Project page : https://dslrproject.github.io/dslr/

DSLR : Dynamic to Static LiDAR Scan Reconstruction Using Adversarially Trained Autoencoder

P. Kumar*[1] S. Sahoo*[1] V. Shah[1] V. Kondameedi[1] A. Jain[1] A. Verma[1]
C. Bhattacharyya[1] V Vinay[23]

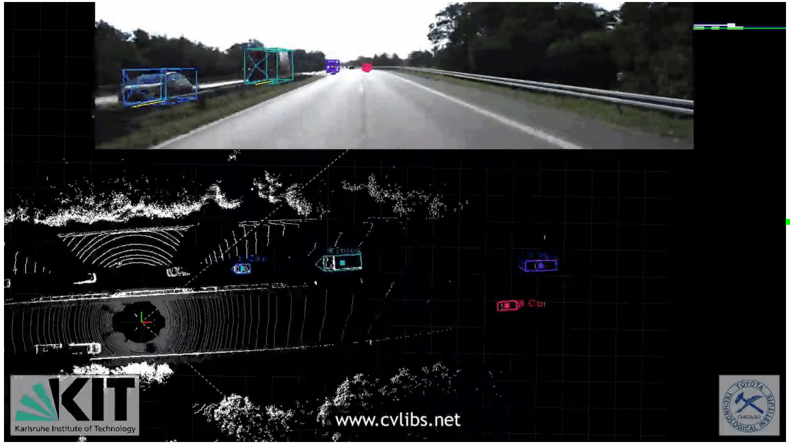
[1]Indian Institute of Science, Bangalore, India

[2]AMIDC Pvt Ltd, Bangalore, India

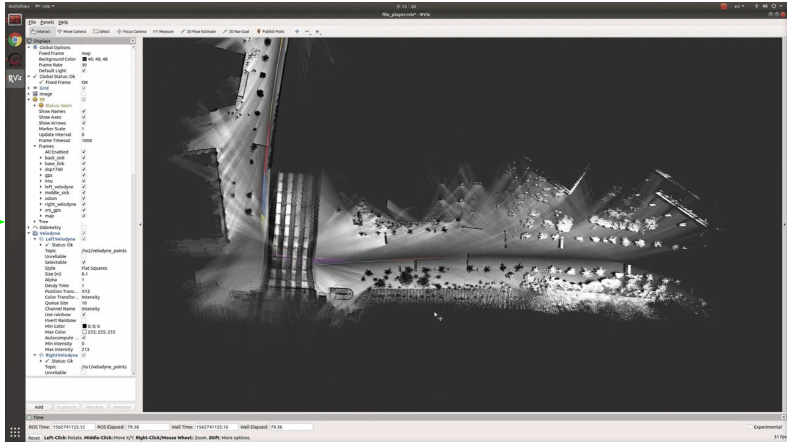
[3]Chennai Mathematical Institute, Chennai, India

Autonomous Navigation using LiDAR based SLAM



SLAM

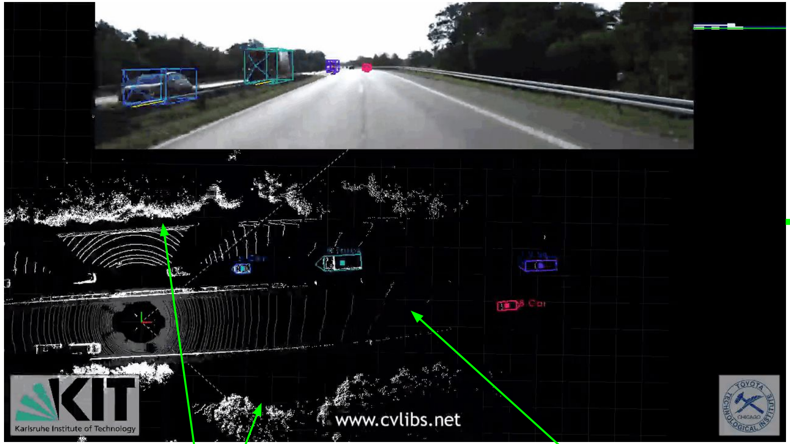


LiDAR scans captured on a highway

Map of city generated using the LiDAR scans

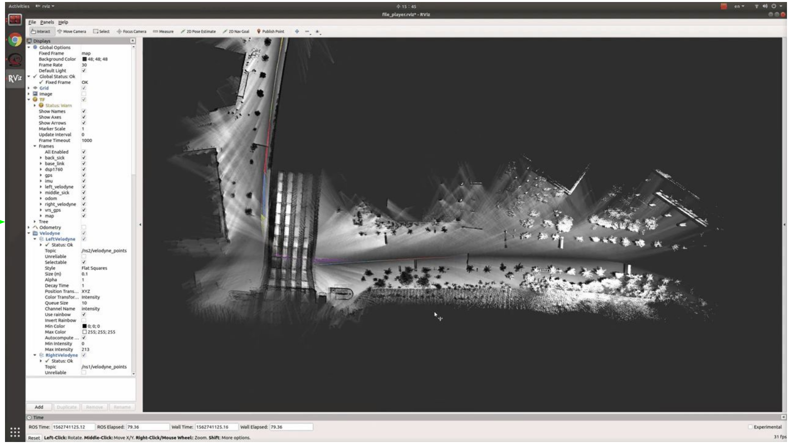
[1] *Geiger et al. 2012. "Are we ready for autonomous driving? the kitti vision benchmark suite.". CVPR 2012.*
[2] *Jeong et al. 2019. "Complex urban dataset with multi-level sensors from highly diverse urban environments". IJRR 2019.*

Autonomous Navigation using LiDAR based SLAM



LiDAR scans captured on a highway

SLAM



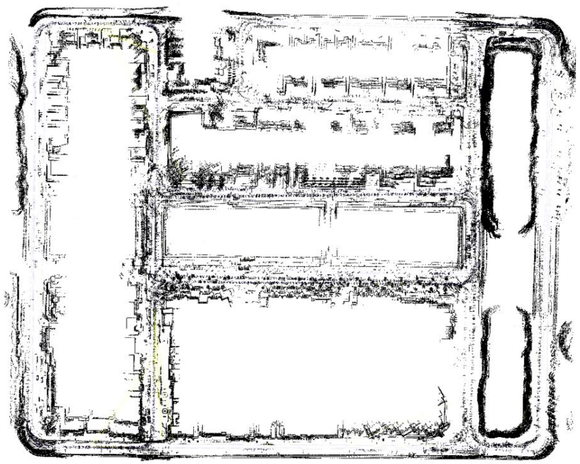
Map of city generated using the LiDAR scans

Static structures like walls or trees

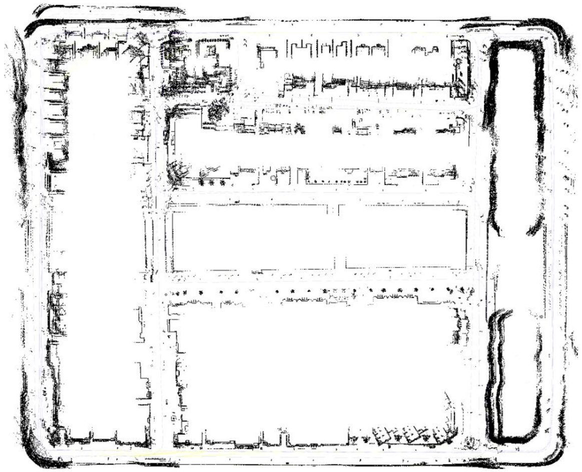
Dynamic objects like vehicles or people

[1] *Geiger et al. 2012. "Are we ready for autonomous driving? the kitti vision benchmark suite.". CVPR 2012.*
[2] *Jeong et al. 2019. "Complex urban dataset with multi-level sensors from highly diverse urban environments". IJRR 2019.*

