Durham University

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

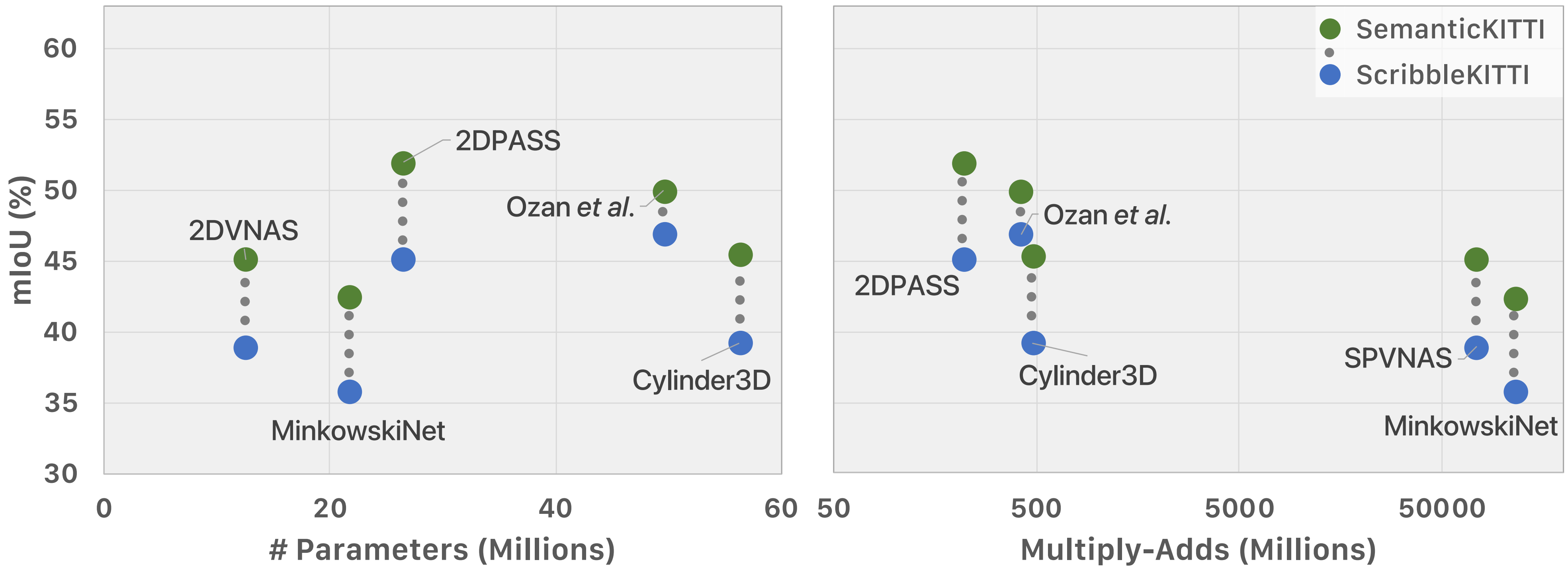
Less is More
Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation

Poster: June 21, 2023 @ WED-AM-108

Li Li[1], Hubert P. H. Shum[1], Toby P. Breckon[1,2]

Department of {Computer Science[1] | Engineering[2]}
Durham University

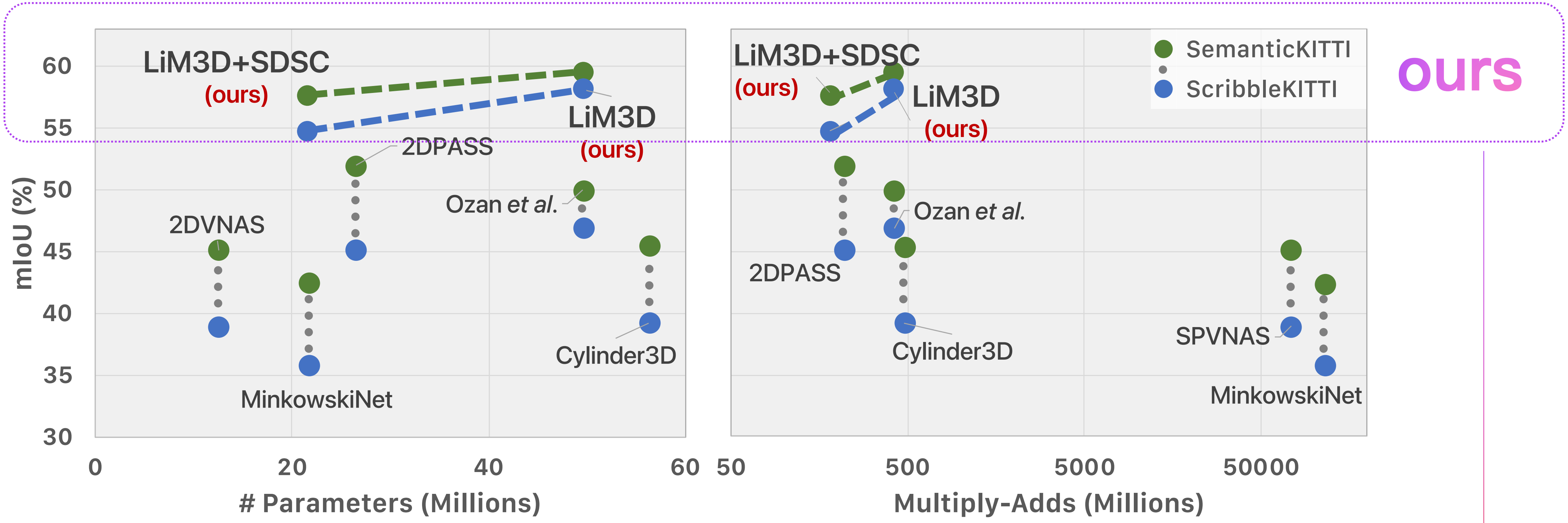
LESS IS MORE



Previous methods

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

LESS IS MORE



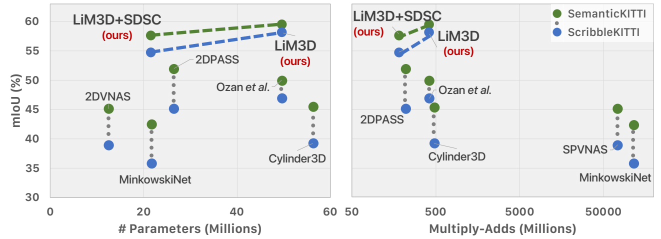
2.3x model size reduction

641x fewer multiply–adds

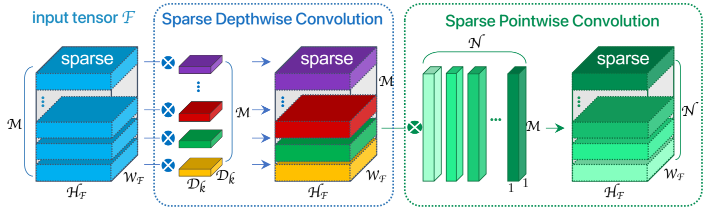
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Contributions

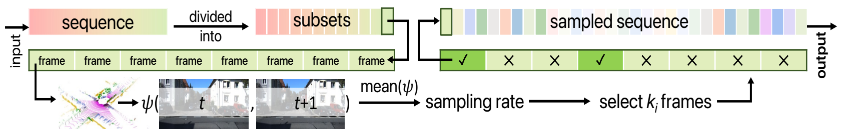
1. semantic segmentation: *less* **parameters** and **(*more*) superior accuracy**.



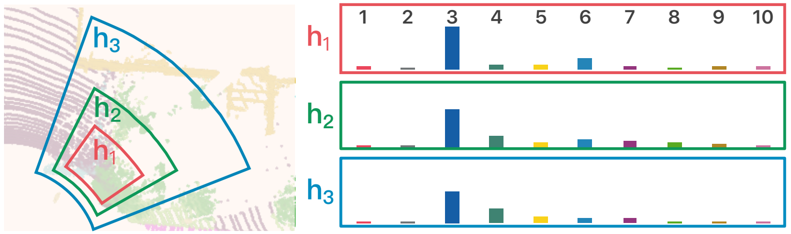
2. **Sparse Depthwise Separable Convolution**: to reduce trainable network without loss.



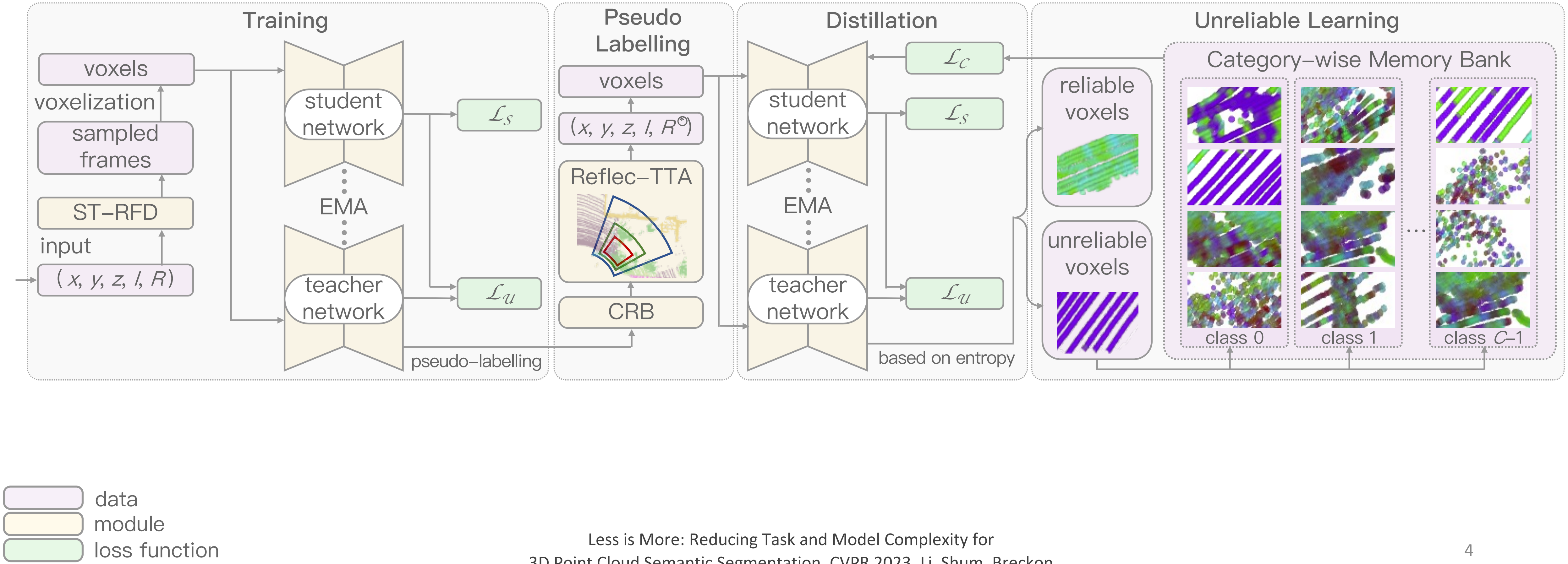
3. **Spatio-Temporal Redundant Frame Downsampling**: to remove temporal redundancy and annotation requirements.



4. Soft pseudo-labeling method informed by **LiDAR reflectivity**: to use limited data annotation effectively.

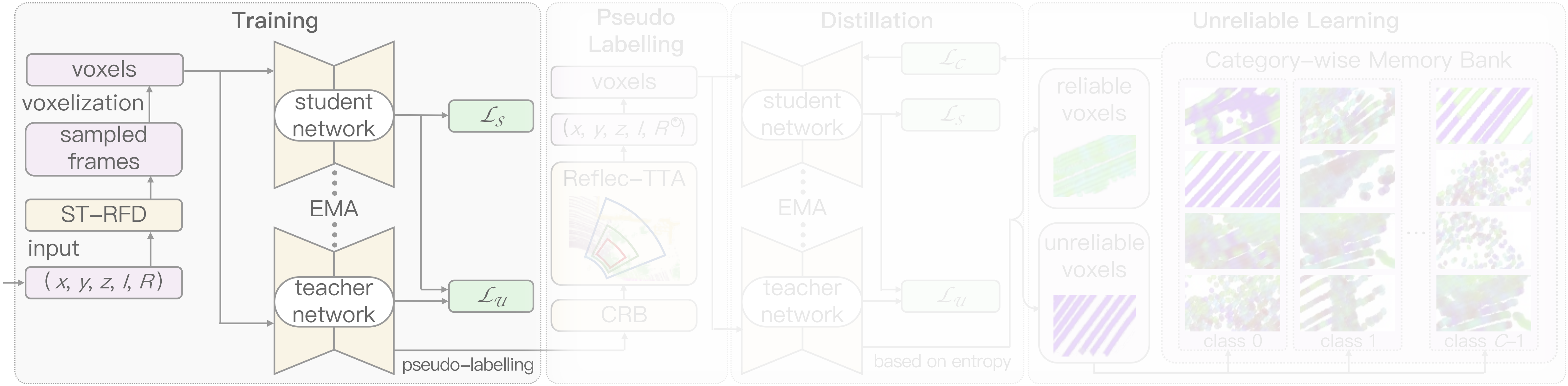


Our Proposed Architecture



Our Proposed Architecture

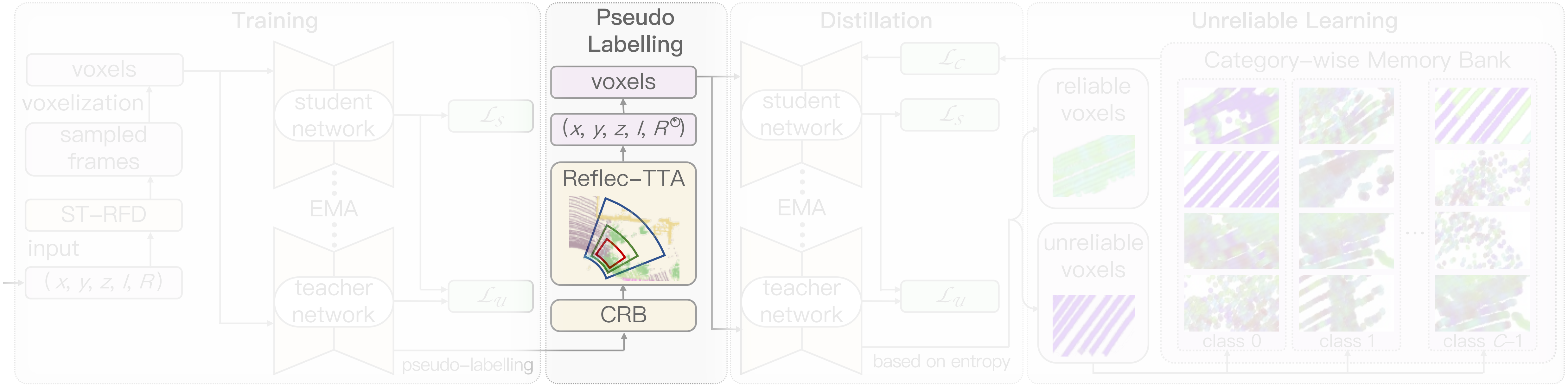
training > to utilize **reflectivity–prior descriptors** and adapt the **Mean Teacher** framework to generate high–quality pseudo–labels



