Adaptive Augmentation-Aware Latent Learning for Robust LiDAR Semantic Segmentation

Wangkai Li[1], Zhaoyang Li[1], Yuwen Pan[1], Rui Sun[1] , Yujia Chen[1], Tianzhu Zhang[1, 2]

[1]University of Science and Technology of China, [2]Deep Space Exploration Laboratory

中国科学技术大学 University of Science and Technology of China

DSEL DEEP SPACE EXPLORATION LAB 深空探测实验室

ICLR

TLDR. Safely utilize aggressive augmentations for robust LiDAR segmentation by semantic shift localization

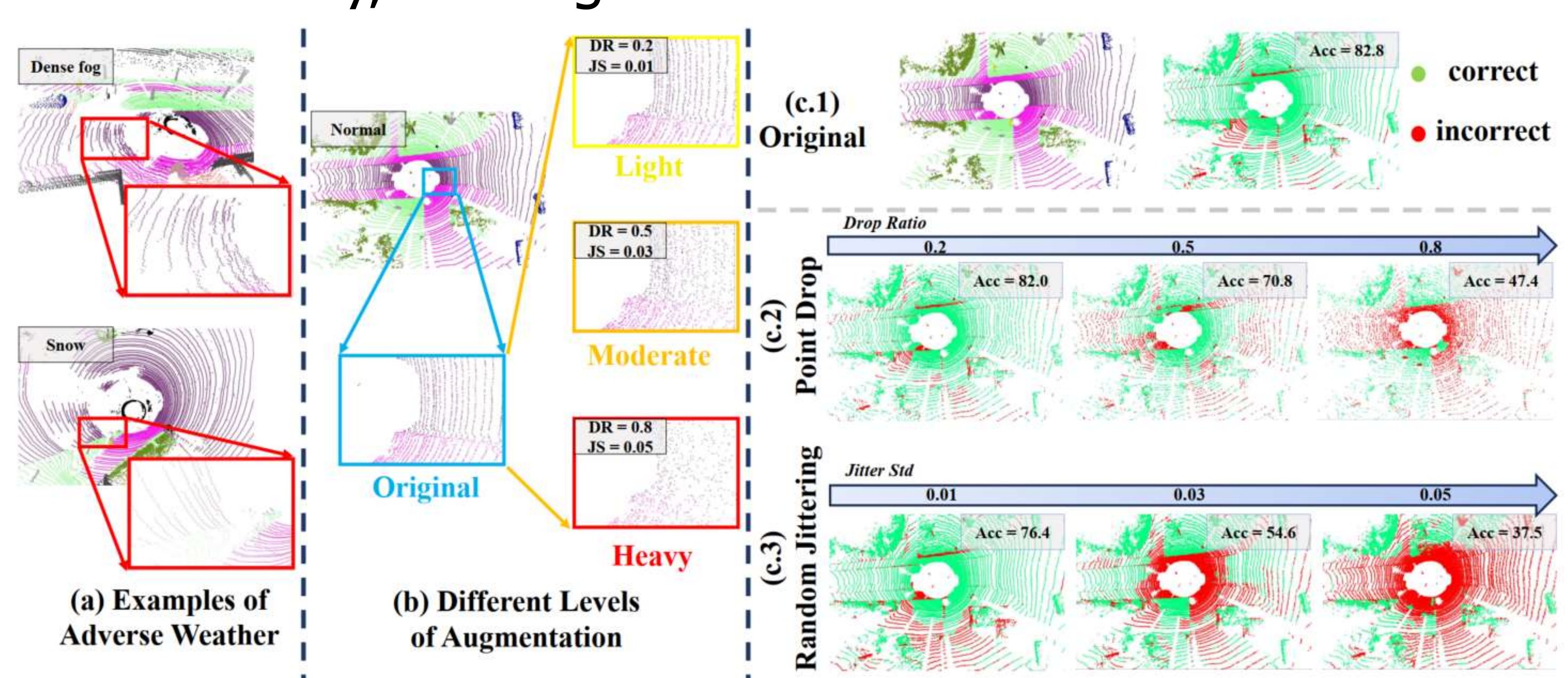
Introduction

Goal. Domain-generalized LiDAR semantic segmentation aims to train on labeled point clouds under normal weather and generalize to unseen adverse-weather conditions without accessing target-domain data during training.

Mainstream Paradigm. Augmentation-based training improves robustness by simulating weather-induced distortions, e.g., geometric perturbation and point drop, on source-domain point clouds.