Dynamic objects deteriorate Autonomous Navigation



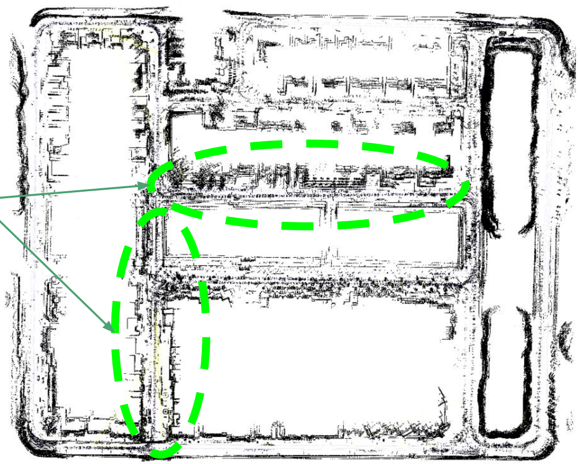
Map generated from Dynamic Scans



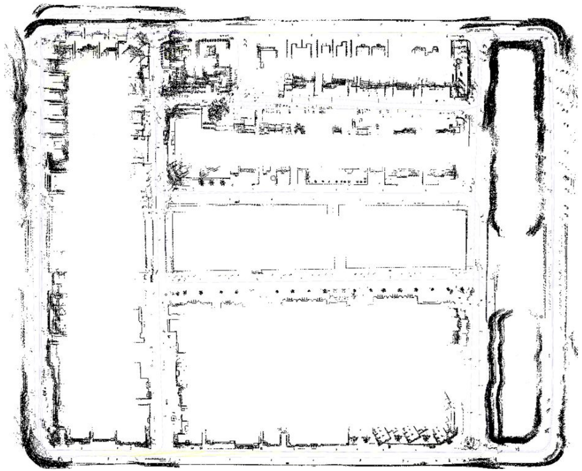
Map generated from Static Scans

Dynamic objects deteriorate Autonomous Navigation

Corruptions
in the map



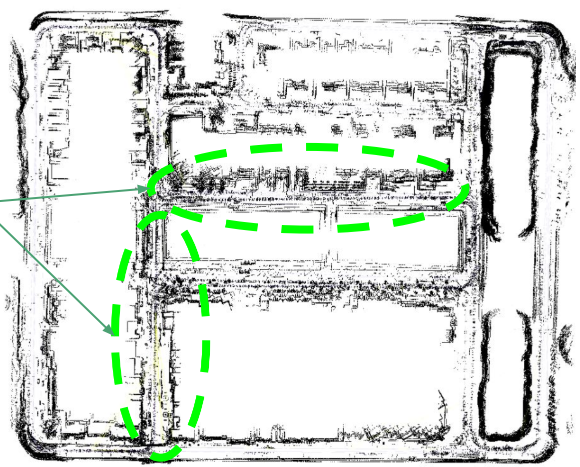
Map generated from Dynamic Scans



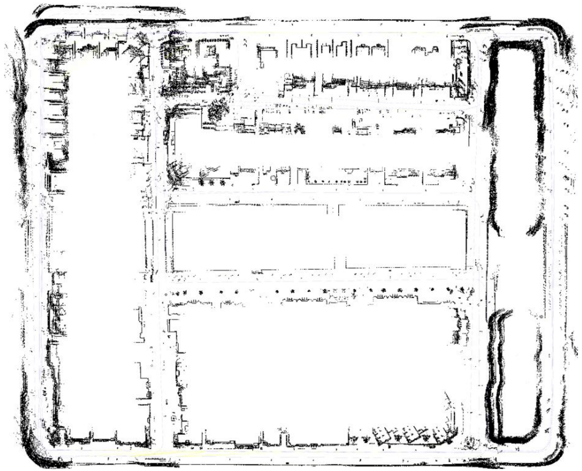
Map generated from Static Scans

Dynamic objects deteriorate Autonomous Navigation

Corruptions in the map



Map generated from Dynamic Scans

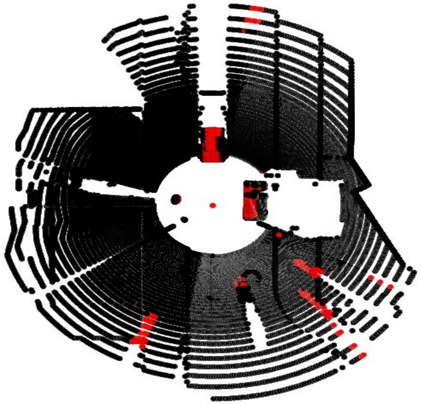


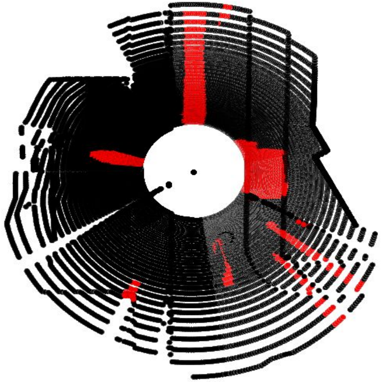
Map generated from Static Scans

Performance in **Dynamic Env** << **Static Env**

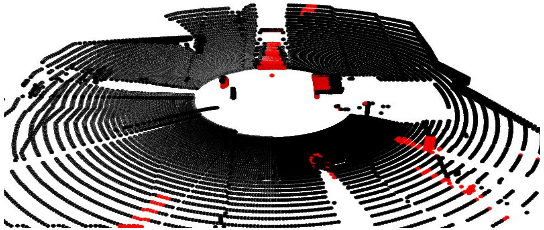
Dynamic and Static LiDAR scans

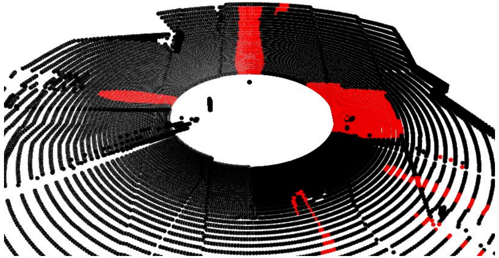
Top View





Angled View





Dynamic LiDAR Scan

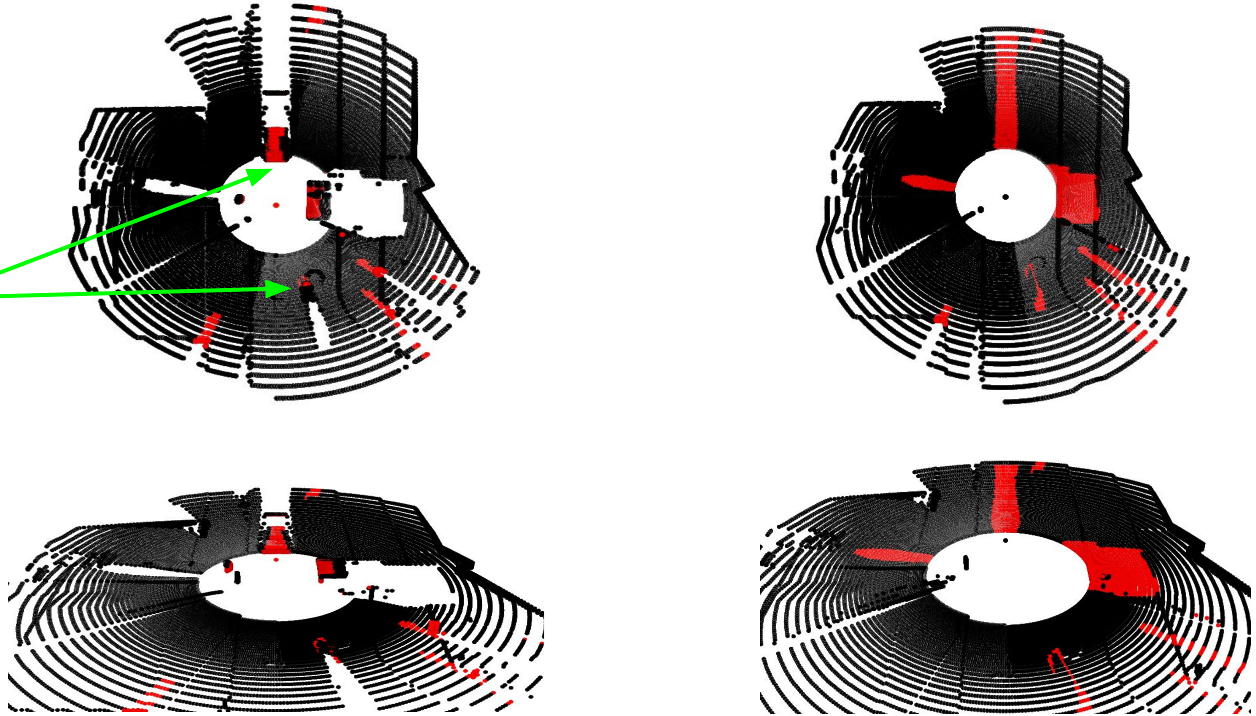
Corresponding Static Lidar Scan

Dynamic and Static LiDAR scans

Top View

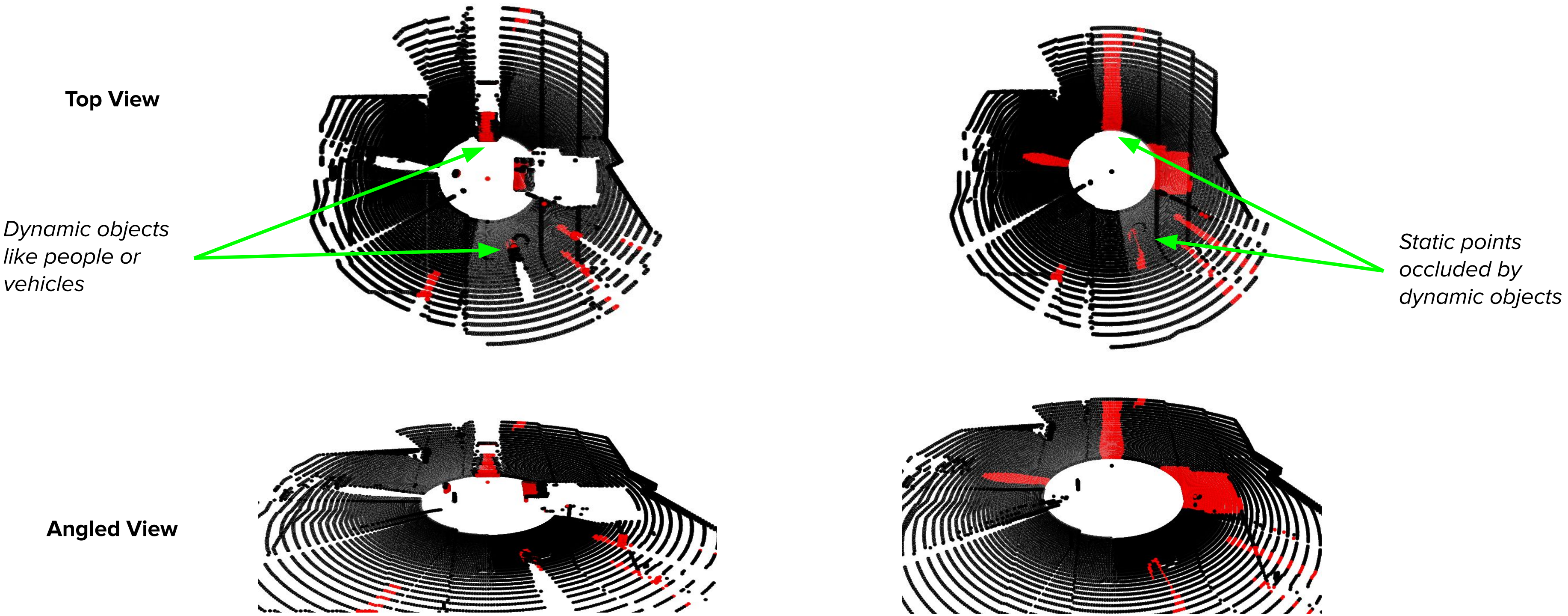
Dynamic objects like people or vehicles

Angled View



Dynamic LiDAR Scan | *Corresponding Static Lidar Scan*

Dynamic and Static LiDAR scans



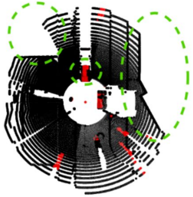
Dynamic LiDAR Scan

Corresponding Static Lidar Scan

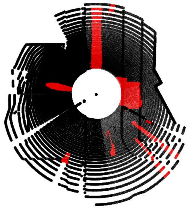
Problem Statement : Dynamic to Static Translation (DST) for LiDAR

Learn mapping, M from dynamic scan D to corresponding static scan S.

min(M(



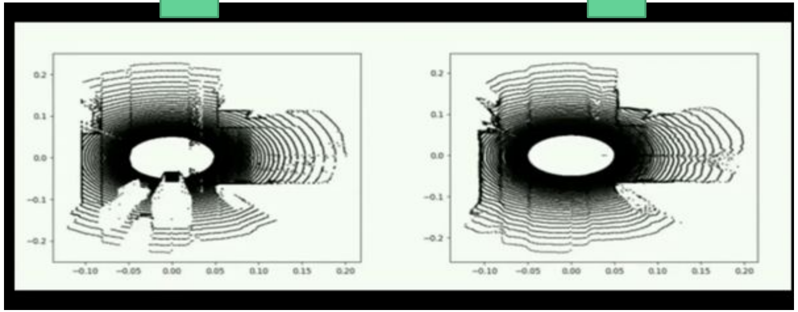
)) -



)^2

Dynamic Scan, D

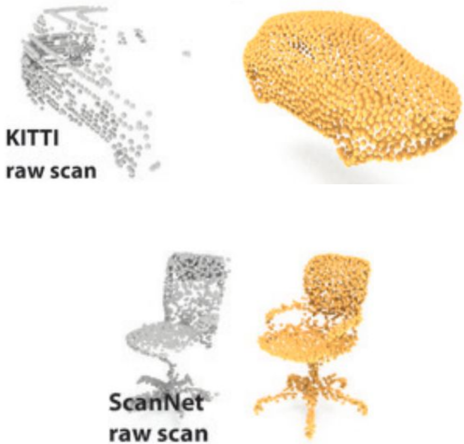
Corresponding Static Scan, S



Existing Work: DST exists for images / object point clouds

Dynamic Input

Static Output



[3] *Bescos et al. 2019. "Empty Cities: Image Inpainting for a Dynamic-Object-Invariant Space". ICRA 2019.*
[4] *Chen et al. 2020. "Unpaired Point Cloud Completion on real scans using adversarial training". ICLR 2020.*
[5] *Groueix et al. 2018. "AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation". CVPR 2018.*
[6] *Wu et al. 2020 "Multimodal Shape Completion via Conditional Generative Adversarial Networks". ECCV 2020.*
[7] *Achlioptas et al. "Learning Representations and Generative Models for 3D Point Clouds". ICML 2018.*

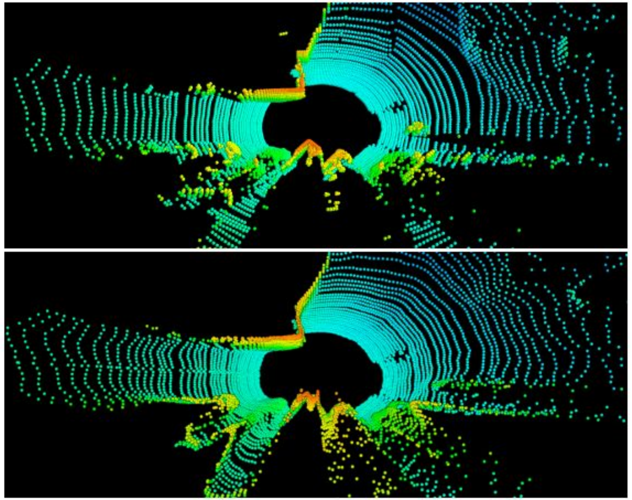
Existing Works- Closest to our Problem

This work uses Generative Modelling to reconstruct and generate new LiDAR frame.