data
module
loss function

Our Proposed Architecture

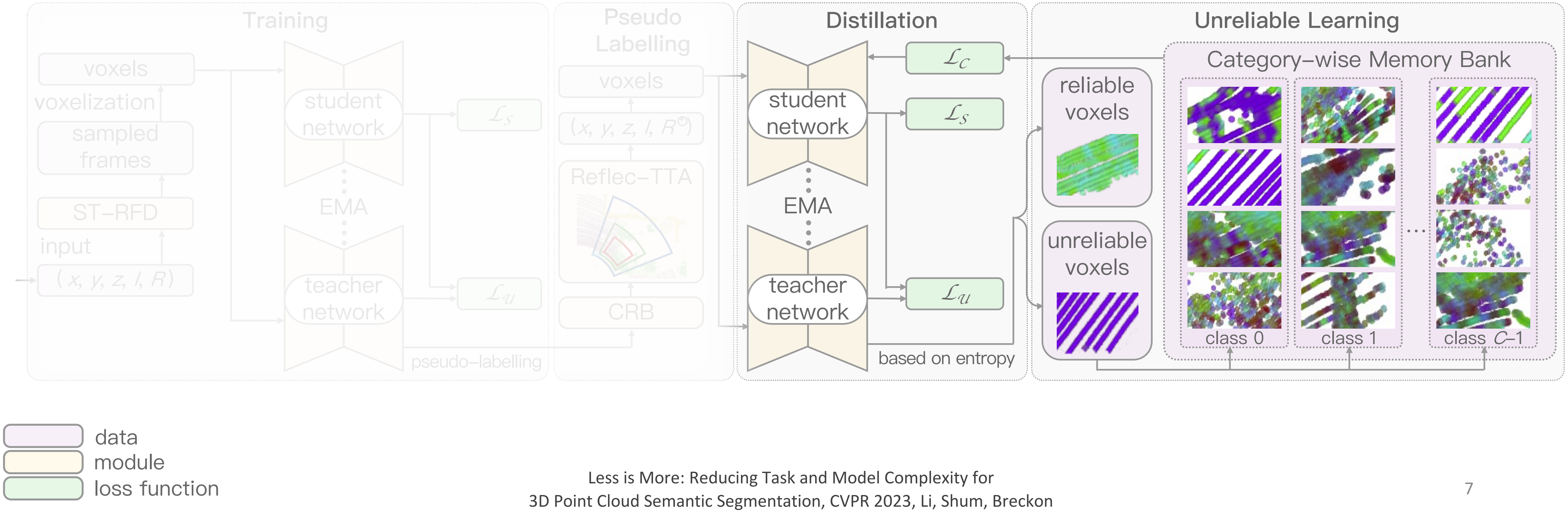
pseudo labelling > to fix the trained teacher model prediction in a **CRB** manner, expanding dataset with **Reflec−TTA** during test time



data
module
loss function

Our Proposed Architecture

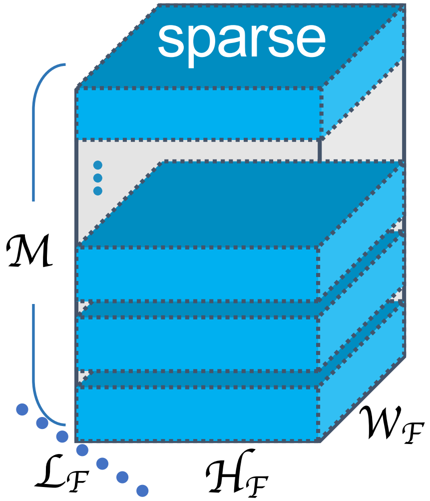
distillation & unreliable learning > to train on the generated pseudo–labels, and **utilize unreliable pseudo–labels** in a memory bank for improved discrimination



Sparse Depthwise Separable Convolution

to reduce trainable network without loss

Input tensor $\mathcal{F}$



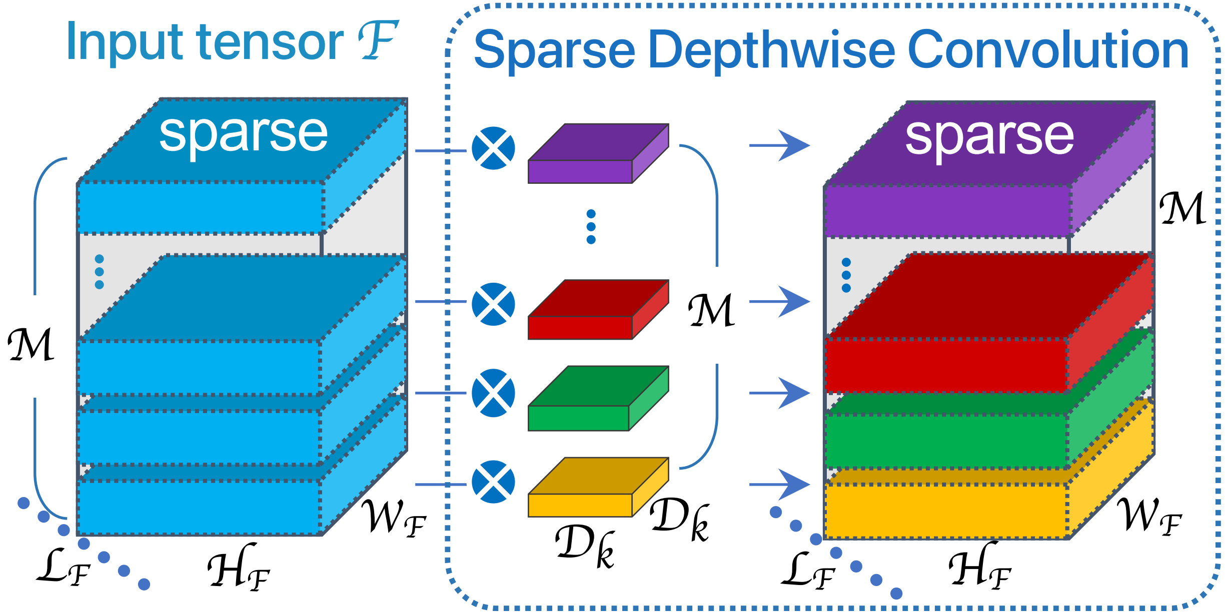
1 taking 3D voxels as input

submanifold sparse convolution

pointwise convolution

Sparse Depthwise Separable Convolution

to reduce trainable network without loss



2 going through the Sparse Depthwise Convolution

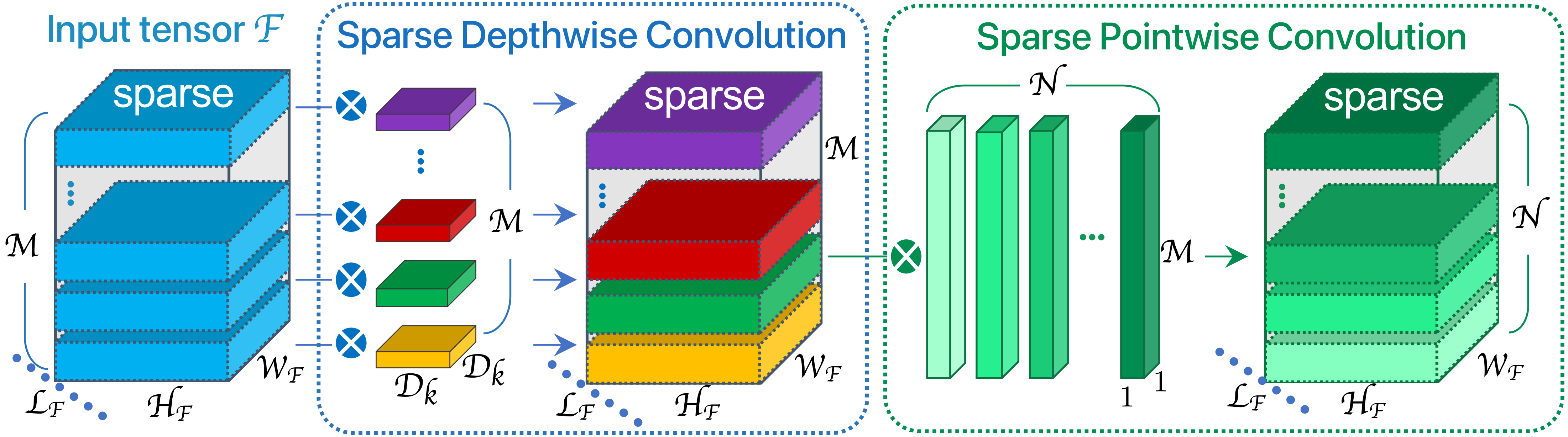
to perform convolution with the trainable parameter reduction

submanifold sparse convolution

pointwise convolution

Sparse Depthwise Separable Convolution

to reduce trainable network without loss



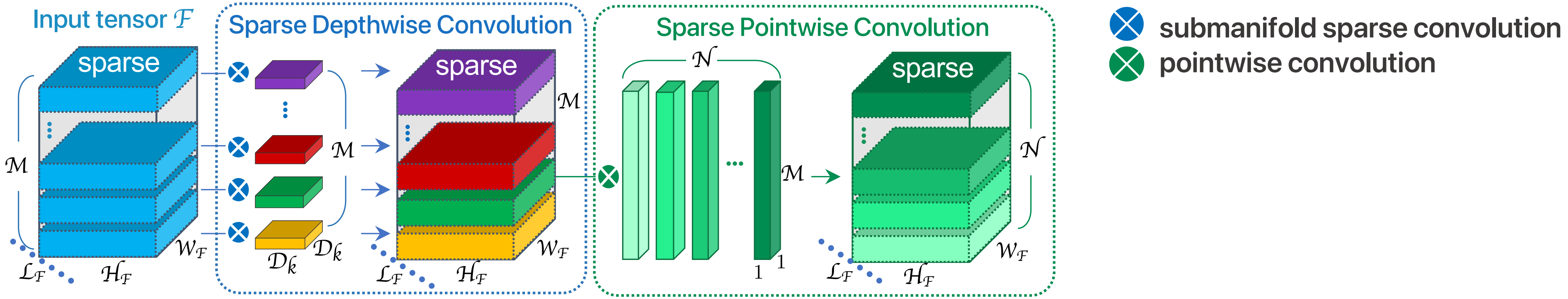
3 going through the Sparse Pointwise Convolution
to mix the information across different channels

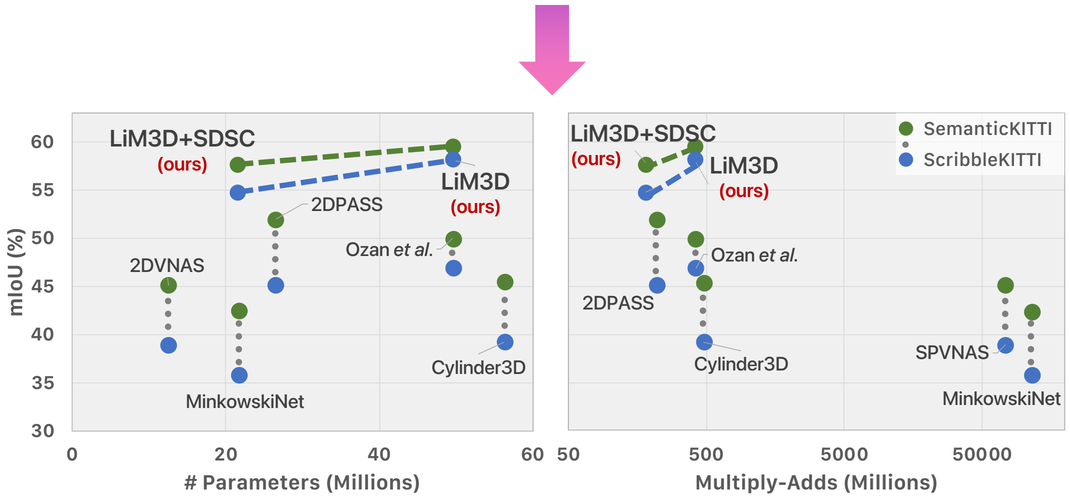
submanifold sparse convolution
pointwise convolution

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Sparse Depthwise Separable Convolution

to reduce trainable network without loss





with our Sparse Depthwise Separable Convolution we can achieve:

