LaserMix for Semi-Supervised LiDAR Semantic Segmentation

Lingdong Kong[1,2,3,*], Jiawei Ren[1,*], Liang Pan[1], Ziwei Liu[1]

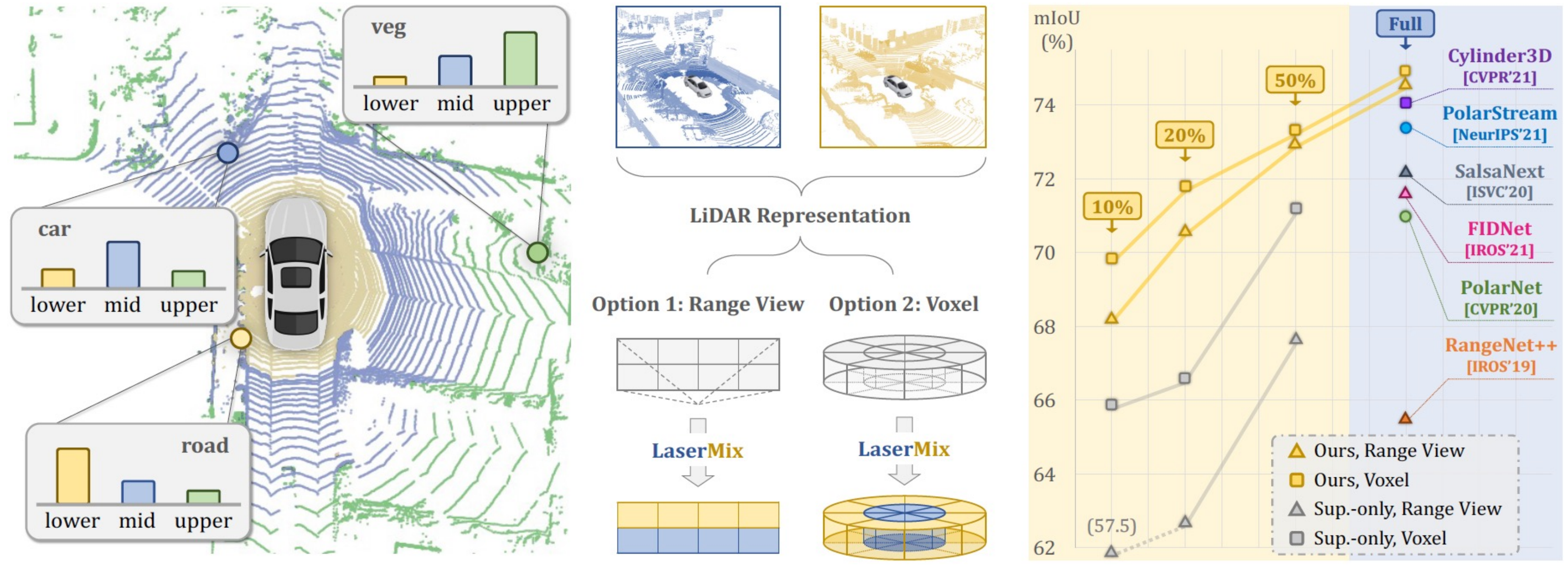
[1]S-Lab, Nanyang Technological University, [2]National University of Singapore, [3]CNRS@CREATE

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE · NUS National University of Singapore · JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Motivation & Contribution

TL;DR

- LaserMix is a semi-supervised learning framework designed for LiDAR segmentation; it leverages spatial prior of driving scenes to construct low-variation areas via laser beam mixing and encourages the model to make confident and consistent predictions before and after mixing.