Static Reconstructions using this work works poorly for SLAM

Original LiDAR

Reconstructed LiDAR



[8] *Caccia et al. 2019. “Deep Generative Modeling of LiDAR Data”. IROS 2019.*

Open Problems in DST for LiDAR

- Existing DST works for images require segmentation information
- Point cloud scan completion based methods do not work for 360° LiDAR
- Existing LiDAR reconstruction methods fail to produce SLAM-worthy reconstructions

Contributions

- **DST for LiDAR**
  - DSLR (Dynamic to Static LiDAR scan reconstruction)
  - DSLR-Seg (with segmentation information, if available)
  - DSLR-UDA (datasets w/o paired static LiDAR scans)
  - Lidar scan Quality Index (LiDAR scan evaluation metric in absence of GT)

- **Dataset Generation** (generate dynamic and corresponding static pairs)

- **DST for LIDAR based SLAM**
  - SLAM Reconstruction Threshold
  - Practically feasible

Outline

DSLR
(Dynamic to Static LiDAR scan Reconstruction)

- DST for LiDAR
  - **DSLR (w/o seg info)**
  - DSLR-UDA (w/o paired static)
  - Lidar scan Quality Index
  - DSLR-Seg (with seg info)
  - Results
- Dataset Generation
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - Results

DSLR - Overview

LiDAR Frame Autoencoder: A **DCGAN based autoencoder** that maps LiDAR scans to their latent representations and reconstructs them back.

DSLR - Overview

LiDAR Frame Autoencoder: A **DCGAN based autoencoder** that maps LiDAR scans to their latent representations and reconstructs them back.

Pair Discriminator: Multilayered neural network that **accepts latent representation vector pairs and discriminates** $(static, static) v/s (static, dynamic)$ latent vector pairs.

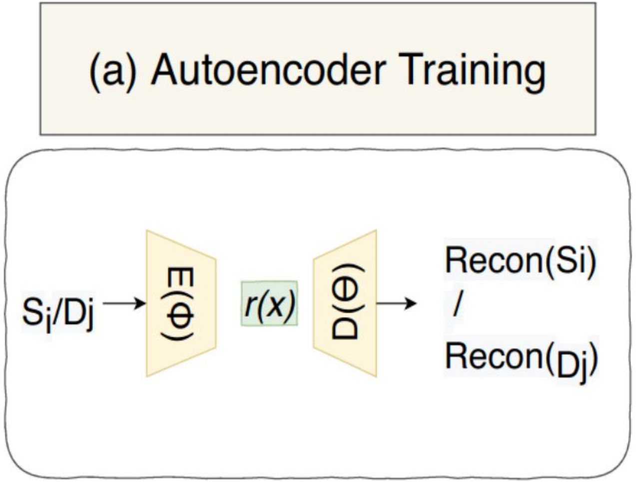
DSLR - Overview

LiDAR Frame Autoencoder: A **DCGAN based autoencoder** that maps LiDAR scans to their latent representations and reconstructs them back.

Pair Discriminator: Multilayered neural network that **accepts latent representation vector pairs and discriminates** $(static, static) v/s (static, dynamic)$ latent vector pairs.

Adversarial Model: **Combines the above 2 networks using an adversarial strategy** to transform dynamic LiDAR frame based latent vector representation to their corresponding static latent vector.

DSLR - LiDAR Frame Autoencoder



$$G : \mathbf{x} \xrightarrow{E_\phi} r(x) \xrightarrow{D_\theta} \bar{\mathbf{x}}$$

$$\mathrm{MSE}(\mathbf{x}, \bar{\mathbf{x}}) = \|\mathbf{x} - \bar{\mathbf{x}}\|^2$$

DSLR - Pair Discriminator (DI)

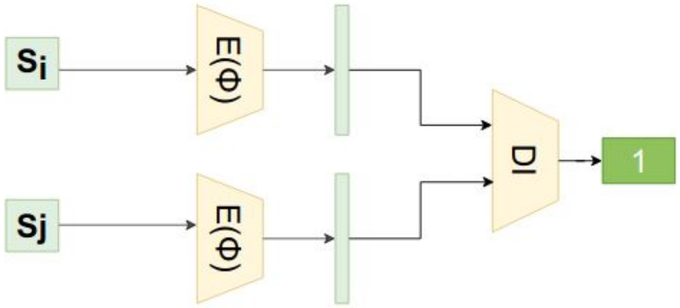
Discriminate (static, static) pairs from (static, dynamic) pairs.

$$DI\left(r(\mathbf{x}_1), (r(\mathbf{x}_2)\right) = \left\{ \begin{array}{ll} 1 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in S \\ 0 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in D \end{array} \right\}$$

DSLR - Pair Discriminator (DI)

Discriminate (static, static) pairs from (static, dynamic) pairs.

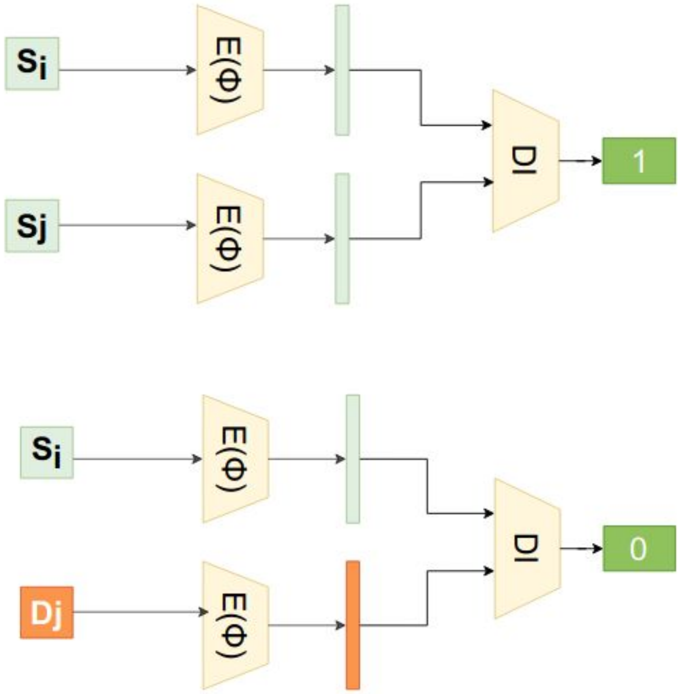
$$DI\left(r(\mathbf{x}_1), (r(\mathbf{x}_2)\right) = \left\{ \begin{array}{ll} 1 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in S \\ 0 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in D \end{array} \right\}$$



DSLR - Pair Discriminator (DI)

Discriminate (static, static) pairs from (static, dynamic) pairs.

$$DI\left(r(\mathbf{x}_1), (r(\mathbf{x}_2)\right) = \left\{ \begin{array}{ll} 1 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in S \\ 0 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in D \end{array} \right\}$$



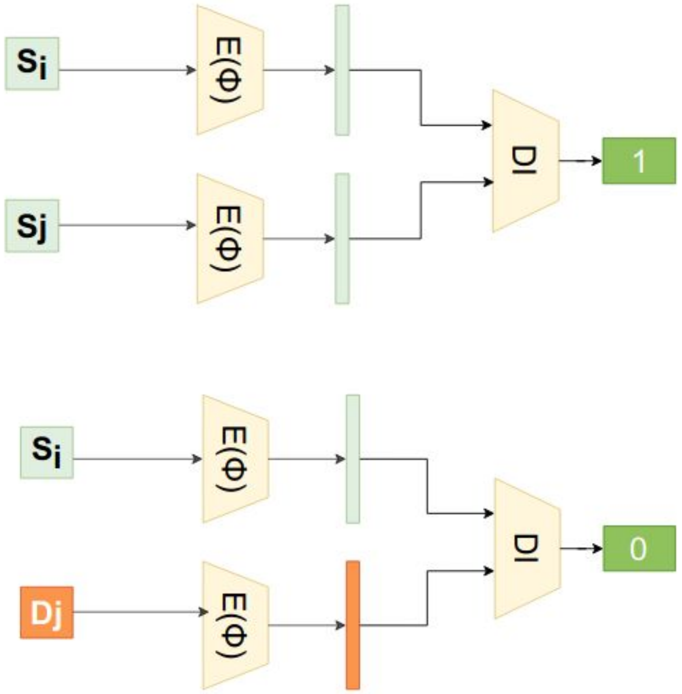
DSLR - Pair Discriminator (DI)

Discriminate (static, static) pairs from (static, dynamic) pairs.

$$DI\left(r(\mathbf{x}_1), (r(\mathbf{x}_2)\right) = \left\{ \begin{array}{ll} 1 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in S \\ 0 & \mathbf{x}_1 \in S, \mathbf{x}_2 \in D \end{array} \right\}$$

Challenging to train DI in the above setting.

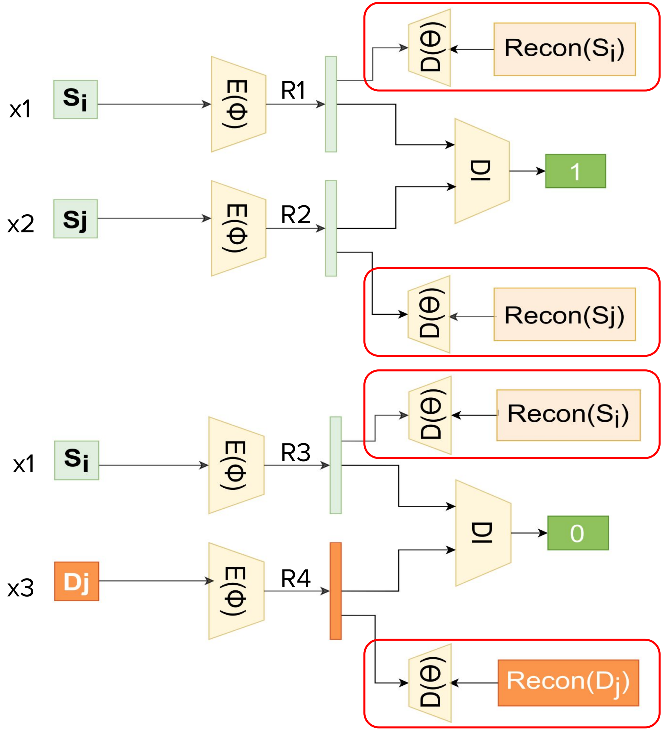
Reason: Latent input vectors that are input to DI had learnt only generative features.



DSLR - Pair Discriminator (DI)

Training Discriminator using a DUAL LOSS.

$$L_{DI} = \boxed{MSE(\mathbf{x}_1, \overline{\mathbf{x}_1}) + MSE(\mathbf{x}_2, \overline{\mathbf{x}_2}) + MSE(\mathbf{x}_3, \overline{\mathbf{x}_3})}$$
$$+ BCE(DI(r(\mathbf{x}_1), r(\mathbf{x}_2)), 1)$$
$$+ BCE(DI(r(\mathbf{x}_1), r(\mathbf{x}_3)), 0)$$

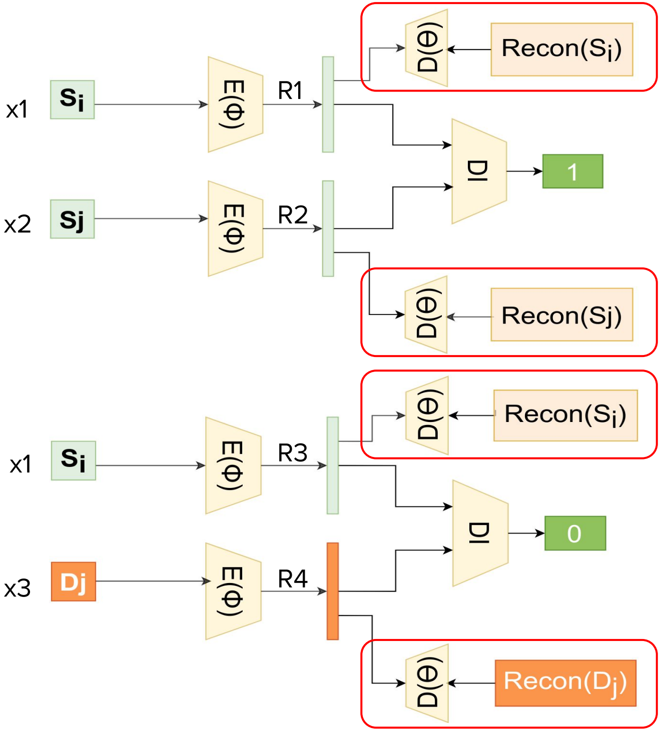


DSLR - Pair Discriminator (DI)

Training Discriminator using a DUAL LOSS.

$$L_{DI} = \boxed{MSE(\mathbf{x}_1, \overline{\mathbf{x}_1}) + MSE(\mathbf{x}_2, \overline{\mathbf{x}_2}) + MSE(\mathbf{x}_3, \overline{\mathbf{x}_3})} \\ + BCE(DI(r(\mathbf{x}_1), r(\mathbf{x}_2)), 1) \\ + BCE(DI(r(\mathbf{x}_1), r(\mathbf{x}_3)), 0)$$

Latent vector pair input to DI contain generative and discriminative features.