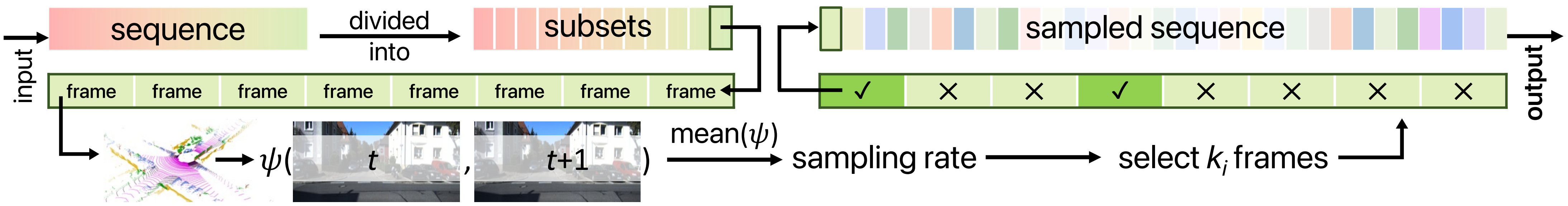
2.3x model size reduction

641x fewer multiply–adds

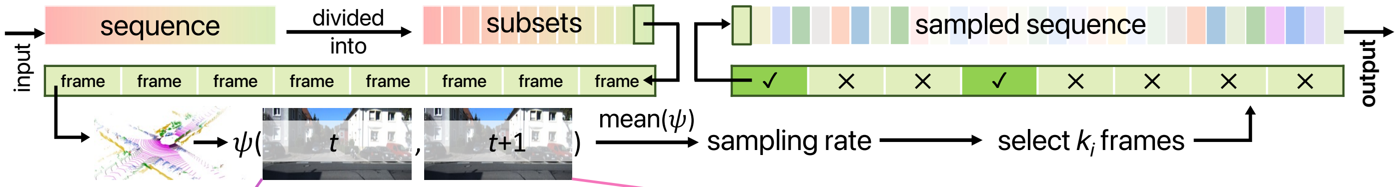
Less is More: Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Spatio-Temporal Redundant Frame Downsampling (ST-RFD)

Using ST-RFD to extract a maximally diverse data subset for training by **removing temporal redundancy** and hence future **annotation requirements**



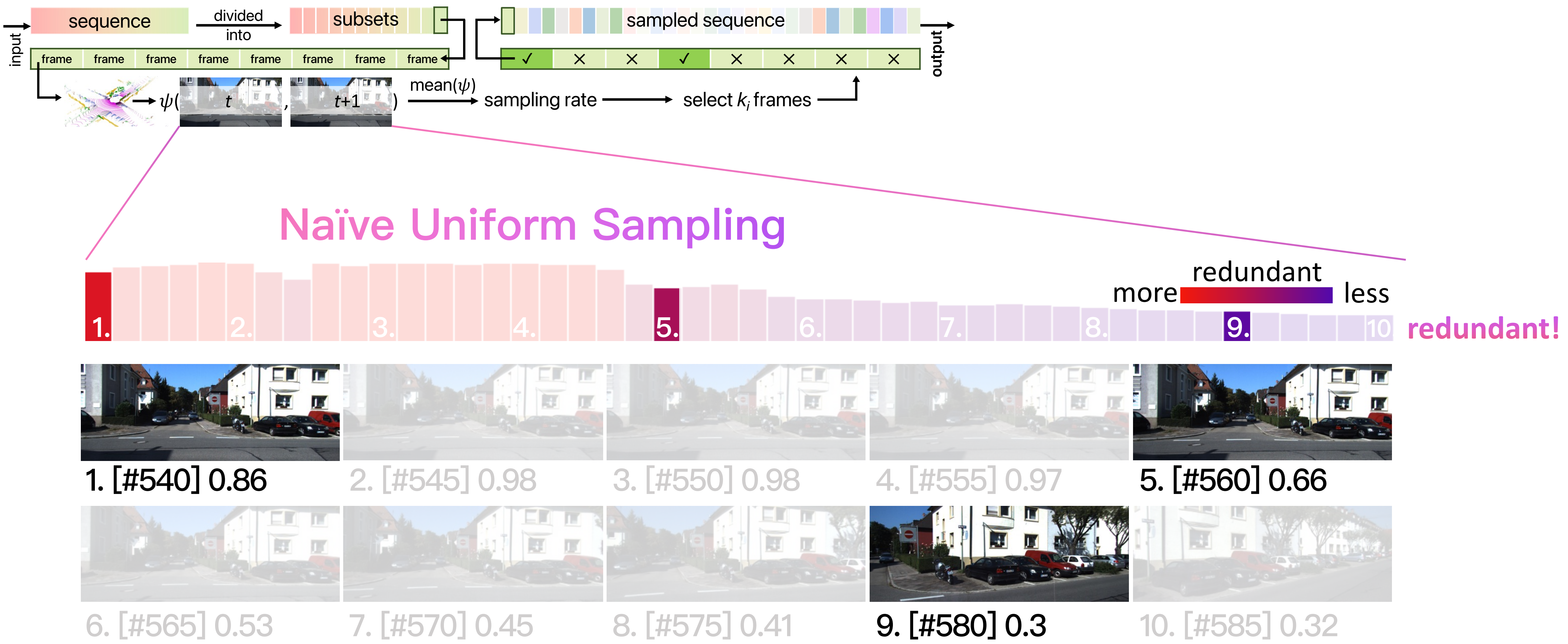
Spatio-Temporal Redundant Frame Downsampling (ST-RFD)



computing the similarity between temporally adjacent frames

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Spatio-Temporal Redundant Frame Downsampling (ST-RFD)

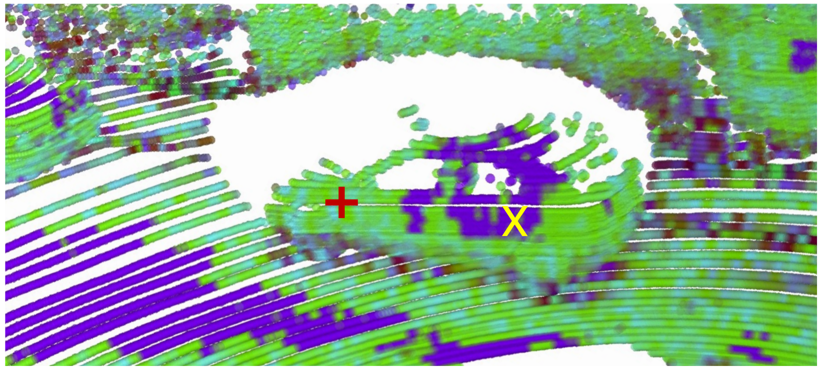


Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Spatio–Temporal Redundant Frame Downsampling (ST–RFD)

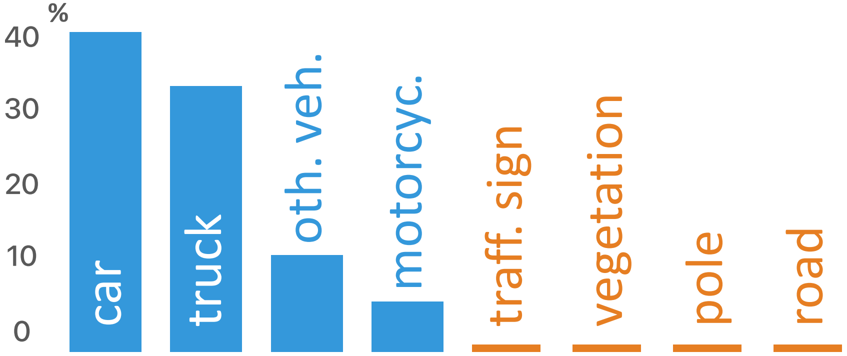
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Using Unreliable Pseudo-labels
to Make Full Use of All Available Labels



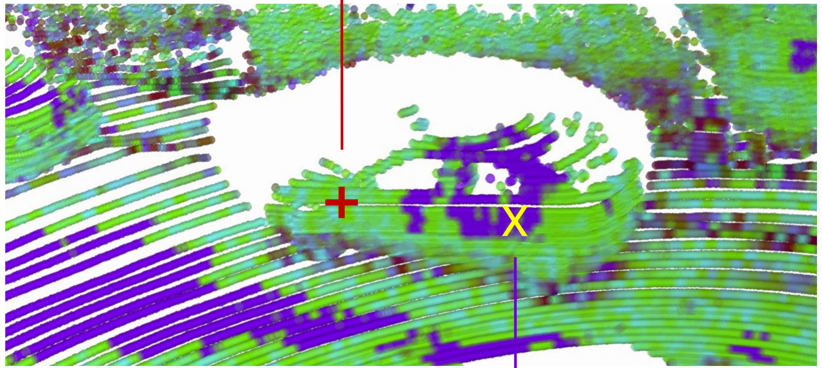
$$\mathcal{L}_C = -\frac{1}{C}\sum_{c=0}^{C-1} \mathop{\mathbb{E}}_{\mathbf{E}_c}\left[\log \frac{f(\mathbf{e}_c, \mathbf{e}_c^{+}, \tau)}{\sum_{\mathbf{e}_{c,j}^{-}\in\mathbf{E}_c^{-}} f(\mathbf{e}_c, \mathbf{e}_{c,j}^{-}, \tau)}\right]$$

$$= -\frac{1}{C}\sum_{c=0}^{C-1} \mathop{\mathbb{E}}_{\mathbf{E}_c}\left[\log \frac{\exp(\langle \mathbf{e}_c, \mathbf{e}_c^{+}\rangle/\tau)}{\exp\left(\langle \mathbf{e}_c, \mathbf{e}_c^{+}\rangle/\tau\right) + \sum_{j=1}^{N-1} \exp\left(\langle \mathbf{e}_c, \mathbf{e}_{c,j}^{-}\rangle/\tau\right)}\right]$$



Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

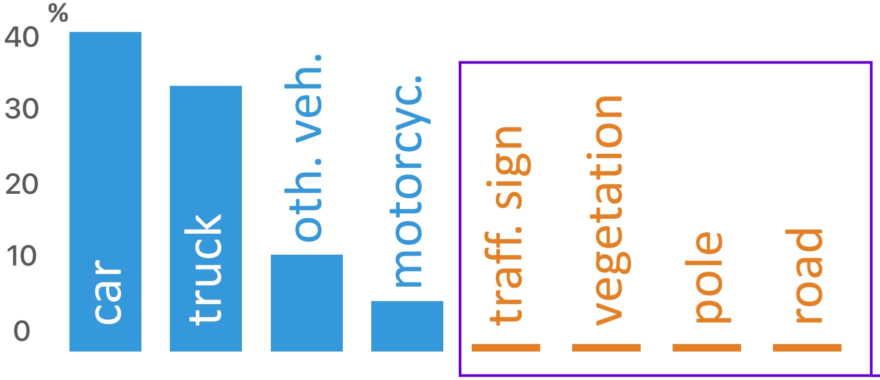
Using Unreliable Pseudo-labels
to Make Full Use of All Available Labels



positive sample

$$\mathcal{L}_C = -\frac{1}{C}\sum_{c=0}^{C-1} \mathbb{E}_{\mathbf{E}_c}\left[\log \frac{f(\mathbf{e}_c, \boxed{\mathbf{e}_c^+}, \tau)}{\sum_{\mathbf{e}_{c,j}^- \in \mathbf{E}_c^-} f(\mathbf{e}_c, \mathbf{e}_{c,j}^-, \tau)}\right]$$

$$= -\frac{1}{C}\sum_{c=0}^{C-1} \mathbb{E}_{\mathbf{E}_c}\left[\log \frac{\exp(\langle \mathbf{e}_c, \mathbf{e}_c^+ \rangle / \tau)}{\exp\left(\langle \mathbf{e}_c, \mathbf{e}_c^+ \rangle / \tau\right) + \sum_{j=1}^{N-1} \exp\left(\langle \mathbf{e}_c, \boxed{\mathbf{e}_{c,j}^-} \rangle / \tau\right)}\right]$$

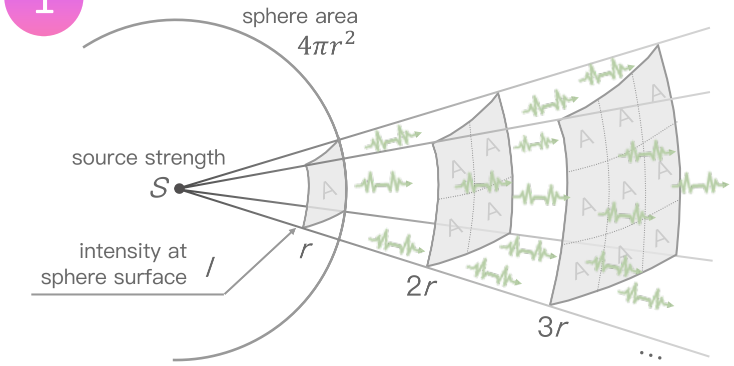


negatives sample

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Using reflectivity-based Test Time Augmentation
to enhance performance of false or non-existent pseudo-labels

1



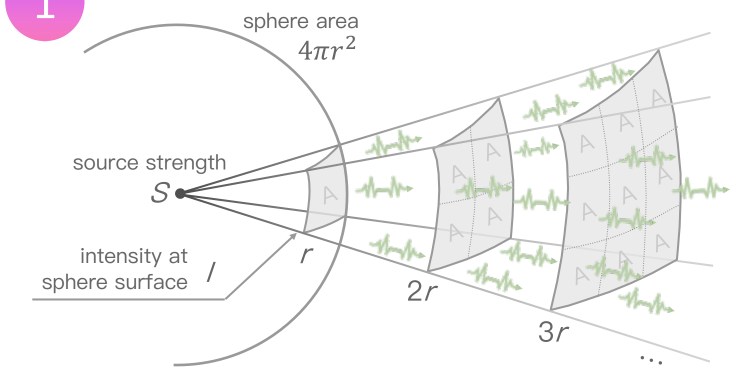
reflectivity

$$R = Ir^2 = \frac{S}{4\pi r^2} \cdot r^2 \propto S$$

Reflectivity is
a **distance-normalized intensity** feature

Using reflectivity-based Test Time Augmentation
to enhance performance of false or non-existent pseudo-labels