Problem. Although augmentation improves robustness, its effectiveness is limited by a fundamental trade-off:

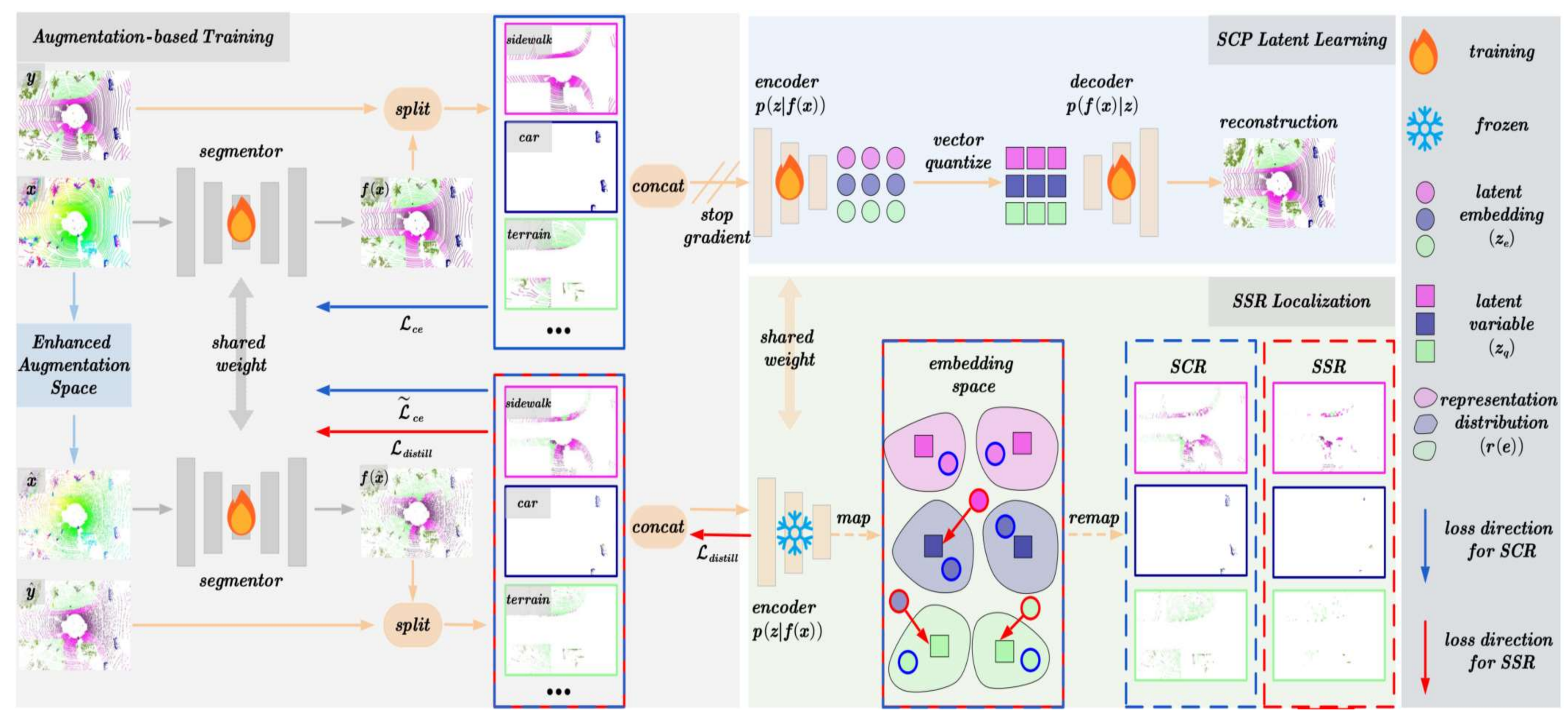
- Mild augmentations are insufficient to cover severe weather conditions.
- Aggressive augmentations distort point cloud geometry and density, causing semantic shift.



Motivation. To fully exploit diverse augmentations, we need to distinguish two factors in augmented point clouds:

- Semantic confusion: inherent ambiguity between similar classes, which should be optimized with original labels.
- Semantic shift: augmentation-induced label–semantics mismatch, which requires adapted supervision.

Method



Enhanced Augmentation Space. We enlarge the augmentation space to better simulate diverse disturbances:

- Random jittering for geometric perturbation
- Point drop for beam attenuation / occlusion effects

We sample augmentation magnitudes from a broad range to obtain light-to-heavy distortions during training.

SCP Latent Learning. To capture the model's inherent semantic confusion, we learn class-wise discrete latent representations from predictions on original point clouds:

- Use a VQ-VAE-style latent learner
- Encode prediction patterns into class-specific codebooks
- Reconstruct prediction to enforce informative embeddings

This module builds a semantic confusion prior, which models normal class-wise ambiguity of the segmentation network.

SSR Localization. We localize semantic shift regions by treating semantic shift detection as anomaly detection.

- Use the frozen prior encoder to map augmented predictions into latent space.
- Compare augmented latent embeddings with the learned semantic confusion priors.

Region-Adaptive Optimization. After localization, we apply different supervision strategies to different regions:

- SCR: optimize with the original labels
- SSR: replace hard supervision with latent prior distillation, using the global nearest latent code as guidance

This avoids misleading optimization caused by label–semantic mismatch in heavily augmented regions.