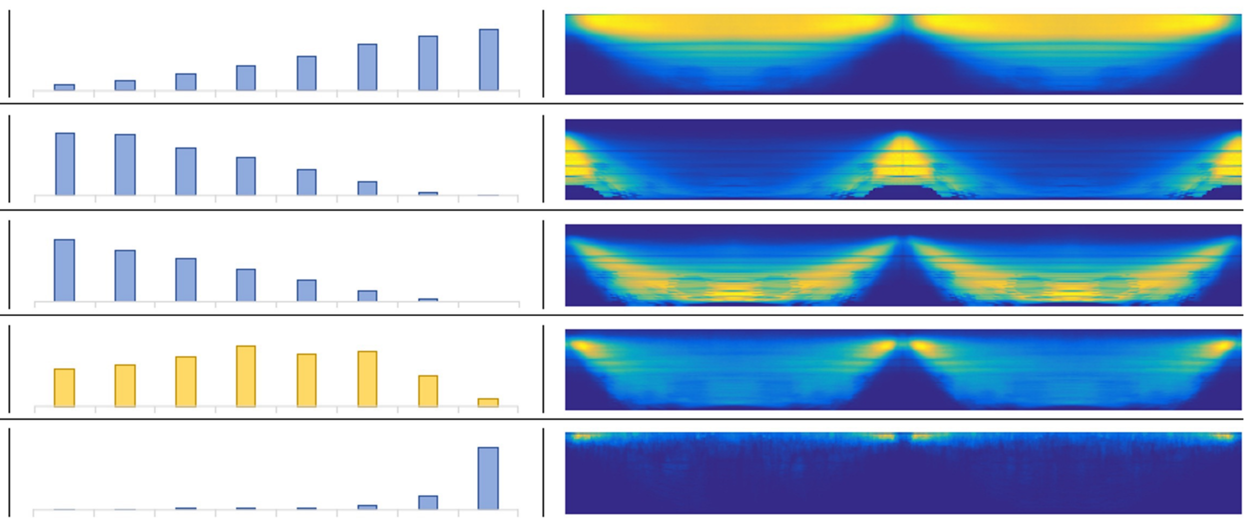
Properties

- Generic to LiDAR representations and can be universally applied.
- Statistically grounded with theoretical explanations and applicability.
- Effective across datasets and semi-supervised scenarios; competitive results over full supervision counterparts with 2x to 5x fewer annotations.

Spatial Prior in 3D

- The distribution of real-world objects and backgrounds is exhibiting a strong correlation to their spatial positions in the LiDAR scan.

vegetation	static	24.825%
road	static	22.545%
sidewalk	static	16.353%
car	dynamic	4.657%
traffic-sign	static	0.061%