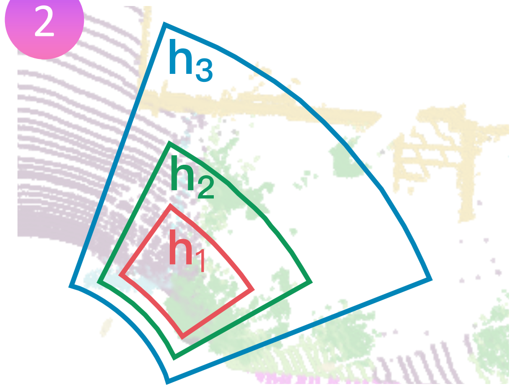
1

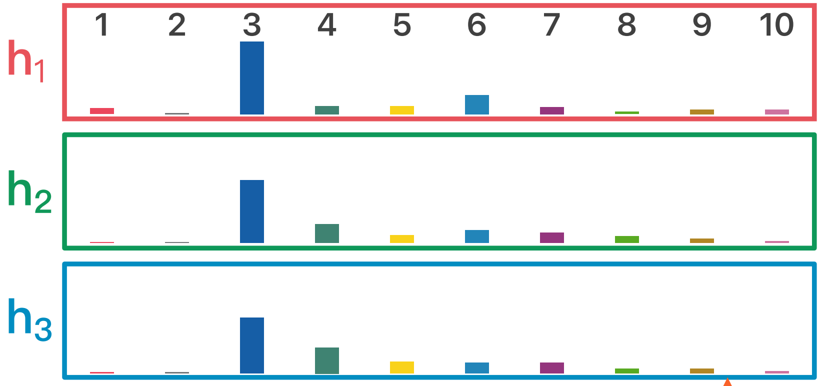


reflectivity

$$R = Ir^2 = \frac{S}{4\pi r^2} \cdot r^2 \propto S$$

2





$$\mathbf{h}_i = \left\{h_i^{(k)} \mid k \in [1, N_b]\right\} \in \mathbb{R}^{N_b}, i \in [1, s],$$

$$h_i^{(k)} = \#\left\{\boxed{R_j} \in r_k, \forall j \mid p_j \in b_i\right\}$$

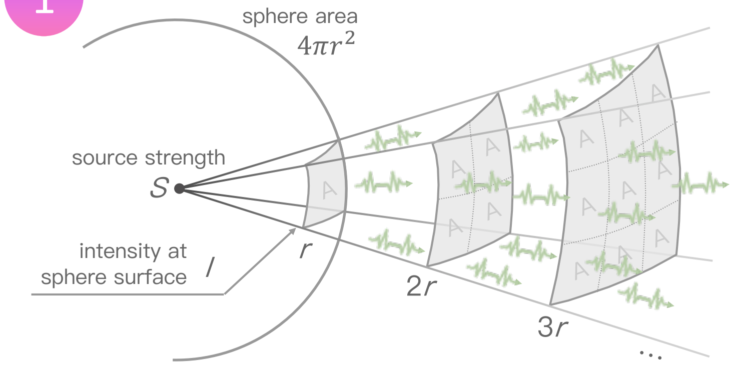
$$r_k = [(k-1)/N_b, k/N_b), k \in [1, N_b].$$

we apply various sizes of bins in cylindrical coordinates to analyze the intrinsic point distribution at varying resolutions (shown in h_1, h_2 and h_3).

Using reflectivity-based Test Time Augmentation

to enhance performance of false or non-existent pseudo-labels

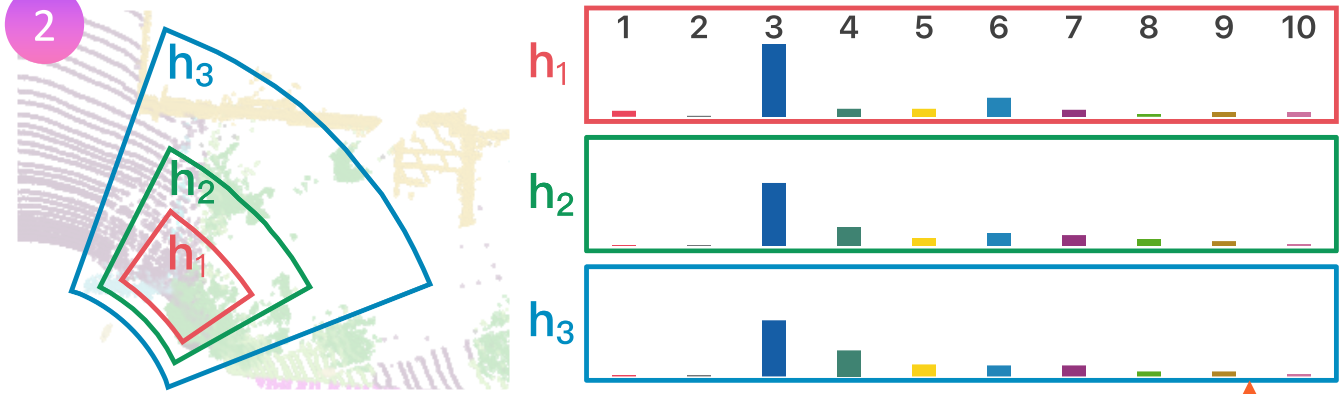
1



reflectivity

$$R = Ir^2 = \frac{S}{4\pi r^2} \cdot r^2 \propto S$$

2



$$\mathbf{h}_i = \left\{ h_i^{(k)} \mid k \in [1, N_b] \right\} \in \mathbb{R}^{N_b}, i \in [1, s],$$

$$h_i^{(k)} = \#\left\{ \boxed{R_j} \in r_k, \forall j \mid p_j \in b_i \right\}$$

$$r_k = [(k-1)/N_b, k/N_b), k \in [1, N_b].$$

3

normalize

augmentation

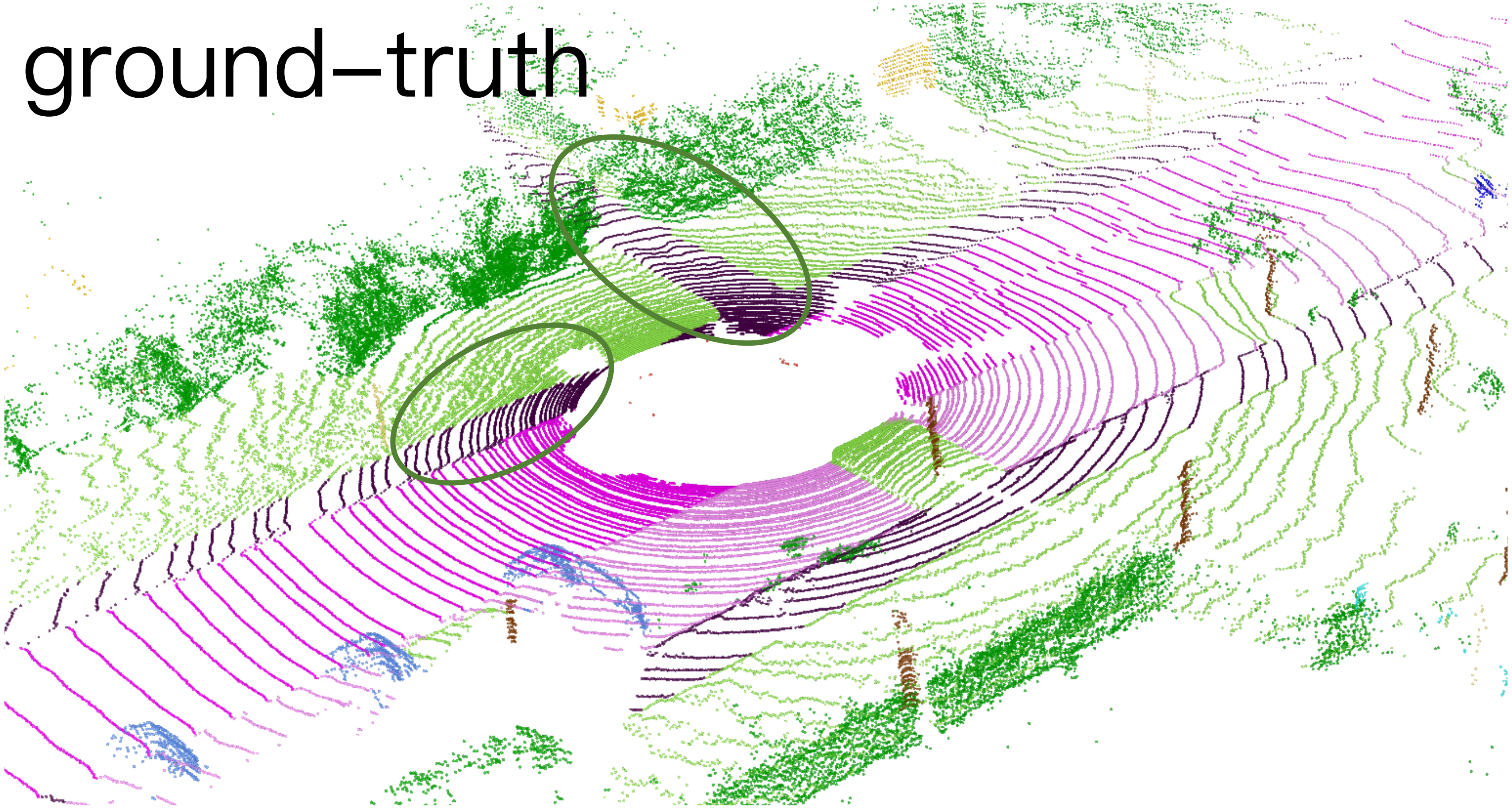
$$\begin{cases} R^{\circledast} = \{\mathbf{h}_i / \max(\mathbf{h}_i) \mid i \in [1, s]\} \in \mathbb{R}^{sN_b} \\ P^{\circledast} = \{p \mid (x, y, z, I, R^{\circledast}) \in \mathbb{R}^{sN_b+4}\} \end{cases}$$