Experiments

Quantitative Results on Different Benchmarks.

Table 1: Comparison results of [A] → [C]. ∗ denotes the reproduced result with the same backbone.

Methods	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	other-g.	build.	fence	veget.	trunk	terr.	pole	traf.	D-fog	L-fog	Rain	Snow	mIoU	gain
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	-
Baseline	67.1	**5.0**	28.1	38.5	14.6	45.8	8.3	13.8	40.1	16.1	26.1	3.3	71.6	52.7	53.8	33.9	39.2	25.3	12.7	30.7	30.1	29.7	25.3	31.4	+0.0
PointDR* (Xiao et al. 2023)	69.2	1.0	8.9	**41.9**	7.6	48.9	17.0	36.2	57.8	15.9	32.3	4.0	75.7	46.4	54.0	36.2	43.9	23.7	24.2	37.3	33.5	35.5	26.9	33.9	+2.5
DGUIL (Kim et al. 2023)	78.2	2.5	33.0	29.7	6.1	49.8	0.8	40.9	67.3	7.2	38.0	2.2	**79.8**	**54.4**	**64.1**	36.8	52.3	31.0	40.0	36.3	34.5	35.5	33.3	37.6	+6.2
WADG (Du et al. 2024)	72.0	0.0	32.9	37.0	1.9	37.7	6.8	52.9	59.9	10.7	31.8	2.2	76.0	48.8	62.7	34.0	49.3	23.6	20.4	39.5	32.5	31.7	29.4	34.8	+3.4
DGLSS (He et al. 2024)	69.6	0.8	42.8	34.4	8.9	41.9	12.8	44.5	52.0	14.5	30.8	**6.0**	77.8	51.1	57.6	38.9	43.2	29.7	30.6	34.2	34.8	36.2	32.1	36.2	+4.8
LiDARWeather (Park et al. 2024)	86.1	4.8	13.8	39.7	**26.6**	**55.4**	8.5	50.4	63.7	14.9	37.9	5.5	75.2	52.7	60.4	39.7	44.9	30.1	40.8	36.0	37.5	37.6	33.1	39.5	+8.1
NTN (Park et al. 2025)	83.3	3.7	31.3	36.2	18.2	53.3	6.8	**55.9**	67.2	**18.1**	37.2	5.4	72.1	41.8	58.0	36.0	46.0	28.2	39.8	35.3	35.1	35.7	32.4	38.9	+7.5
A3Point (ours)	**88.3**	4.1	**57.5**	29.0	7.6	45.3	**24.0**	46.4	**69.2**	16.9	**38.4**	3.3	74.8	48.2	63.1	**42.9**	**49.2**	**32.6**	**41.8**	**41.1**	**38.5**	**38.2**	**37.2**	**41.3**	+9.9

Table 2: Comparison results of [B] → [C]. ∗ denotes the reproduced result with the same backbone.

Methods	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	other-g.	build.	fence	veget.	trunk	terr.	pole	traf.	D-fog	L-fog	Rain	Snow	mIoU	gain
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	-
Baseline	33.8	1.7	3.3	15.5	0.2	25.5	1.6	3.4	15.3	9.2	16.8	0.1	33.4	21.9	39.5	18.7	**44.0**	8.8	0.8	15.2	16.0	16.8	12.8	15.5	+0.0
PointDR* (Xiao et al. 2023)	41.1	2.8	3.4	18.1	0.2	31.3	**2.8**	3.3	34.4	**10.2**	19.7	1.0	52.7	22.0	48.5	21.3	38.3	19.2	5.6	19.1	20.3	25.3	19.0	19.8	+4.3
WADG (Du et al. 2024)	33.8	1.1	2.9	17.0	0.2	26.8	1.0	4.3	53.9	5.0	20.6	2.2	**64.3**	27.1	53.8	27.0	37.0	**28.6**	8.6	21.6	23.4	27.2	21.4	21.9	+6.4
LiDARWeather (Park et al. 2024)	39.3	2.9	0.9	19.4	0.8	27.7	2.2	3.8	42.5	9.4	21.6	0.3	51.9	33.5	47.4	23.1	33.3	23.2	6.8	19.0	21.2	23.1	17.3	20.5	+5.0
NTN (Park et al. 2025)	48.4	1.5	2.4	19.4	0.2	29.1	3.2	8.9	43.5	6.7	20.5	0.0	52.2	30.1	49.8	20.0	32.9	24.7	7.5	-	-	-	-	21.1	+5.6
A3Point (ours)	**76.7**	**4.0**	**5.0**	**29.6**	**1.3**	**35.1**	1.7	**9.5**	**55.4**	3.9	**24.0**	**3.5**	61.7	**34.5**	**60.1**	**34.1**	33.3	28.1	**14.8**	**26.8**	**26.6**	**31.9**	**28.6**	**27.2**	+11.7

Qualitative Results on SemanticKITTI→SemanticSTF.