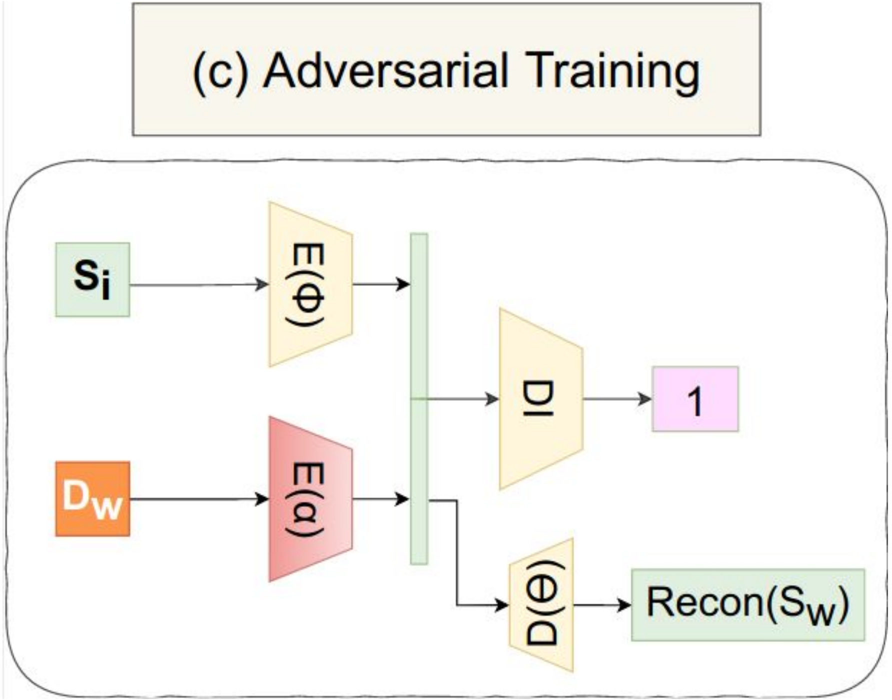


DSLR - Adversarial Loss Training for DST

Adversarial strategy.

FORCE autoencoder to output static latent vector for a given dynamic frame.

FOOLING discriminator with adversarial labels.



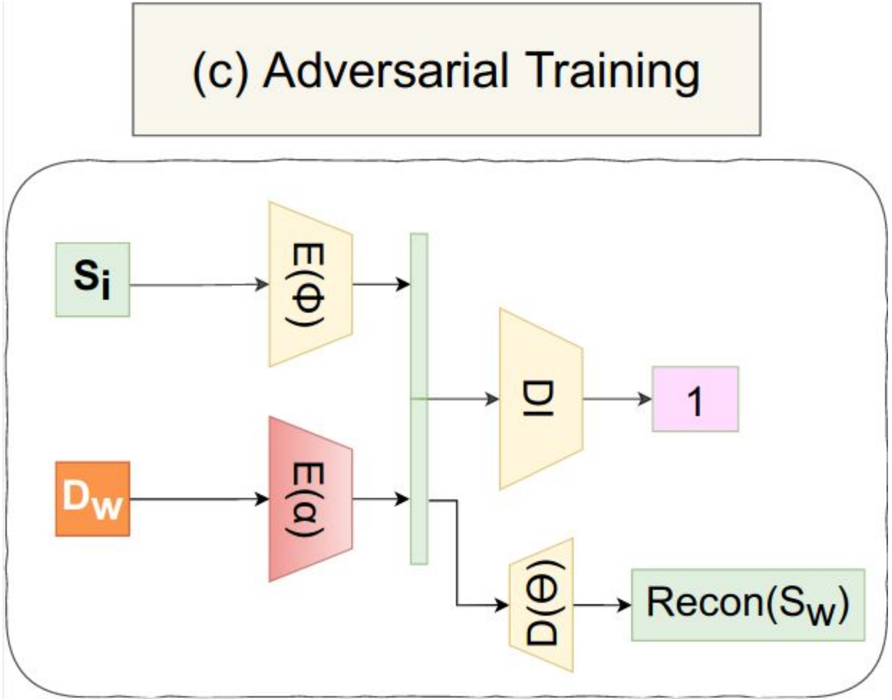
DSLR - Adversarial Loss Training for DST

Adversarial strategy

FORCE autoencoder to output static latent vector for a given dynamic frame.

FOOLING discriminator with adversarial labels.

Working on the latent space of LiDAR scan unlike existing baselines.



Qualitative LiDAR Reconstruction Results

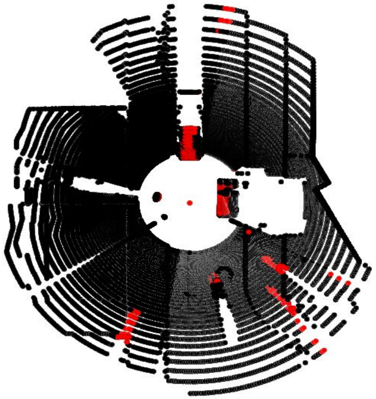


Fig. A

*Fig A**:** CARLA dynamic LiDAR scan* ***(red indicates dynamic objects)***

Qualitative LiDAR Reconstruction Results

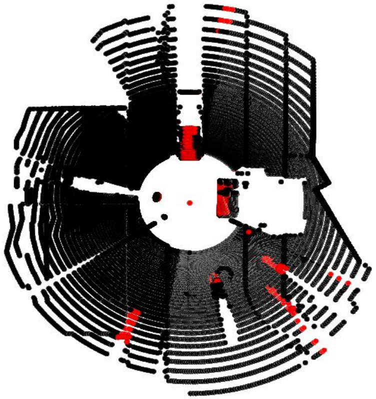


Fig. A

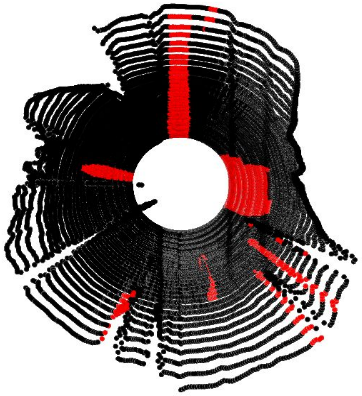


Fig. B

*Fig A**:** CARLA dynamic LiDAR scan* ***(red indicates dynamic objects)***

*Fig B**:** Model reconstructed static LiDAR scan* ***(red indicates reconstructed static background)***

Qualitative LiDAR Reconstruction Results

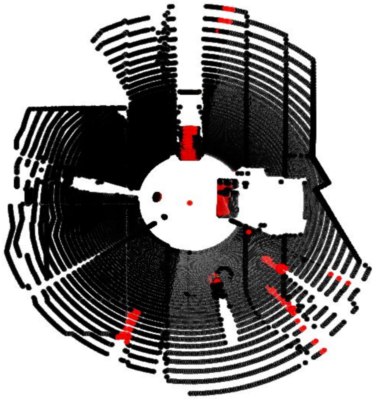


Fig. A

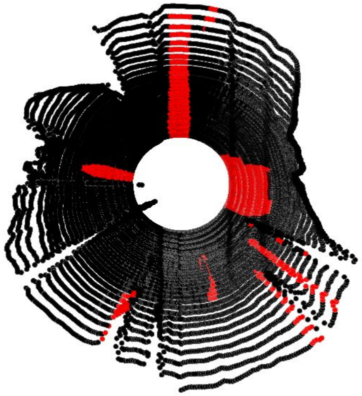


Fig. B

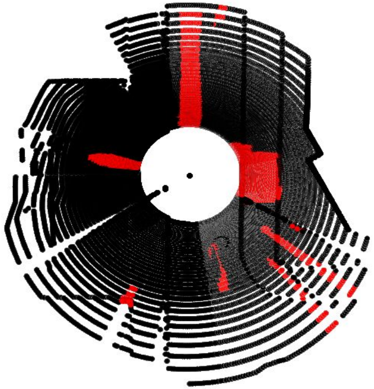


Fig. C

*Fig A: CARLA dynamic LiDAR scan **(red indicates dynamic objects)***

*Fig B: Model reconstructed static LiDAR scan **(red indicates reconstructed static background)***

*Fig C: Ground Truth static **(red indicates ground truth static points corresponding to dynamic regions)***

Qualitative LiDAR Reconstruction Results

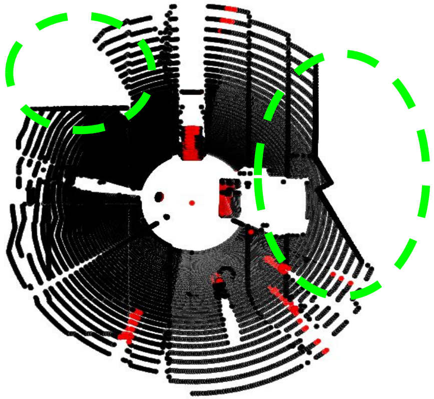


Fig. A

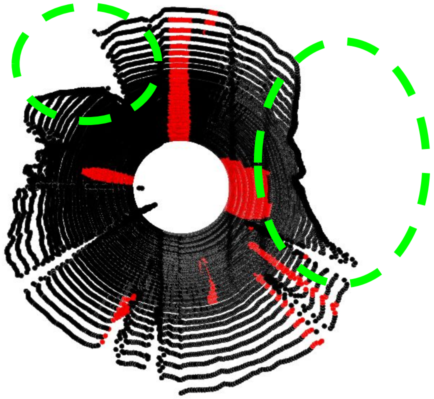


Fig. B

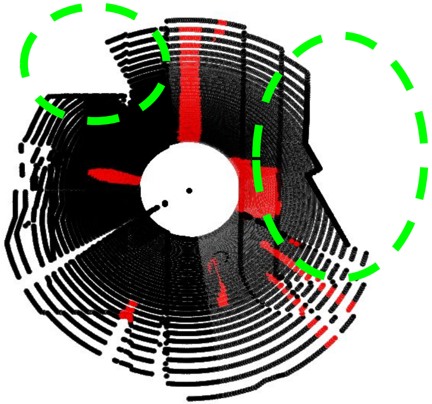


Fig. C

Fig A: CARLA dynamic LiDAR scan ***(red indicates dynamic objects)***

Fig B: Model reconstructed static LiDAR scan ***(red indicates reconstructed static background)***

Fig C: Ground Truth static ***(red indicates ground truth static points corresponding to dynamic regions)***

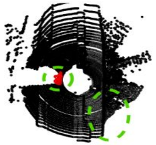
Outline

DSLR-UDA
(Unsupervised Domain Adaptation)

- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - **DSLR-UDA (w/o paired static)**
  - Lidar scan Quality Index
  - DSLR-Seg (with seg info)
  - Results
- Dataset Generation
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - Results

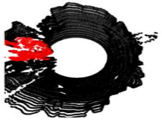
DSLR-UDA (Unsupervised Domain Adaptation)

- “Corresponding dynamic-static pairs” might not always be available in real world scenarios