we then normalize it and append to the point set

Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

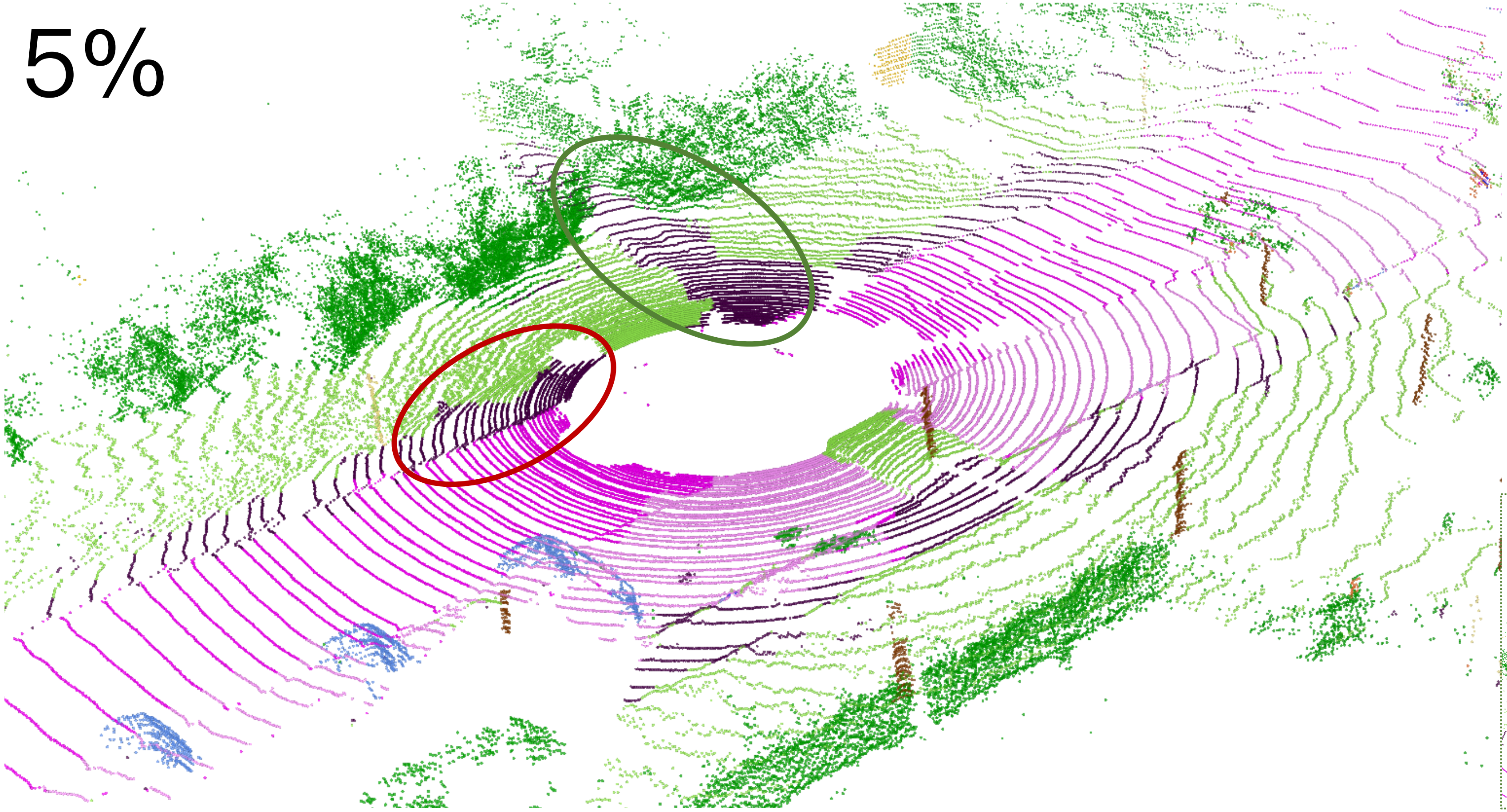
ground–truth

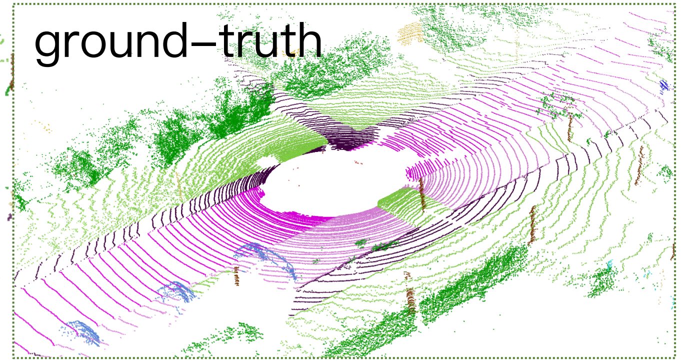


Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

5%



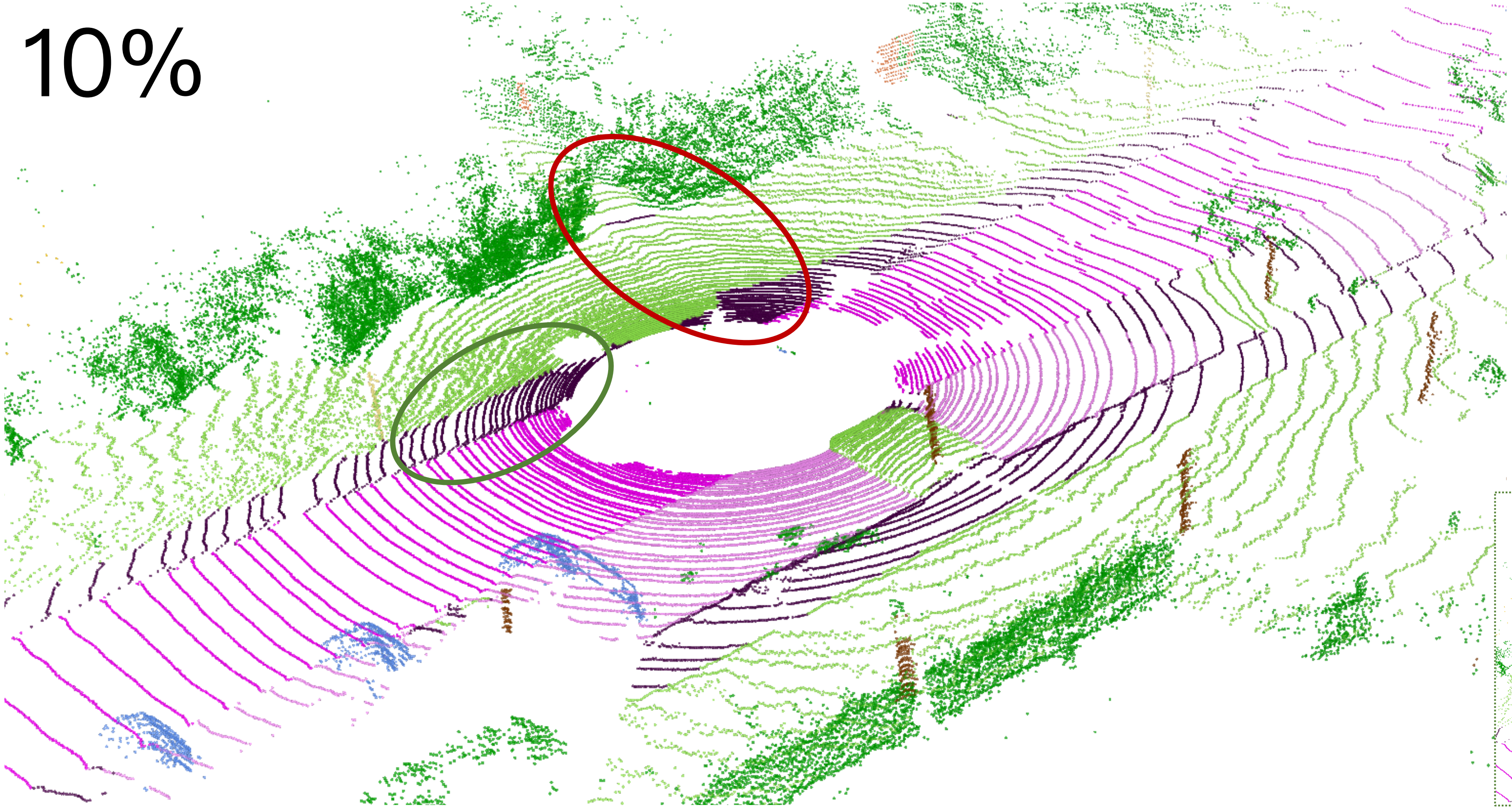


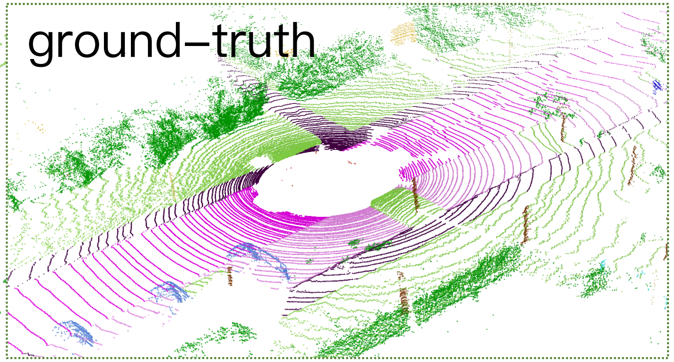
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

10%



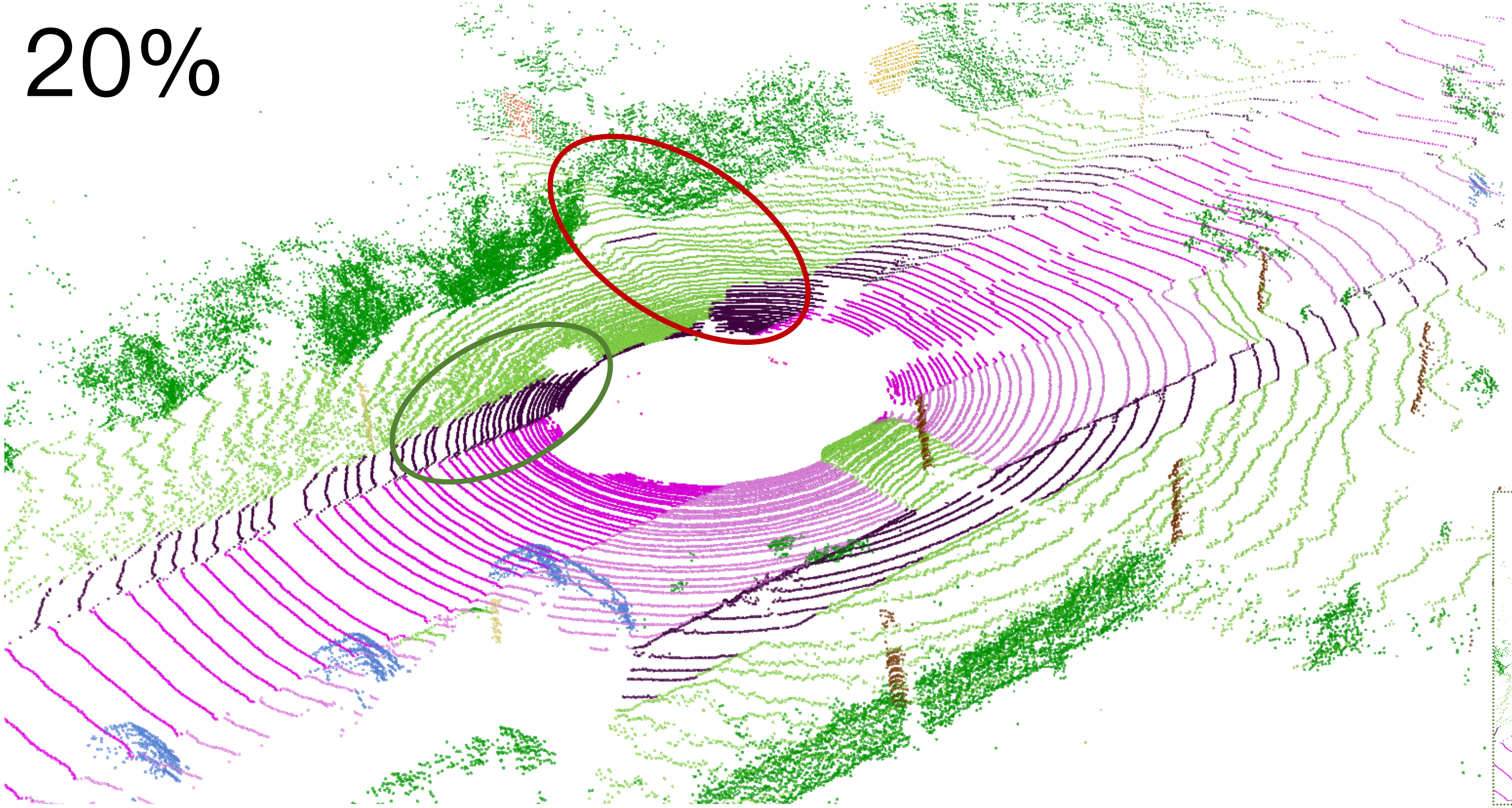


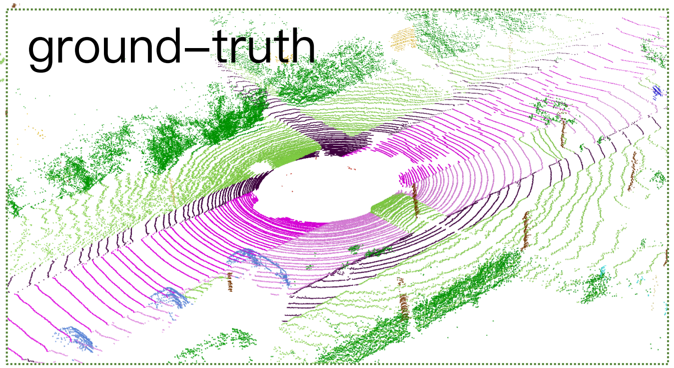
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

20%

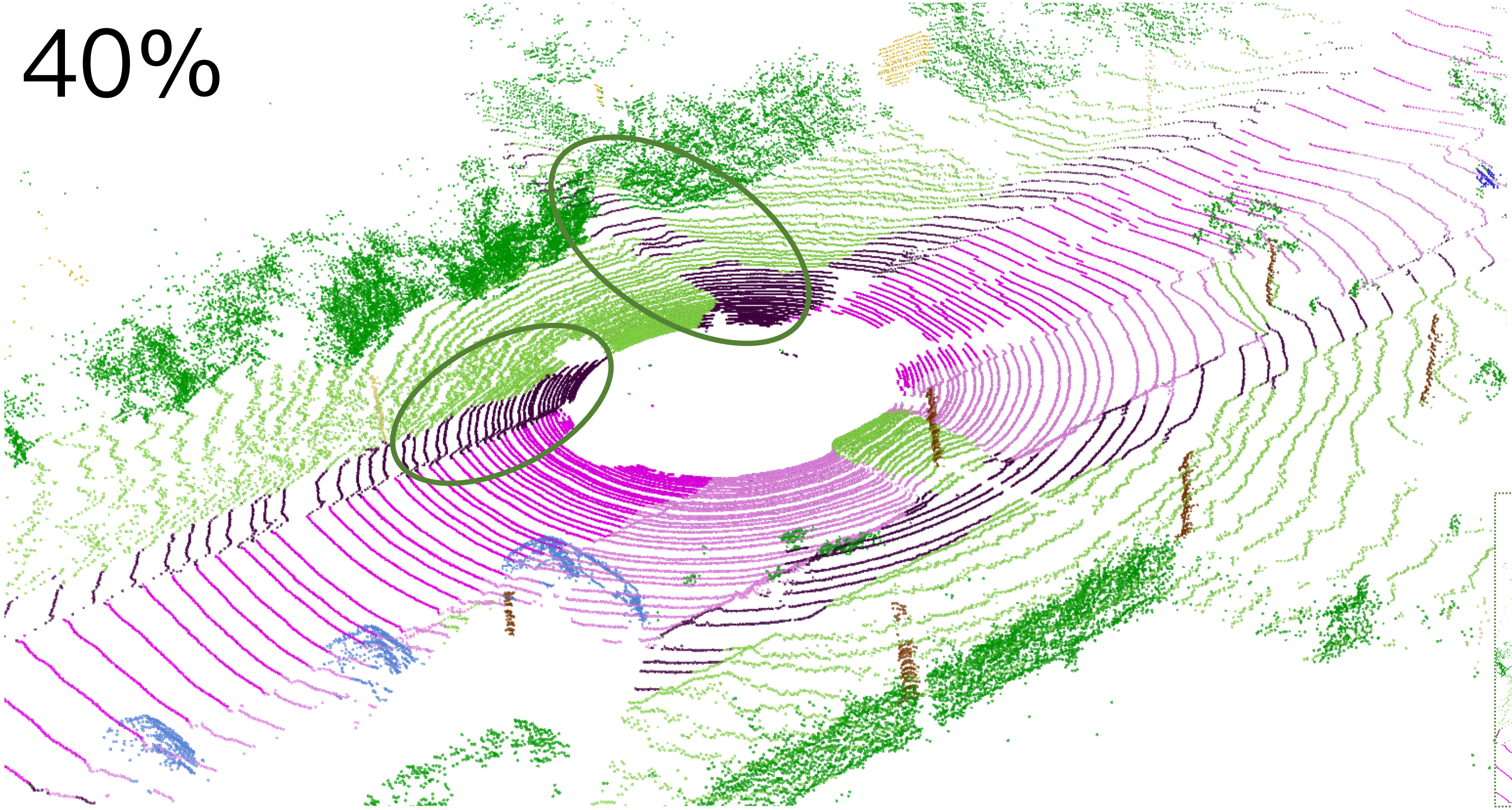


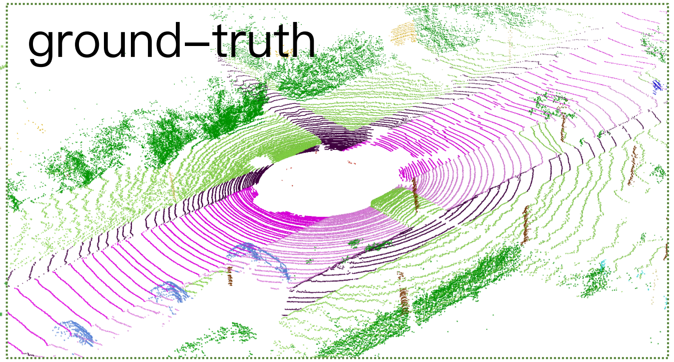


Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results
Comparing {5%, 10%, 20%, 40%} labeled splits

