan Foroosh. Polarnet: An improved grid representation for online lidar point clouds semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9601–9610, 2020. 2

[55] Shanshan Zhao, Mingming Gong, Tongliang Liu, Huan Fu, and Dacheng Tao. Domain generalization via entropy regularization. *Advances in Neural Information Processing Systems*, 33:16096–16107, 2020. 2

[56] Sicheng Zhao, Yezhen Wang, Bo Li, Bichen Wu, Yang Gao, Pengfei Xu, Trevor Darrell, and Kurt Keutzer. epointda: An end-to-end simulation-to-real domain adaptation framework for lidar point cloud segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3500–3509, 2021. 2

[57] Kaiyang Zhou, Ziwei Liu, Yu Qiao, Tao Xiang, and Chen Change Loy. Domain generalization: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 2

[58] Zijian Zhu, Yichi Zhang, Hai Chen, Yinpeng Dong, Shu Zhao, Wenbo Ding, Jiachen Zhong, and Shibao Zheng. Understanding the robustness of 3d object detection with bird's-eye-view representations in autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21600–21610, 2023. 3

[59] Yang Zou, Zhiding Yu, Xiaofeng Liu, BVK Kumar, and Jinsong Wang. Confidence regularized self-training. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5982–5991, 2019. 6, 8