Real World Dynamic Scan

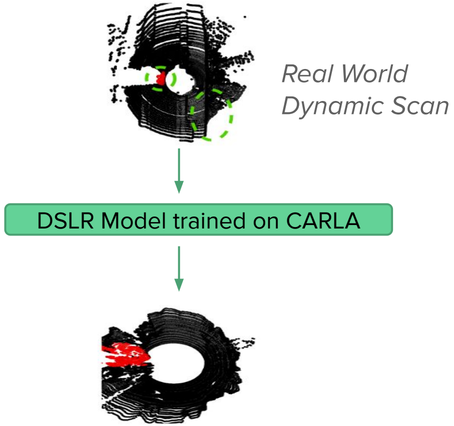
DSLR Model trained on CARLA



Static Reconstruction are quite ***poor*** *for such real world datasets*

DSLR-UDA (Unsupervised Domain Adaptation)

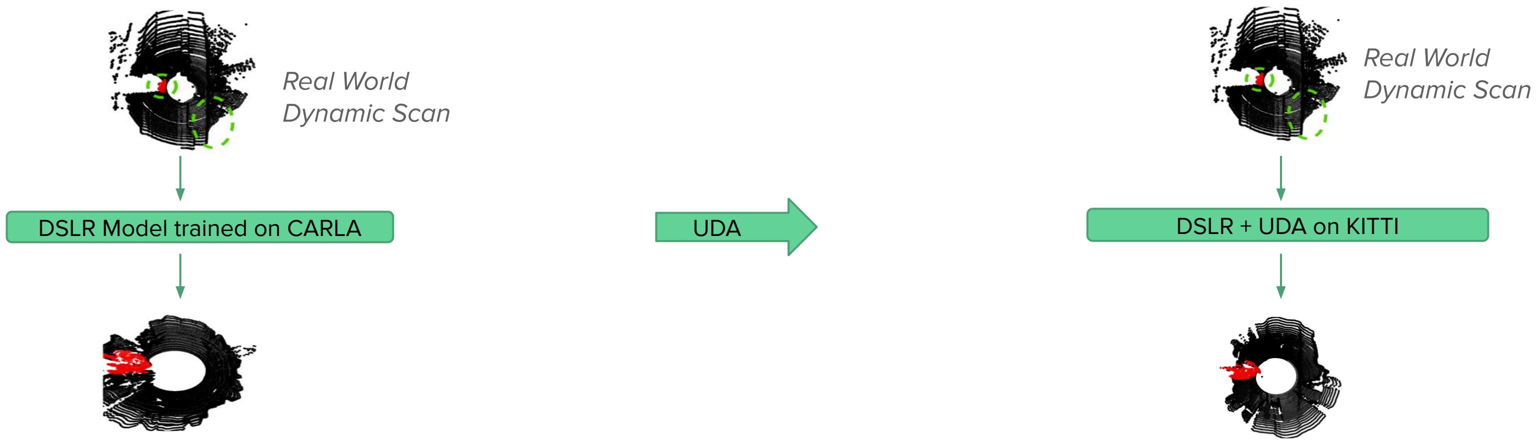
- “Corresponding dynamic-static pairs” might not always be available in real world scenarios
- Unsupervised domain adaptation between the simulated CARLA scans and the real world dynamic scans, help adapt DSLR to such environments



*Static Reconstruction are quite **poor** for such real world datasets*

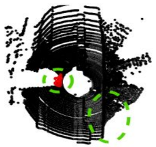
DSLR-UDA (Unsupervised Domain Adaptation)

- "Corresponding dynamic-static pairs" might not always be available in real world scenarios
- Unsupervised domain adaptation between the simulated CARLA scans and the real world dynamic scans, help adapt DSLR to such environments



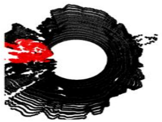
*Static Reconstruction are quite **poor** for such real world datasets*

*Static Reconstruction are **better** when trained using **DSLR + UDA***



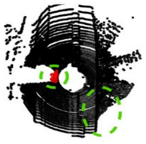
Real World Dynamic Scan

DSLR Model trained on CARLA



*Static Reconstruction are quite **poor** for such real world datasets*

UDA



Real World Dynamic Scan

DSLR + UDA on KITTI

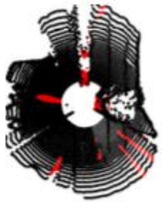
*Static Reconstruction are **better** when trained using **DSLR + UDA***

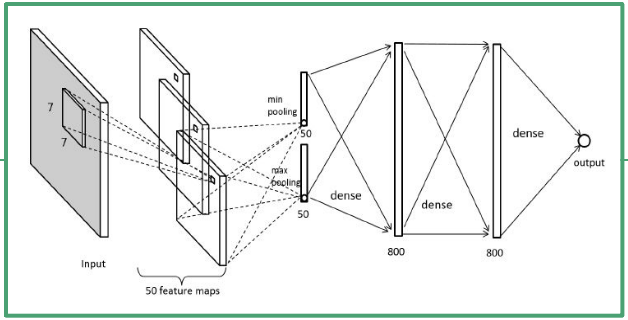
Outline

Lidar scan Quality Index or LQI
(LiDAR scan evaluation metric in absence of Ground Truth)

- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - DSLR-UDA (w/o paired static)
  - **Lidar scan Quality Index**
  - DSLR-Seg (with seg info)
  - Results
- Dataset Generation
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - Results

LiDAR Quality Index (LQI)

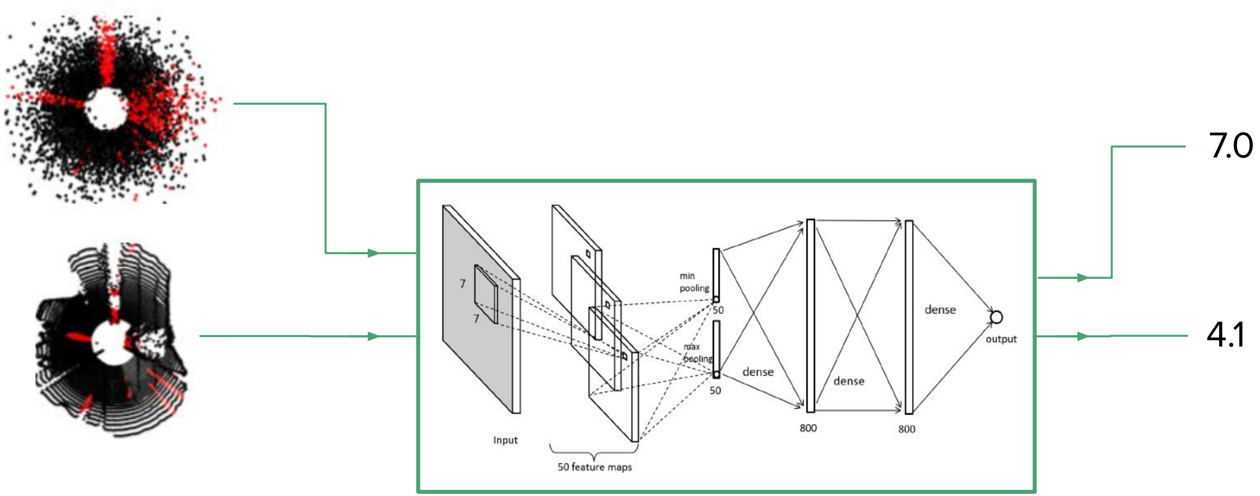




4.1

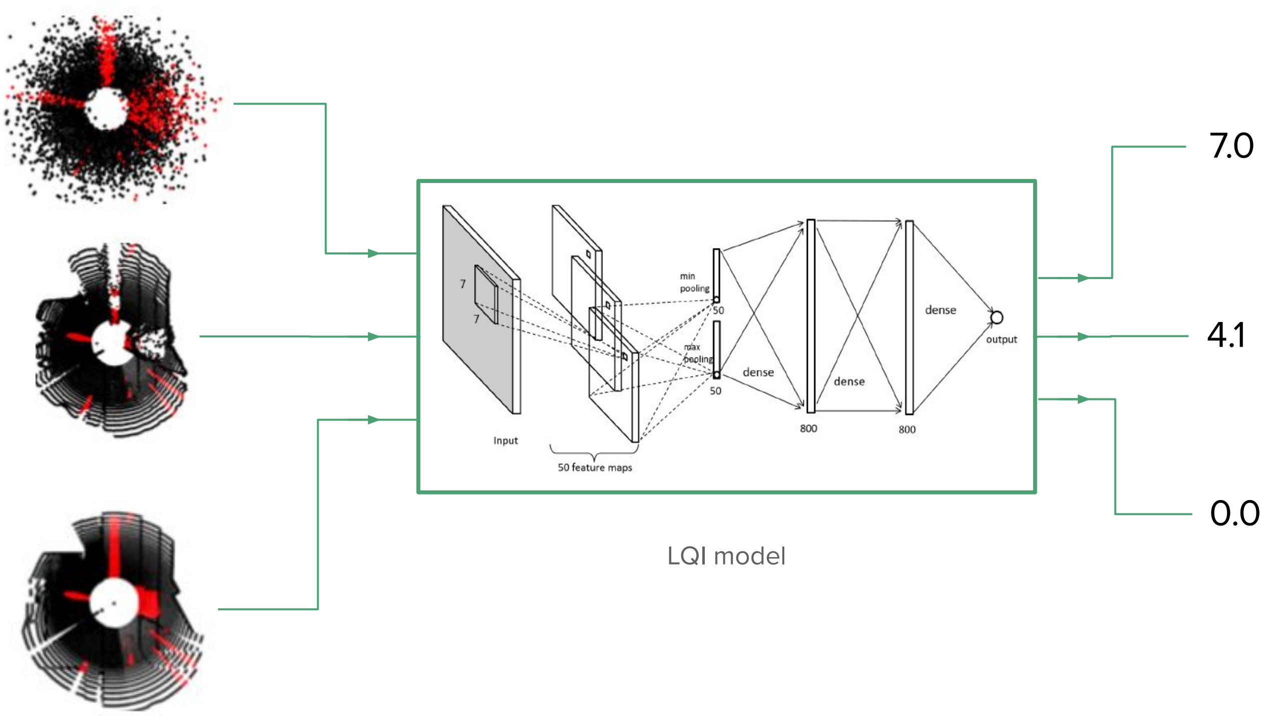
LQI model

LiDAR Quality Index (LQI)



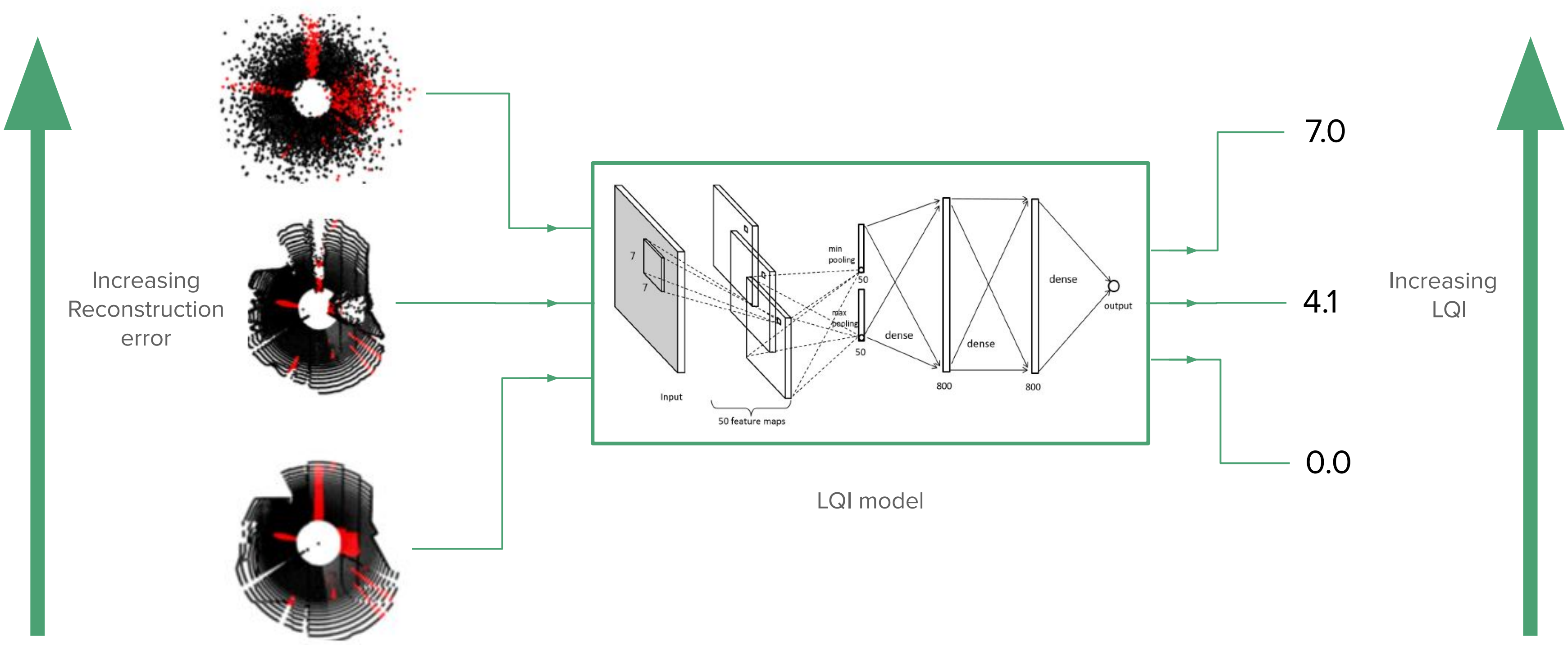
LQI model

LiDAR Quality Index (LQI)