40%

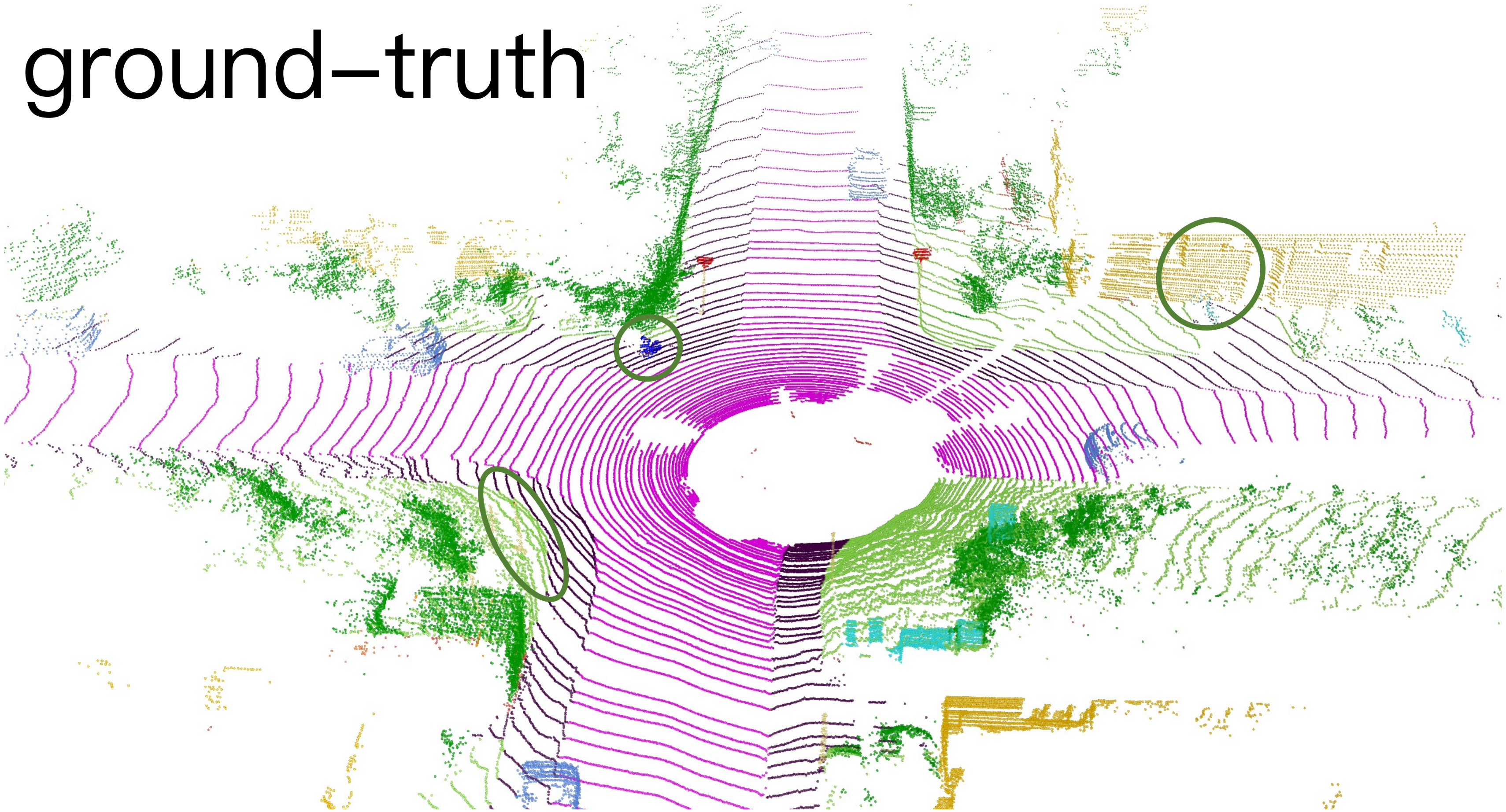




Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results
Comparing {5%, 10%, 20%, 40%} labeled splits

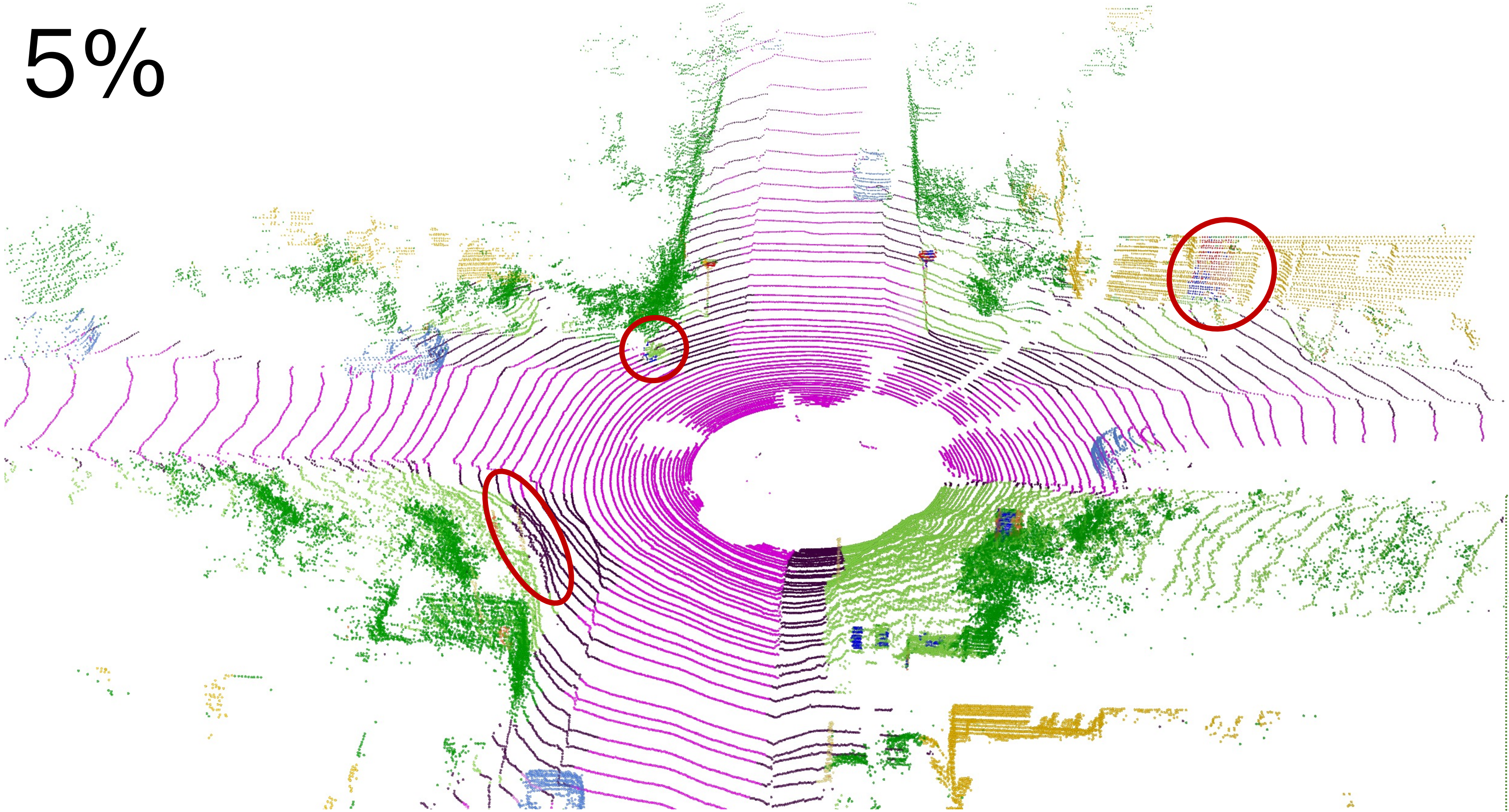
ground-truth

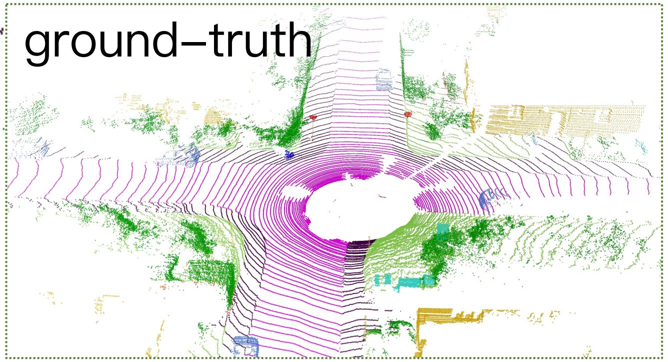


Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

5%



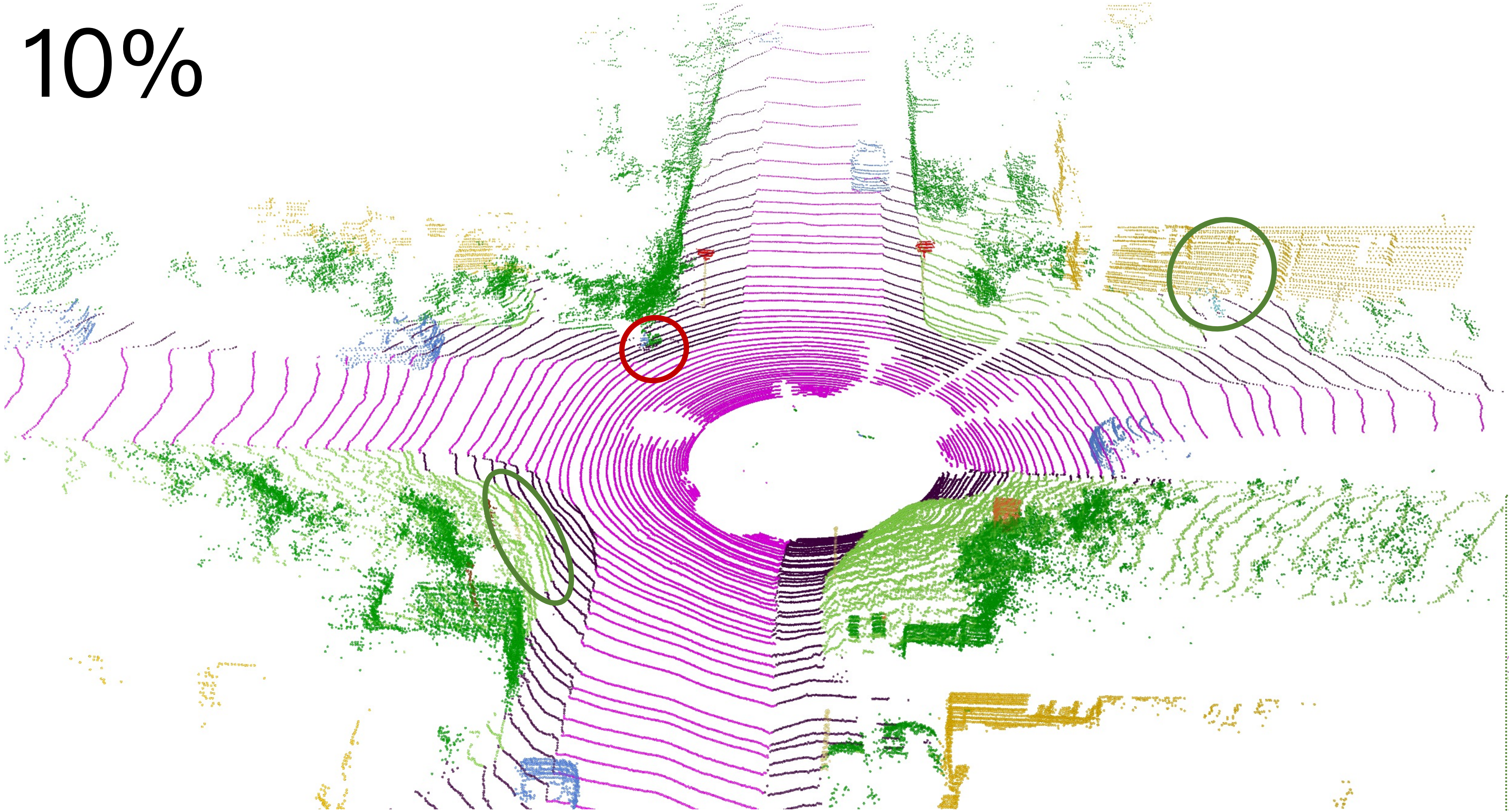


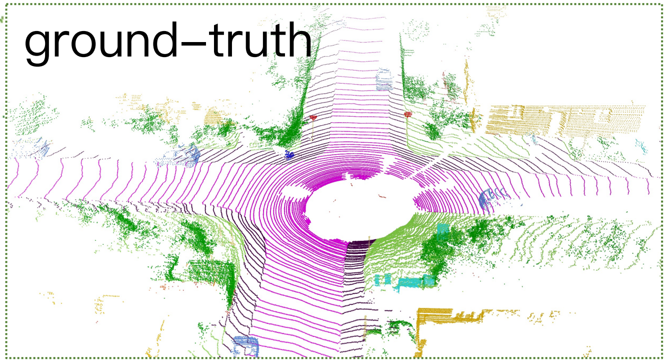
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