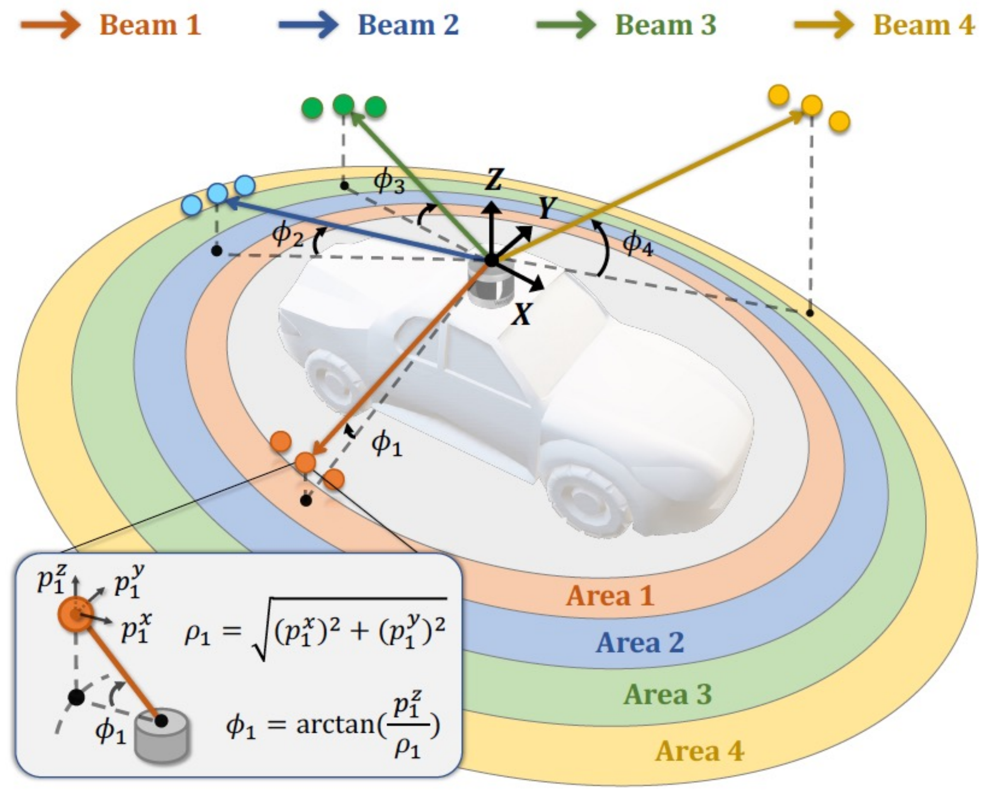


- We leverage such strong distribution patterns on the laser beams and thus propose the laser partition. Our framework effectively "excites" spatial prior and mixes LiDAR scans in an efficient and scalable manner.

Methodology

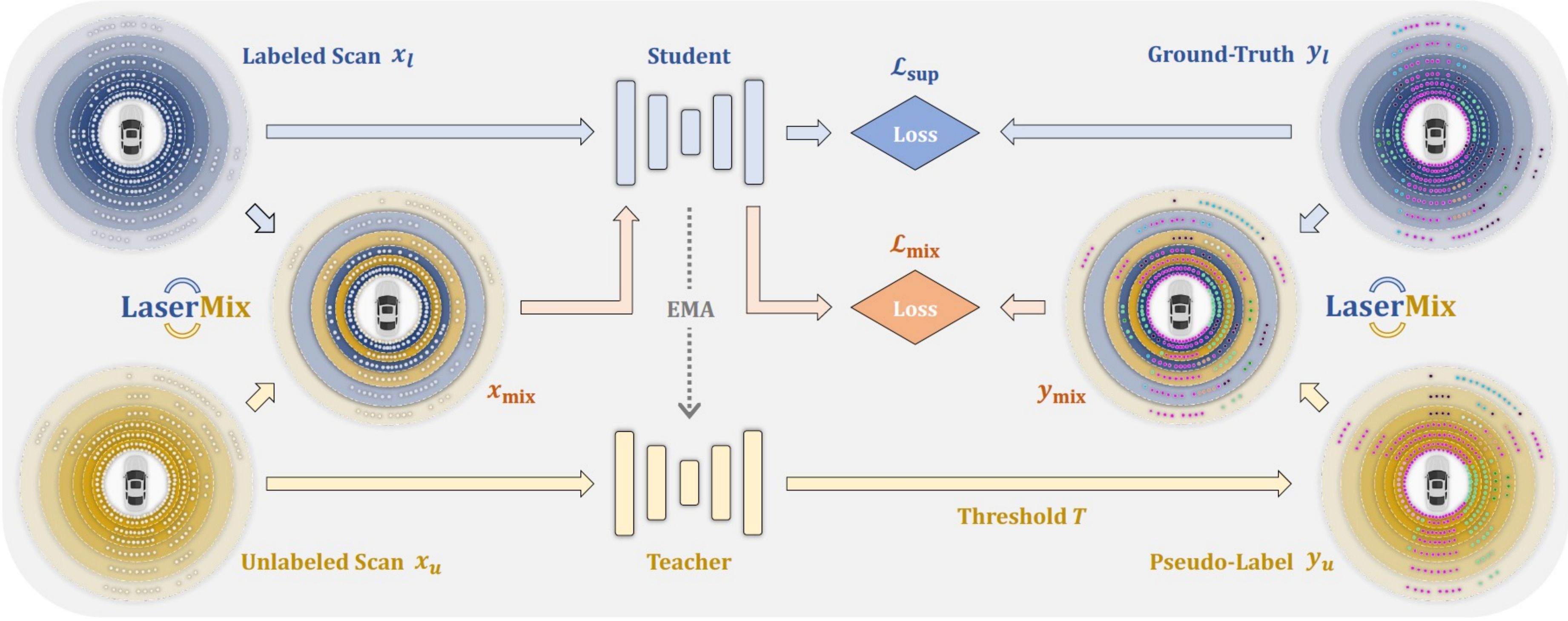
Laser Beam Partition

- LiDAR scans serve as a perfect reflection of real-world patterns, which are highly dependent on the spatial areas in the LiDAR-centered 3D coordinates.
- We group LiDAR points whose inclinations are within the same range into the same area, as depicted in the color regions.