LQI model

LiDAR Quality Index (LQI)

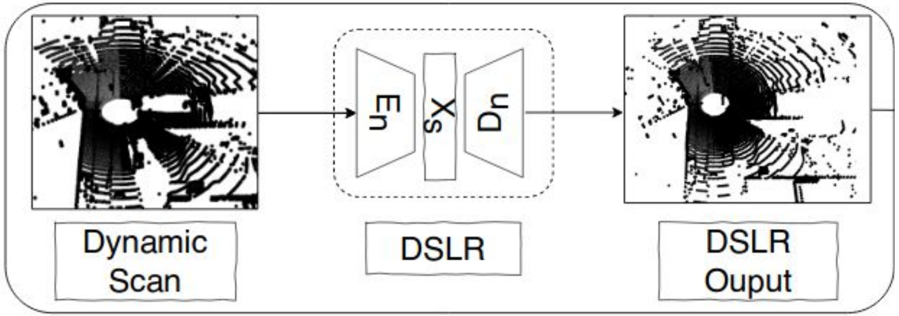


Outline

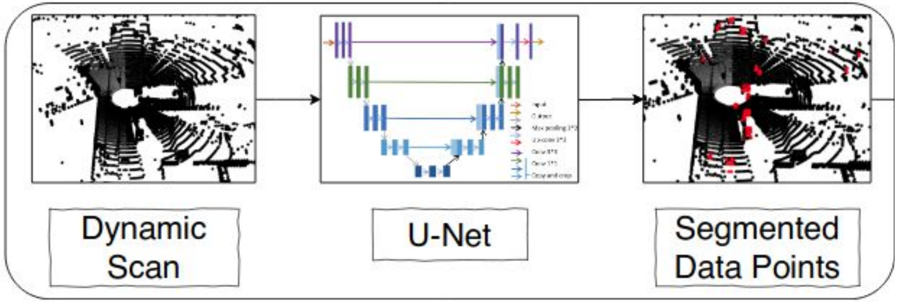
DSLR-Seg
(if segmentation information is available)

- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - DSLR-UDA (w/o paired static)
  - Lidar scan Quality Index
  - **DSLR-Seg (with seg info)**
  - Results
- Dataset Generation
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - Results

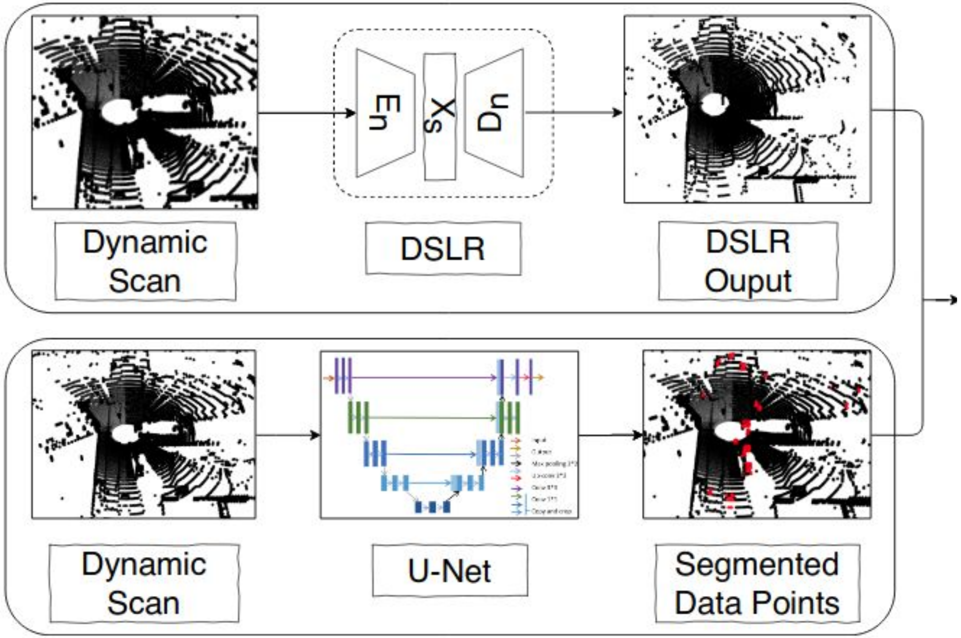
DSLR-Seg



DSLR-Seg

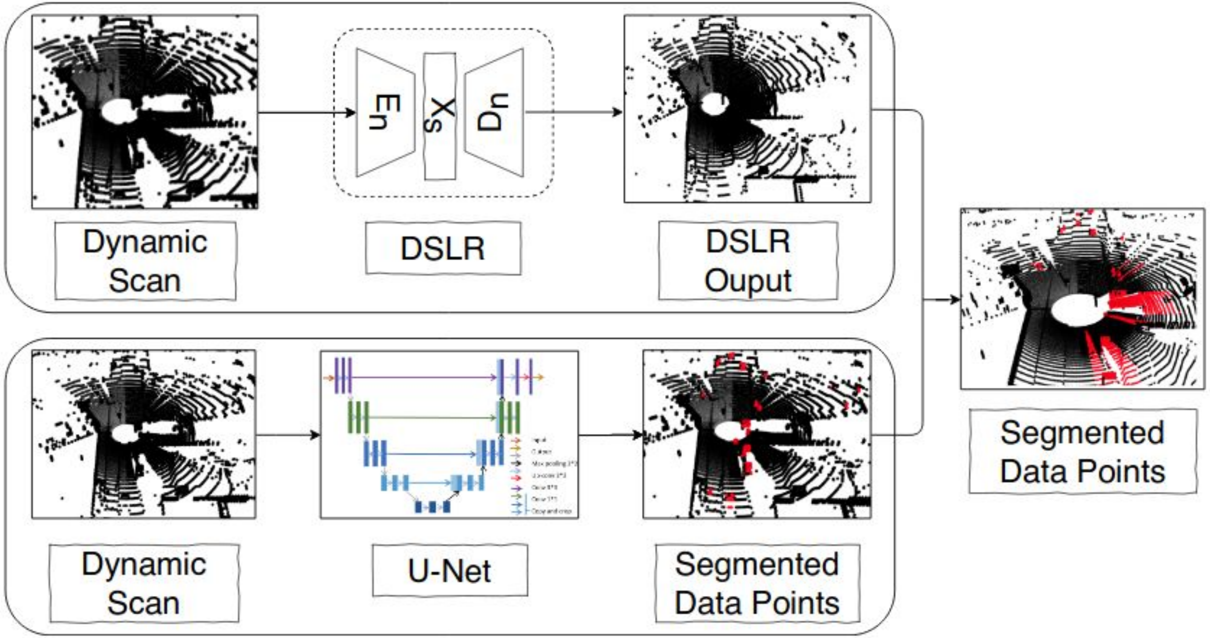


DSLR-Seg



En
Xs
Dn
Dynamic Scan
DSLR
DSLR Ouput
Dynamic Scan
U-Net
Segmented Data Points

DSLR-Seg



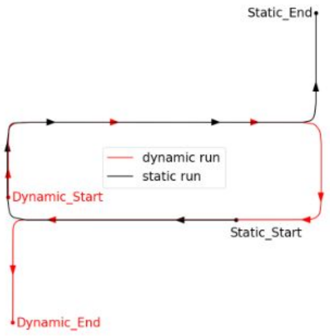
En
Xs
Dn
Dynamic Scan
DSLR
DSLR Ouput
Dynamic Scan
U-Net
Segmented Data Points
Segmented Data Points

Outline

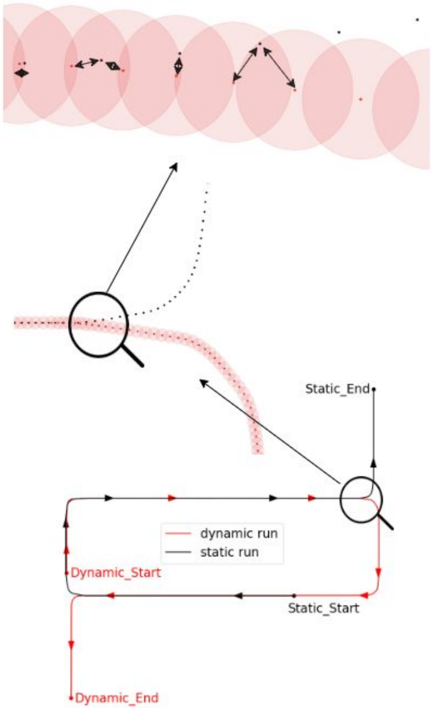
Dataset Generation
(generate dynamic and corresponding static pairs)

- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - DSLR-UDA (w/o paired static)
  - Lidar scan Quality Index
  - DSLR-Seg (with seg info)
  - Results
- **Dataset Generation**
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - Results