10%



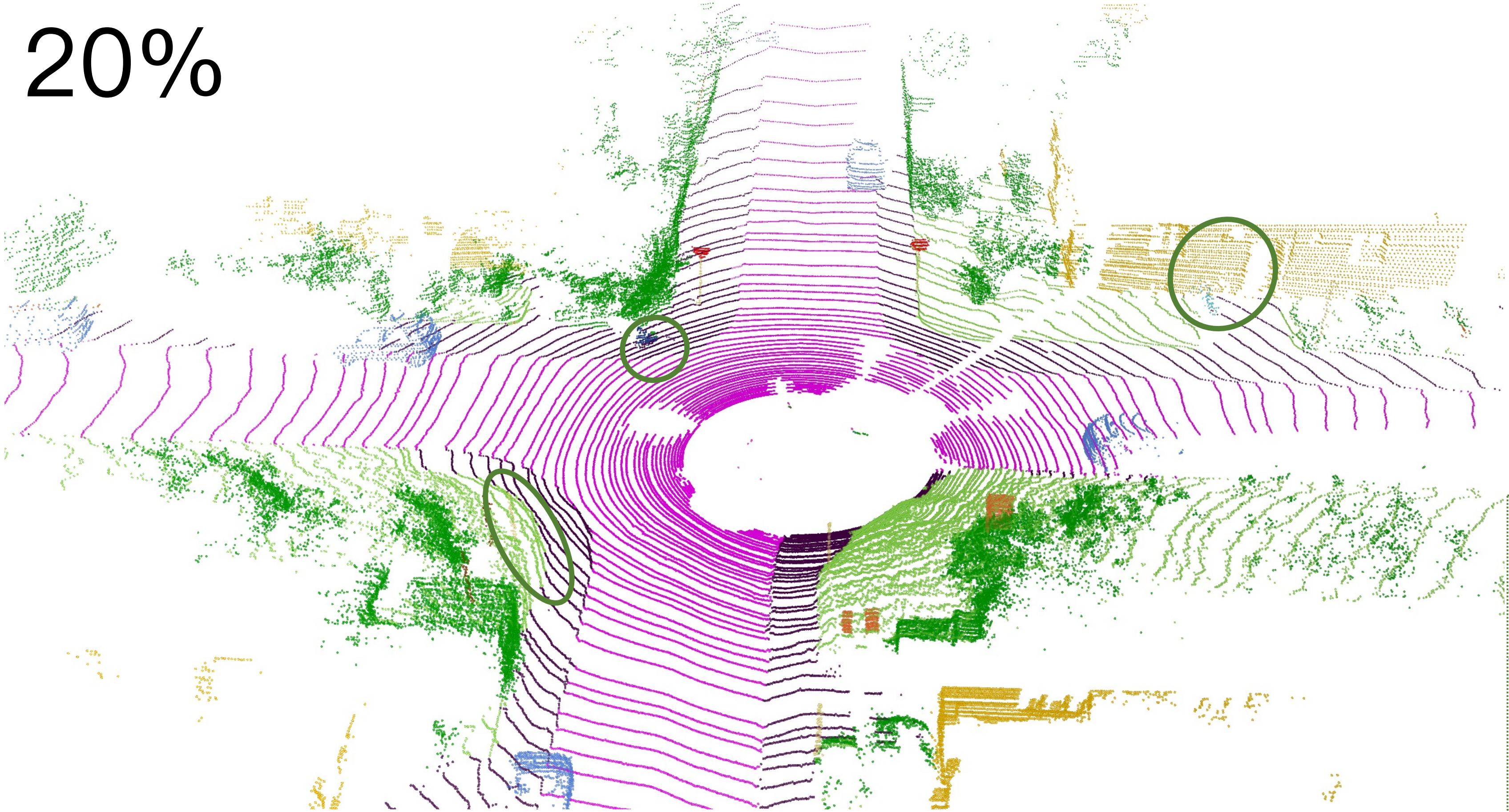


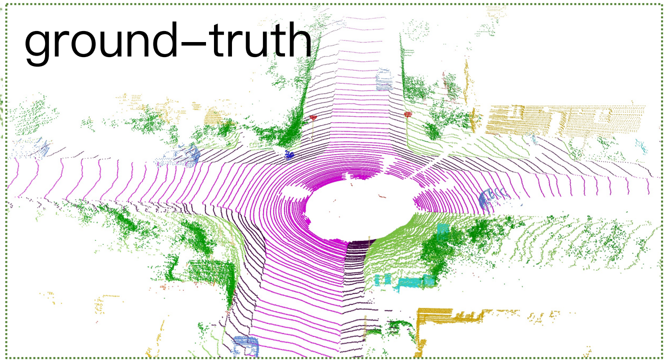
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

20%



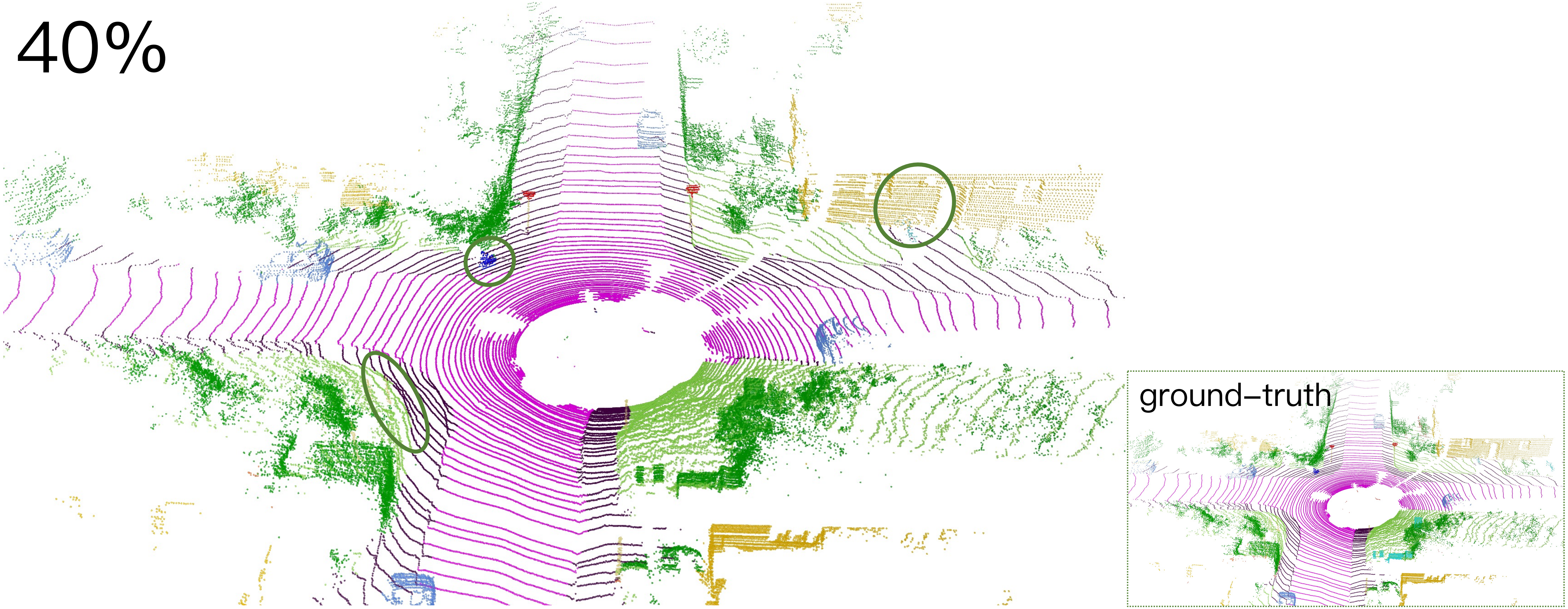


Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results

Comparing {5%, 10%, 20%, 40%} labeled splits

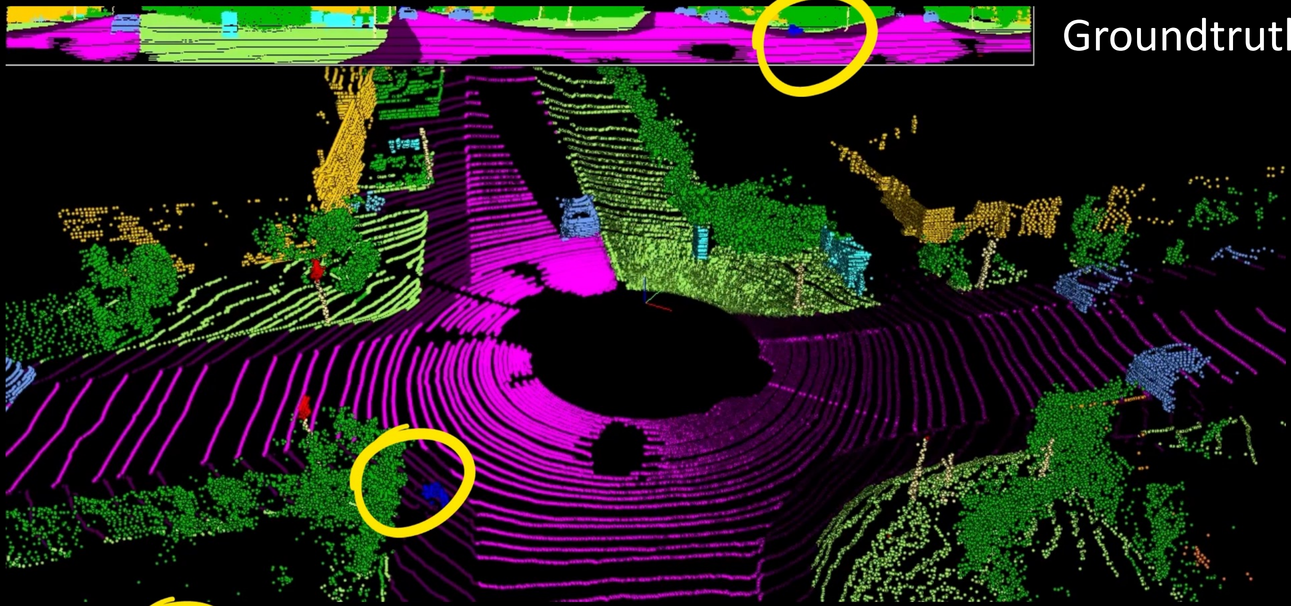
40%



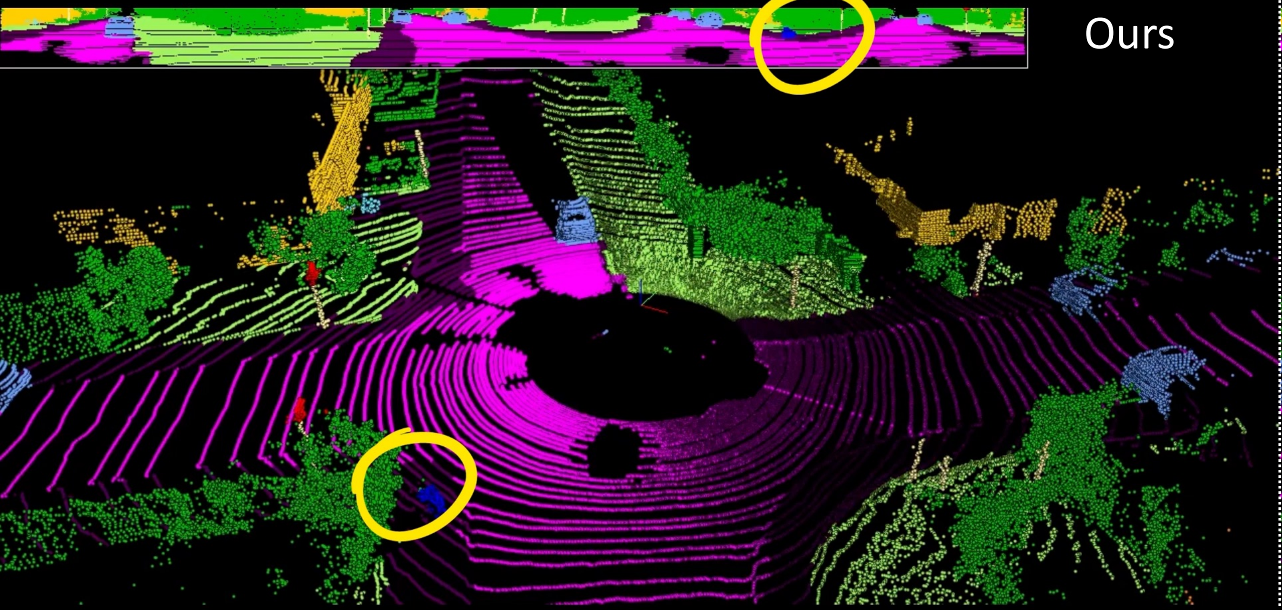
Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Qualitative results