Three-Step Procedure

- Partitioning the captured LiDAR scan into low-variation areas.
- Efficiently mixing every area in the LiDAR scan with foreign data.
- Encouraging the LiDAR segmentation models to make confident and consistent predictions on the same area in different mixing.



Semi-Supervised Learning Framework

- LaserMix mixes two LiDAR scans by intertwining the areas so that the neighbors of each area are filled with data from the other scan.
- As a result, we obtain the prediction on all areas of two scans from only two predictions, which reduces the computational cost effectively.
- A two-branch framework is constructed, where the Student and Teacher nets take the mixed and unlabeled scans, respectively, as their inputs.

Experiments & Analysis

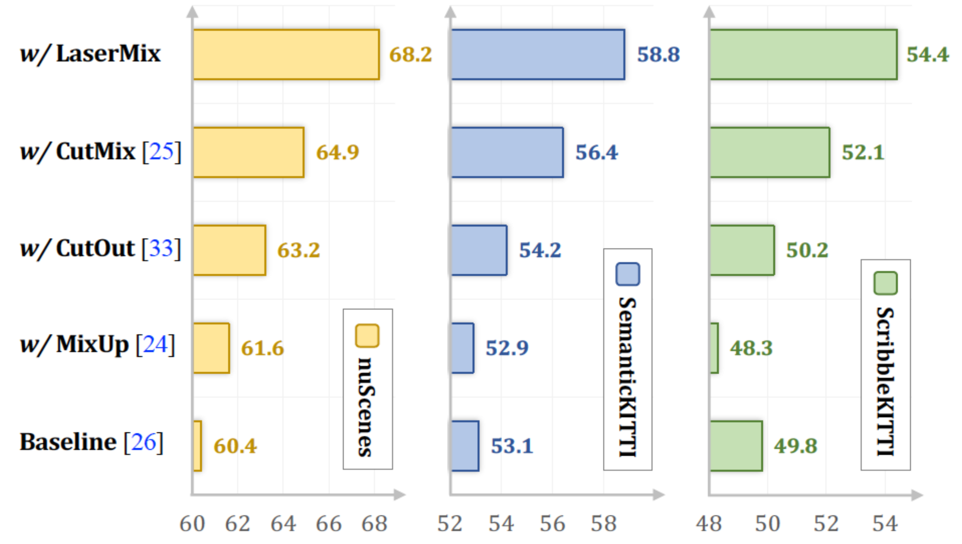
Data-Efficient LiDAR Segmentation

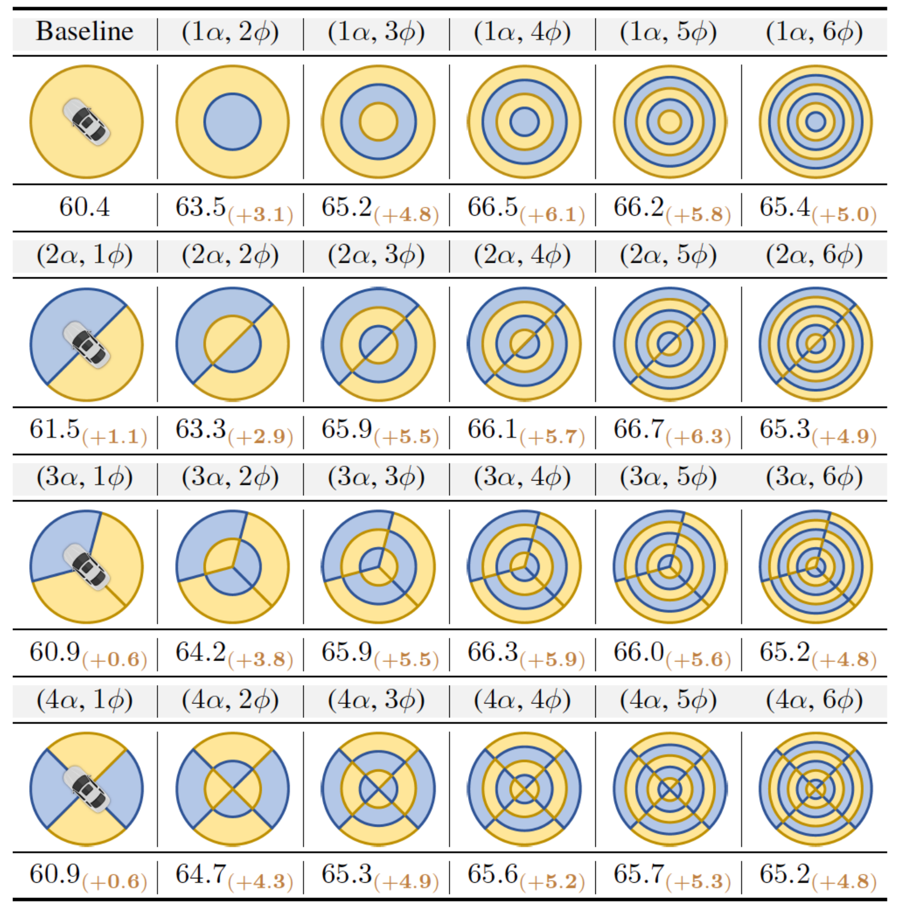
- LaserMix's effectiveness is verified across different sensor setups, LiDAR representations, annotation budgets, and data splits. Superior results over prior arts have been constantly achieved under all tested settings.

Repr.	Method	nuScenes [15] 1%	10%	20%	50%	SemanticKITTI [16] 1%	10%	20%	50%	ScribbleKITTI [4] 1%	10%	20%	50%
Range View	*Sup.-only*	38.3	57.5	62.7	67.6	36.2	52.2	55.9	57.2	33.1	47.7	49.9	52.5