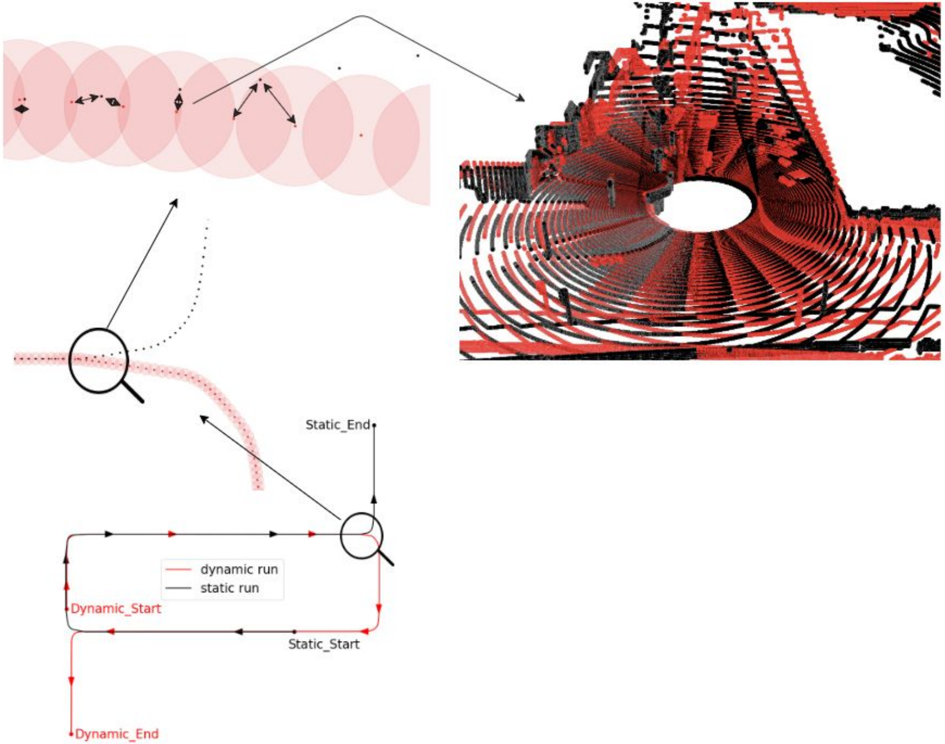
Dataset Generation



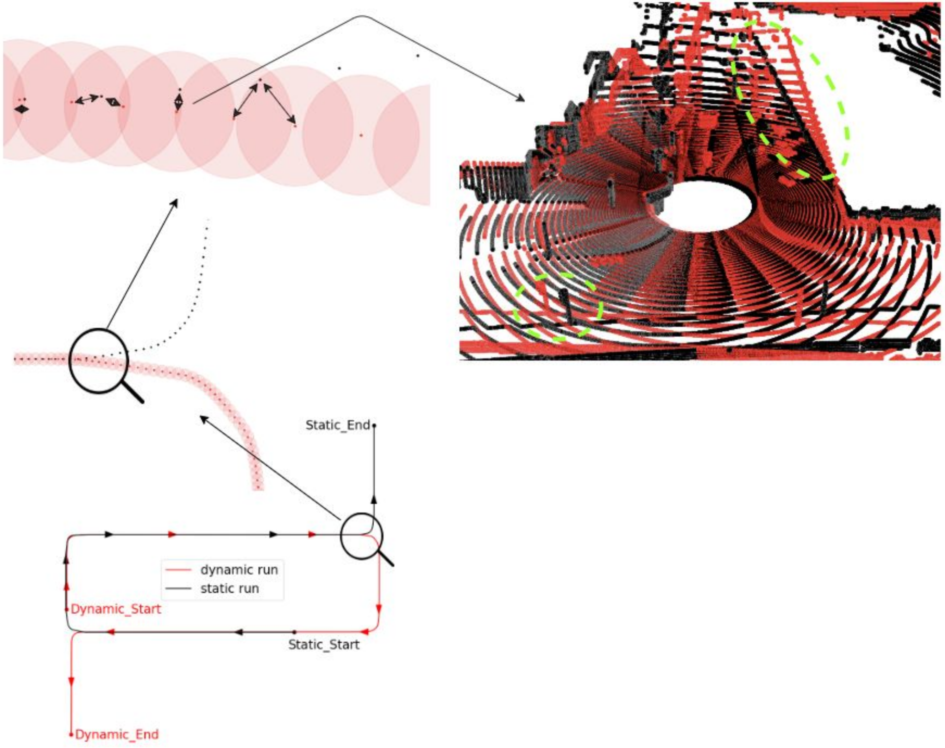
Dataset Generation



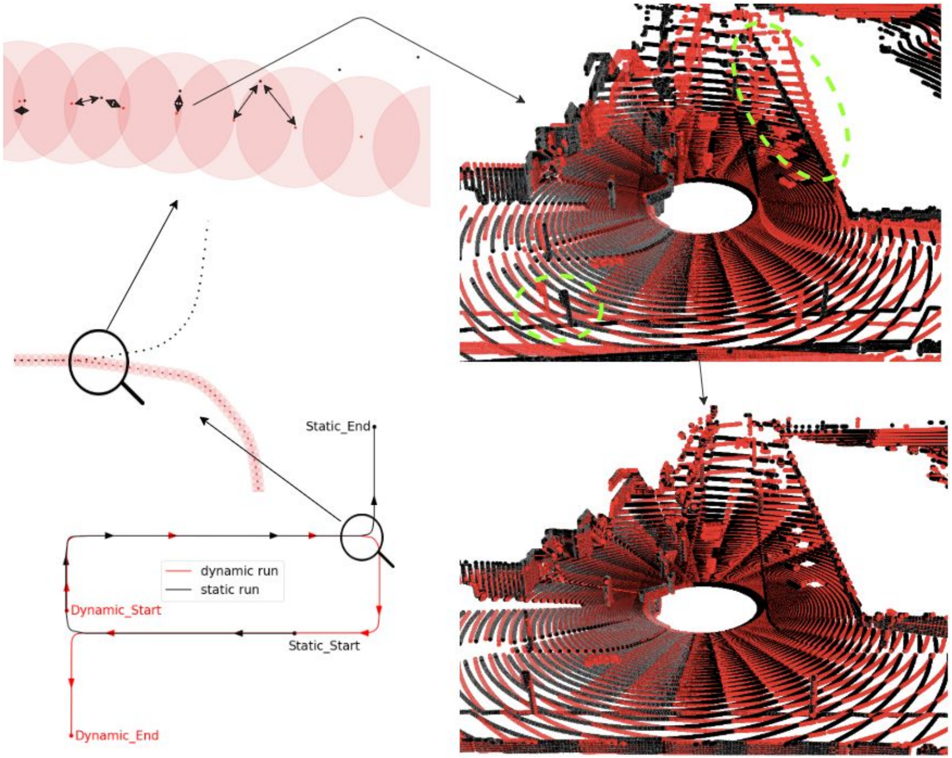
Dataset Generation



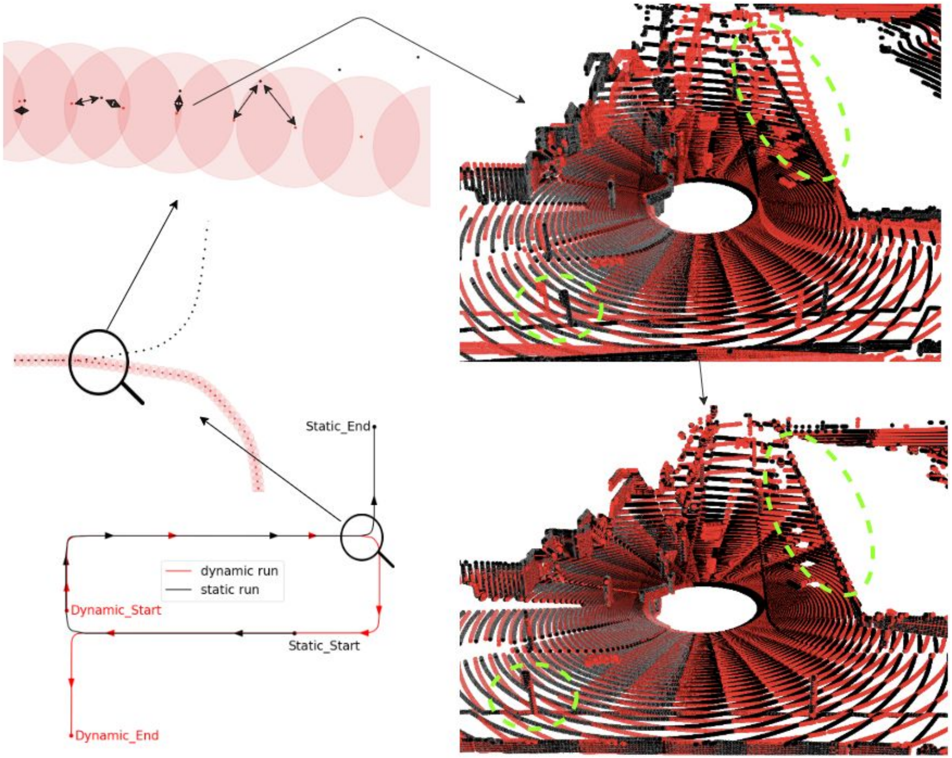
Dataset Generation



Dataset Generation

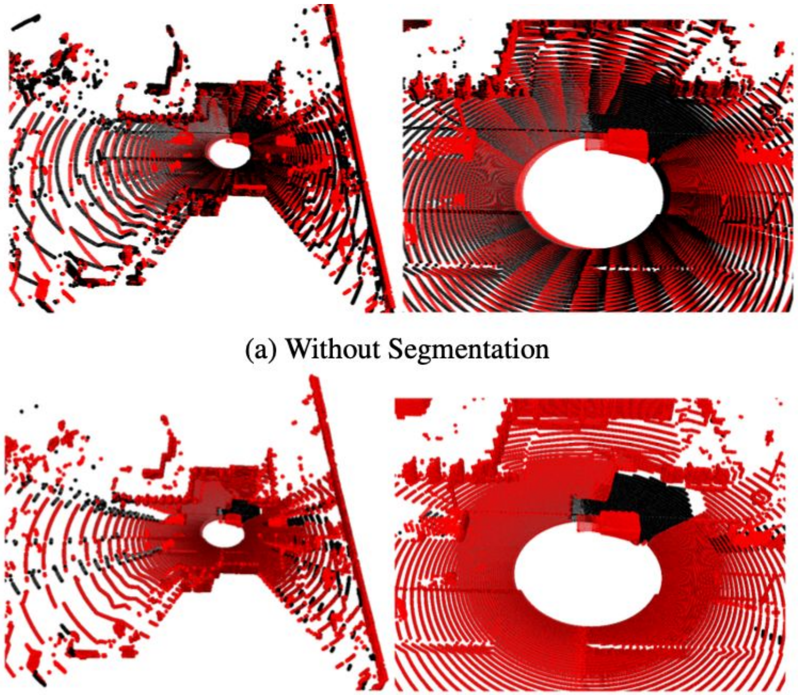


Dataset Generation



DSLR++

Segmentation can improve dataset quality



(a) Without Segmentation

(b) With Segmentation

Outline

DST for LiDAR Results

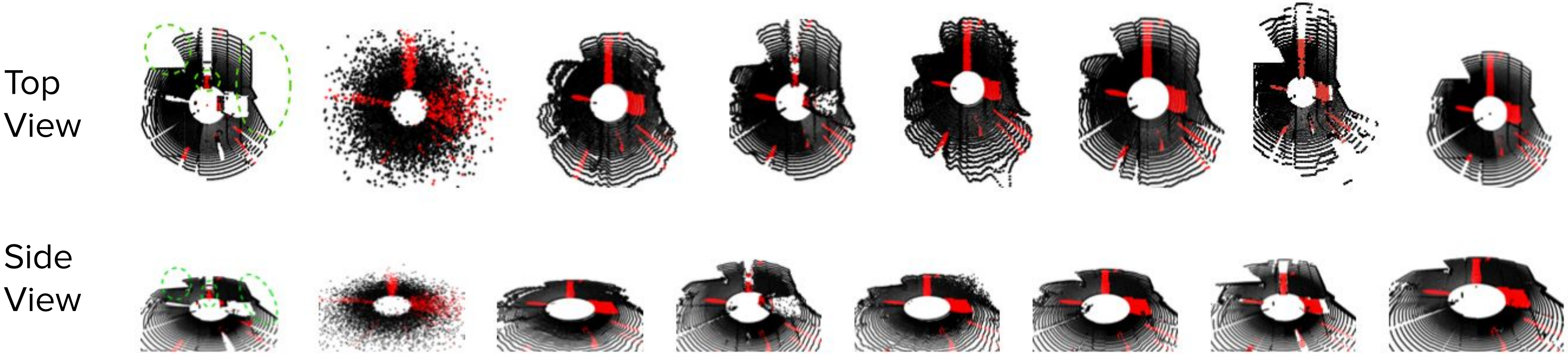
- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - DSLR-UDA (w/o paired static)
  - Lidar scan Quality Index
  - DSLR-Seg (with seg info)
  - **Results**
- Dataset Generation
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - Results

Quantitative LiDAR Reconstruction Results

Model	Uses Seg	Carla-64			ARD-16		KITTI-64
		EMD	Chamfer	LQI	EMD	Chamfer	LQI
AtlasNet	No	5681.85	5109.98	-	1464.61	176.46	-
ADMG	No	397.94	6.23	7.049	309.64	1.62	2.911
CHCP-VAE	No	343.98	9.58	4.080	88.94	0.67	1.128
CHCP-GAN	No	329.38	8.19	3.519	65.24	0.38	1.133
CHCP-AE	No	253.91	4.05	3.720	65.40	0.31	1.738
WCZC	Yes	$2.73*10^6$	478.12	-	-	-	-
EmptyCities	Yes	640.97	29.39	-	-	-	-
DSLR (Ours)	No	**232.51**	**1.00**	**3.350**	**57.75**	**0.20**	1.120
DSLR++ (Ours)	Yes	205.48	0.49	-	-	-	
DSLR-Seg (Ours)	Yes	**150.90**	**0.02**	-	-	-	
DSLR-UDA(Ours)	Yes	-	-	-	-	-	**1.119**