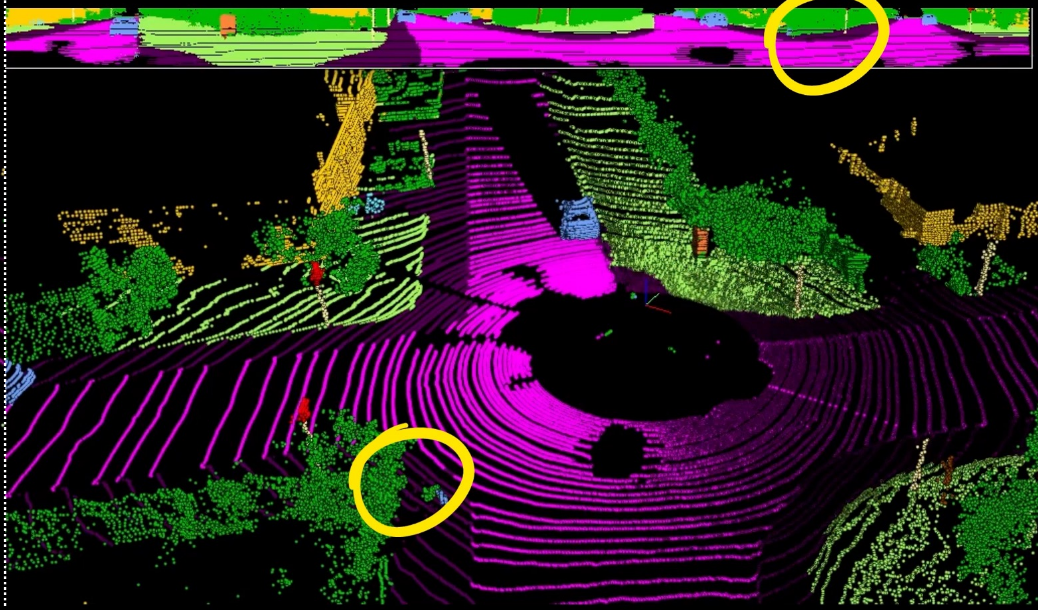
5%–Labeled Frames



Groundtruth



Ours



(Ozan et al)

Comparative mIoU for Semi-supervised Methods

Repr.	Samp.	Method		SemanticKITTI [7] 1%	5%	10%	20%	40%	50%	100%	ScribbleKITTI [46] 1%	5%	10%	20%	40%	50%	100%
Range	U	LaserMix [32]	(2022)	43.4	–	58.8	59.4	–	61.4	–	38.3	–	54.4	55.6	–	58.7	–
Voxel	U	Cylinder3D [63]	(CVPR'21)	–	45.4	56.1	57.8	58.7	–	67.8	–	39.2	48.0	52.1	53.8	–	56.3
	U	LaserMix [32]	(2022)	50.6	–	60.0	61.9	–	62.3	–	44.2	–	53.7	55.1	–	56.8	–
	P	Jiang *et al.* [29]	(ICCV'21)	–	41.8	49.9	58.8	59.9	–	65.8	–	–	–	–	–	–	–
	U	Unal *et al.* [46]	(CVPR'22)	–	49.9*	58.7*	59.1*	60.9	–	68.2*	–	46.9*	54.2*	56.5*	58.6*	–	61.3
	S	LiM3D+SDSC	(ours)	57.2	57.6	61.0	61.7	62.1	62.7	67.5	55.8	56.1	56.9	57.2	58.9	59.3	60.7
	S	LiM3D	(ours)	**58.4**	**59.5**	**62.2**	**63.1**	**63.3**	**63.6**	**69.5**	**57.0**	**58.1**	**61.0**	**61.2**	**62.0**	**62.1**	**62.4**

Comparative mIoU for Semi-supervised Methods

Repr.	Samp.	Method		SemanticKITTI [7] 1%	5%	10%	20%	40%	50%	100%	ScribbleKITTI [46] 1%	5%	10%	20%	40%	50%	100%
Range	U	LaserMix [32]	(2022)	43.4	–	58.8	59.4	–	61.4	–	38.3	–	54.4	55.6	–	58.7	–
Voxel	U	Cylinder3D [63]	(CVPR'21)	–	45.4	56.1	57.8	58.7	–	67.8	–	39.2	48.0	52.1	53.8	–	56.3
	U	LaserMix [32]	(2022)	50.6	–	60.0	61.9	–	62.3	–	44.2	–	53.7	55.1	–	56.8	–
	P	Jiang *et al.* [29]	(ICCV'21)	–	41.8	49.9	58.8	59.9	–	65.8	–	–	–	–	–	–	–
	U	Unal *et al.* [46]	(CVPR'22)	–	49.9*	58.7*	59.1*	60.9	–	68.2*	–	46.9*	54.2*	56.5*	58.6*	–	61.3
	S	LiM3D+SDSC	(ours)	57.2	57.6	61.0	61.7	62.1	62.7	67.5	55.8	56.1	56.9	57.2	58.9	59.3	60.7
	S	LiM3D	(ours)	**58.4**	**59.5**	**62.2**	**63.1**	**63.3**	**63.6**	**69.5**	**57.0**	**58.1**	**61.0**	**61.2**	**62.0**	**62.1**	**62.4**

Component-wise Ablation (Ours)

	UP	RF	RT	ST	SD	Training mIoU (%) 5%	10%	20%	40%	Validation mIoU (%) 5%	10%	20%	40%	#Params (M)
						82.8	87.5	87.8	88.2	54.8	58.1	59.3	60.8	49.6
	✓					–	–	–	–	55.9	58.8	59.9	61.2	49.6
	✓	✓				83.6	88.3	88.7	89.1	56.8	59.6	60.5	61.4	49.6
	✓		✓			–	–	–	–	57.5	59.8	61.2	62.6	49.6
	✓	✓	✓			–	–	–	–	58.7	61.3	62.4	62.8	49.6
LiM3D	✓	✓	✓	✓		**85.2**	**89.1**	**89.5**	**89.7**	**59.5**	**62.2**	**63.1**	**63.3**	49.6
LiM3D+SDSC	✓	✓	✓	✓	✓	83.8	88.6	89.0	89.2	57.6	61.0	61.7	62.1	**21.5**

UP Unreliable Pseudo labeling

RF Reflectivity Feature

RT Reflec-TTA

ST ST-RFD

SD SDSC module

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Component-wise Ablation (Ours)

	UP	RF	RT	ST	SD	Training mIoU (%) 5%	10%	20%	40%	Validation mIoU (%) 5%	10%	20%	40%	#Params (M)
						82.8	87.5	87.8	88.2	54.8	58.1	59.3	60.8	49.6
	✓					–	–	–	–	55.9	58.8	59.9	61.2	49.6
	✓	✓				83.6	88.3	88.7	89.1	56.8	59.6	60.5	61.4	49.6
	✓		✓			–	–	–	–	57.5	59.8	61.2	62.6	49.6
	✓	✓	✓			–	–	–	–	58.7	61.3	62.4	62.8	49.6
LiM3D	✓	✓	✓	✓		**85.2**	**89.1**	**89.5**	**89.7**	**59.5**	**62.2**	**63.1**	**63.3**	49.6
LiM3D+SDSC	✓	✓	✓	✓	✓	83.8	88.6	89.0	89.2	57.6	61.0	61.7	62.1	**21.5**

UP Unreliable Pseudo labeling

RF Reflectivity Feature

RT Reflec-TTA

ST ST-RFD

SD SDSC module

The Computation Cost and mIoU

Under 5%-labeled Training Results

Method	# Parameters	# Mult-Adds	SeK [7]	ScK [45]
Cylider3D [61]	56.3	476.9M	45.4	39.2
Ozan *et al.* [45]	49.6	420.2M	49.9	46.9
2DPASS [56]	26.5	217.4M	51.7	45.1
MinkowskiNet [13]	21.7	114.0G	42.4	35.8
SPVNAS [43]	**12.5**	73.8G	45.1	38.9
LiM3D+SDSC (ours)	21.5	**182.0M**	57.6	54.7
LiM3D (ours)	49.6	420.2M	**59.5**	**58.1**

The Computation Cost and mIoU

Under 5%-labeled Training Results

Method	# Parameters	# Mult-Adds	SeK [7]	ScK [45]
Cylider3D [61]	56.3	476.9M	45.4	39.2
Ozan *et al.* [45]	49.6	420.2M	49.9	46.9
2DPASS [56]	26.5	217.4M	51.7	45.1
MinkowskiNet [13]	21.7	114.0G	42.4	35.8
SPVNAS [43]	**12.5**	73.8G	45.1	38.9
LiM3D+SDSC (ours)	21.5	**182.0M**	57.6	54.7
LiM3D (ours)	49.6	420.2M	**59.5**	**58.1**

2.3x model size reduction

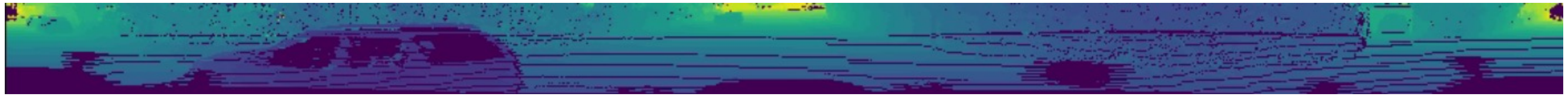
641x fewer multiply-adds

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Effects of ST–RFD Sampling

Sampling	SemanticKITTI [7] 5%	SemanticKITTI [7] 10%	SemanticKITTI [7] 20%	SemanticKITTI [7] 40%	ScribbleKITTI [45] 5%	ScribbleKITTI [45] 10%	ScribbleKITTI [45] 20%	ScribbleKITTI [45] 40%
Random	58.5	61.6	62.6	62.7	57.1	60.3	60.5	60.9
Uniform	58.7	61.3	62.4	62.8	56.9	60.6	60.3	61.0
ours ST-RFD-R	59.1	**62.4**	62.9	**63.4**	58.0	60.7	**61.2**	61.8
ours ST-RFD	**59.5**	62.2	**63.1**	63.3	**58.1**	**61.0**	**61.2**	**62.0**

ST–RFD–R
Range Image:

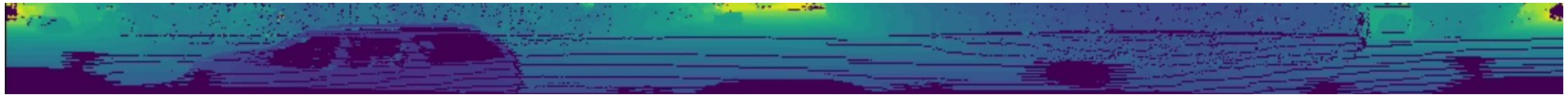


Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Effects of ST-RFD Sampling

	Sampling	SemanticKITTI [7]				ScribbleKITTI [45]			
		5%	10%	20%	40%	5%	10%	20%	40%
	Random	58.5	61.6	62.6	62.7	57.1	60.3	60.5	60.9
	Uniform	58.7	61.3	62.4	62.8	56.9	60.6	60.3	61.0
ours	ST-RFD-R	59.1	**62.4**	62.9	**63.4**	58.0	60.7	**61.2**	61.8
ours	ST-RFD	**59.5**	62.2	**63.1**	63.3	**58.1**	**61.0**	**61.2**	**62.0**

ST–RFD–R
Range Image:



Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Effects of Differing Reliability Using Pseudo Voxels

Ratio	Unreliable		Reliable		Random	
	mIoU	SS/FF	mIoU	SS/FF	mIoU	SS/FF
5%	**59.5**	**85.6**	57.2	82.3	56.4	81.2
10%	**62.2**	**89.5**	60.8	87.5	59.7	85.9
20%	**63.1**	**90.8**	61.4	88.3	60.5	87.1
40%	**63.3**	**91.1**	62.8	90.4	61.3	88.2

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Effects of Differing Reliability Using Pseudo Voxels

Ratio	Unreliable		Reliable		Random	
	mIoU	SS/FF	mIoU	SS/FF	mIoU	SS/FF
5%	**59.5**	**85.6**	57.2	82.3	56.4	81.2
10%	**62.2**	**89.5**	60.8	87.5	59.7	85.9
20%	**63.1**	**90.8**	61.4	88.3	60.5	87.1
40%	**63.3**	**91.1**	62.8	90.4	61.3	88.2

Less is More: Reducing Task and Model Complexity for
3D Point Cloud Semantic Segmentation, CVPR 2023, Li, Shum, Breckon

Reflectivity (Reflec–TTA) vs. Intensity (Intensity–based TTA)

TTA	SemanticKITTI [7]				ScribbleKITTI [45]			
	5%	10%	20%	40%	5%	10%	20%	40%
Intensity	56.2	59.1	59.8	60.9	55.7	57.5	57.9	59.2
Reflectivity	**59.5**	**62.2**	**63.1**	**63.3**	**58.1**	**61.0**	**61.2**	**62.0**

Reflectivity (Reflec–TTA) vs. Intensity (Intensity–based TTA)

TTA	SemanticKITTI [7]				ScribbleKITTI [45]			
	5%	10%	20%	40%	5%	10%	20%	40%
Intensity	56.2	59.1	59.8	60.9	55.7	57.5	57.9	59.2
Reflectivity	**59.5**	**62.2**	**63.1**	**63.3**	**58.1**	**61.0**	**61.2**	**62.0**

Durham University

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Less is More

Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation

homepage

paper

code

{li.li4, hubert.shum, toby.breckon}@durham.ac.uk