	MeanTeacher [26]	42.1	60.4	65.4	69.4	37.5	53.1	56.1	57.4	34.2	49.8	51.6	53.3
	CBST [30]	40.9	60.5	64.3	69.3	39.9	53.4	56.1	56.9	35.7	50.7	52.7	54.6
	CutMix-Seg [29]	43.8	63.9	64.8	69.8	37.4	54.3	56.6	57.6	36.7	50.7	52.9	54.3
	CPS [13]	40.7	60.8	64.9	68.0	36.5	52.3	56.3	57.4	33.7	50.0	52.8	54.6
	LaserMix (Ours)	**49.5**	**68.2**	**70.6**	**73.0**	**43.4**	**58.8**	**59.4**	**61.4**	**38.3**	**54.4**	**55.6**	**58.7**
	Δ ↑	+11.2	+10.7	+7.9	+5.4	+7.2	+6.6	+3.5	+4.2	+5.2	+6.7	+5.7	+6.2
Voxel	*Sup.-only*	50.9	65.9	66.6	71.2	45.4	56.1	57.8	58.7	39.2	48.0	52.1	53.8
	MeanTeacher [26]	51.6	66.0	67.1	71.7	45.4	57.1	59.2	60.0	41.0	50.1	52.8	53.9
	CBST [30]	53.0	66.5	69.6	71.6	48.8	58.3	59.4	59.7	41.5	50.6	53.3	54.5
	CPS [13]	52.9	66.3	70.0	72.5	46.7	58.7	59.6	60.5	41.4	51.8	53.9	54.8
	LaserMix (Ours)	**55.3**	**69.9**	**71.8**	**73.2**	**50.6**	**60.0**	**61.9**	**62.3**	**44.2**	**53.7**	**55.1**	**56.8**
	Δ ↑	+4.4	+4.0	+5.2	+2.0	+5.2	+3.9	+4.1	+3.6	+5.0	+5.7	+3.0	+3.0

Ablation Study

- LaserMix exhibits superiority over other mixing-based techniques in encouraging spatial prior in 3D.





Summary & Conclusion

- We proposed LaserMix, an effective and scalable framework for data-efficient scene understanding in autonomous driving.
- Our code and other resources are openly accessible at MMDetection3D platform.

Paper Code