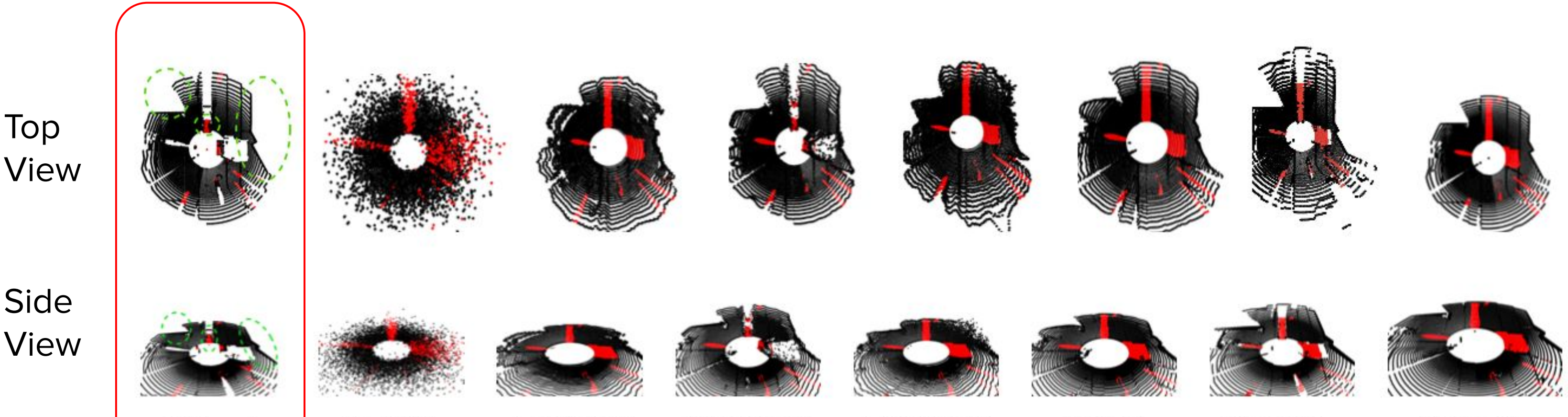
Comparison of reconstruction of DSLR and its variants (DSLR-Seg, DSLR-UDA, DSLR++) against baselines on Earth Mover Distance, Chamfer Distance and LQI metrics for the 3 datasets.

CARLA-64 Reconstruction Results



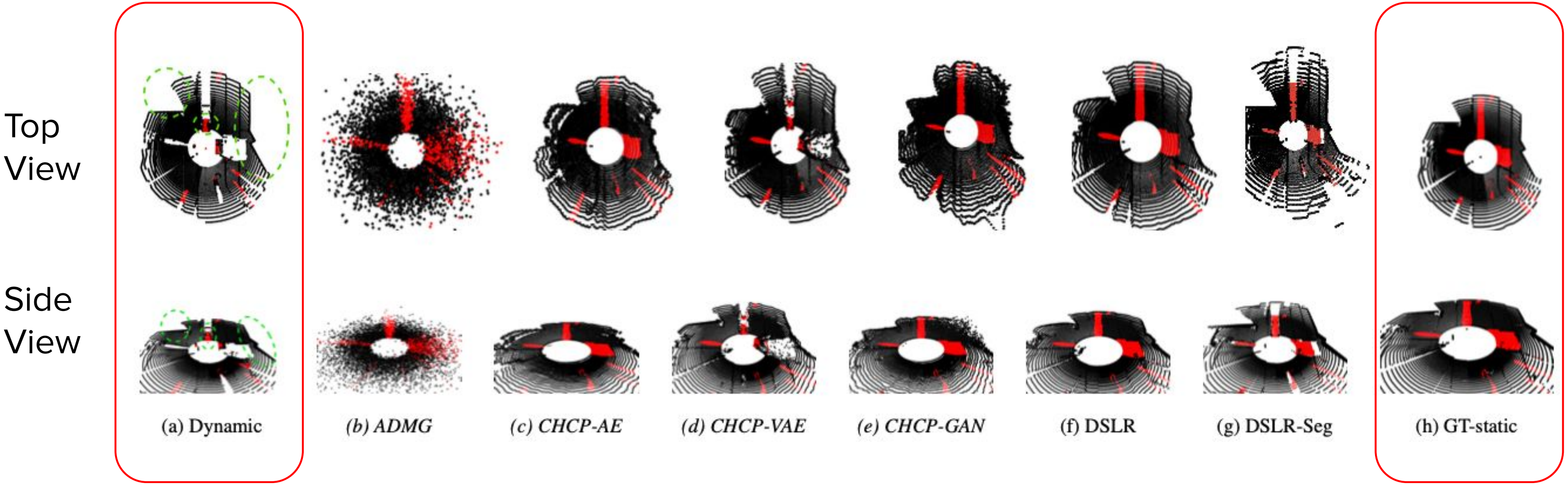
(a) Dynamic (b) *ADMG* (c) *CHCP-AE* (d) *CHCP-VAE* (e) *CHCP-GAN* (f) DSLR (g) DSLR-Seg (h) GT-static

CARLA-64 Reconstruction Results

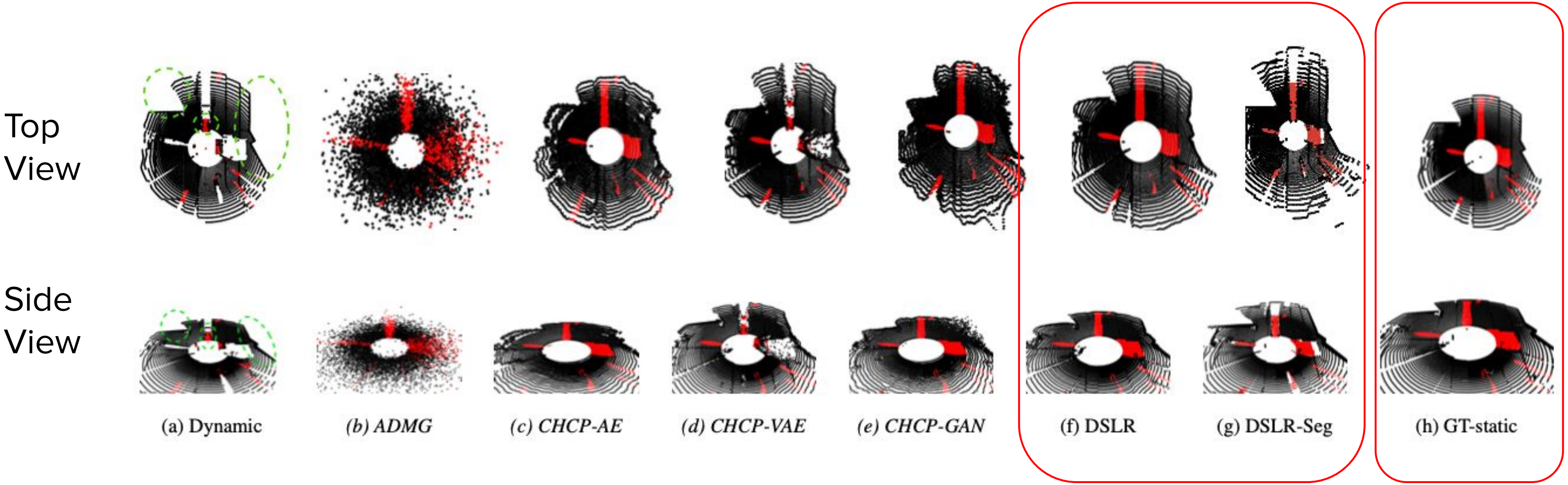


(a) Dynamic (b) ADMG (c) CHCP-AE (d) CHCP-VAE (e) CHCP-GAN (f) DSLR (g) DSLR-Seg (h) GT-static

CARLA-64 Reconstruction Results

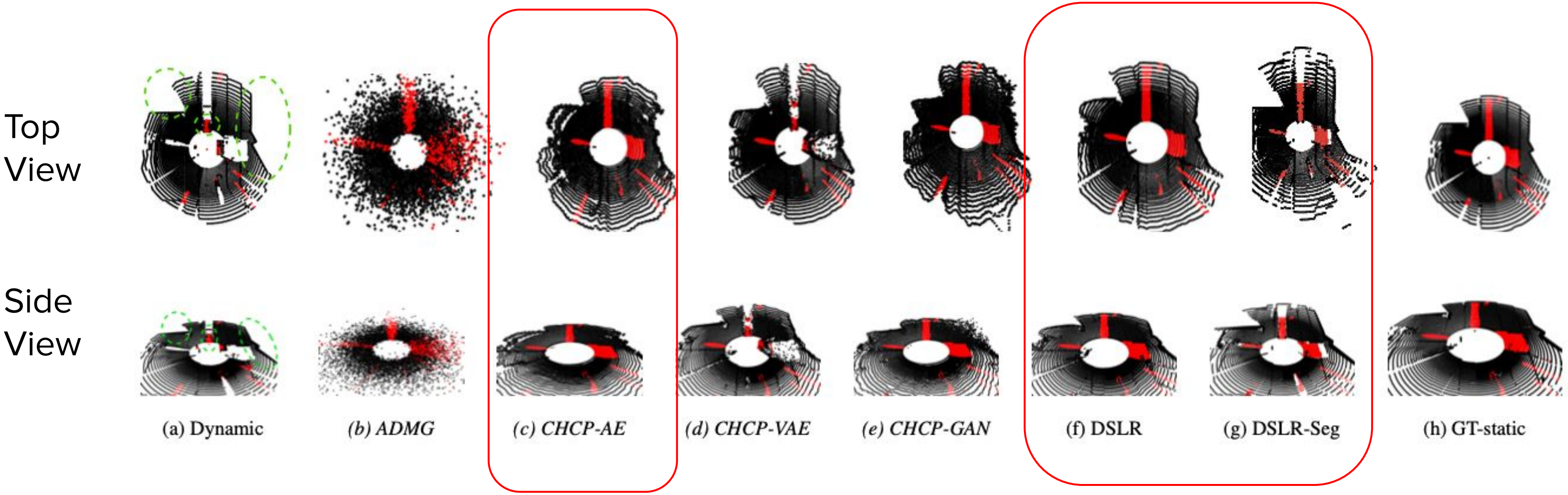


CARLA-64 Reconstruction Results



(a) Dynamic (b) ADMG (c) CHCP-AE (d) CHCP-VAE (e) CHCP-GAN (f) DSLR (g) DSLR-Seg (h) GT-static

CARLA-64 Reconstruction Results



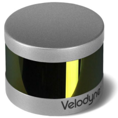
(a) Dynamic (b) *ADMG* (c) *CHCP-AE* (d) *CHCP-VAE* (e) *CHCP-GAN* (f) DSLR (g) DSLR-Seg (h) GT-static

Outline

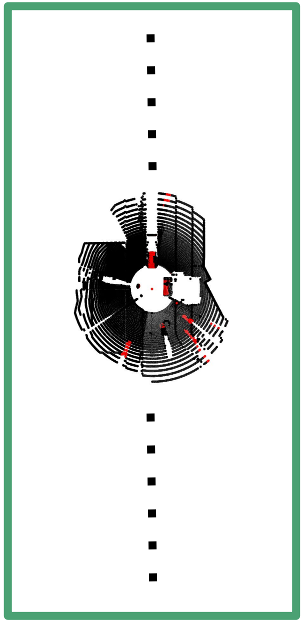
SLAM Setup

- Solve DST for LiDAR
    - DSLR (w/o seg info)
    - DSLR-UDA (w/o paired static)
    - Lidar scan Quality Index
    - DSLR-Seg (with seg info)
    - Results
- Dataset Generation
- Solve DST for LIDAR based SLAM
    - **SLAM Setup**
    - SLAM Reconstruction Threshold
    - Results

Simultaneous Localisation and Mapping (SLAM)



Sensory Input

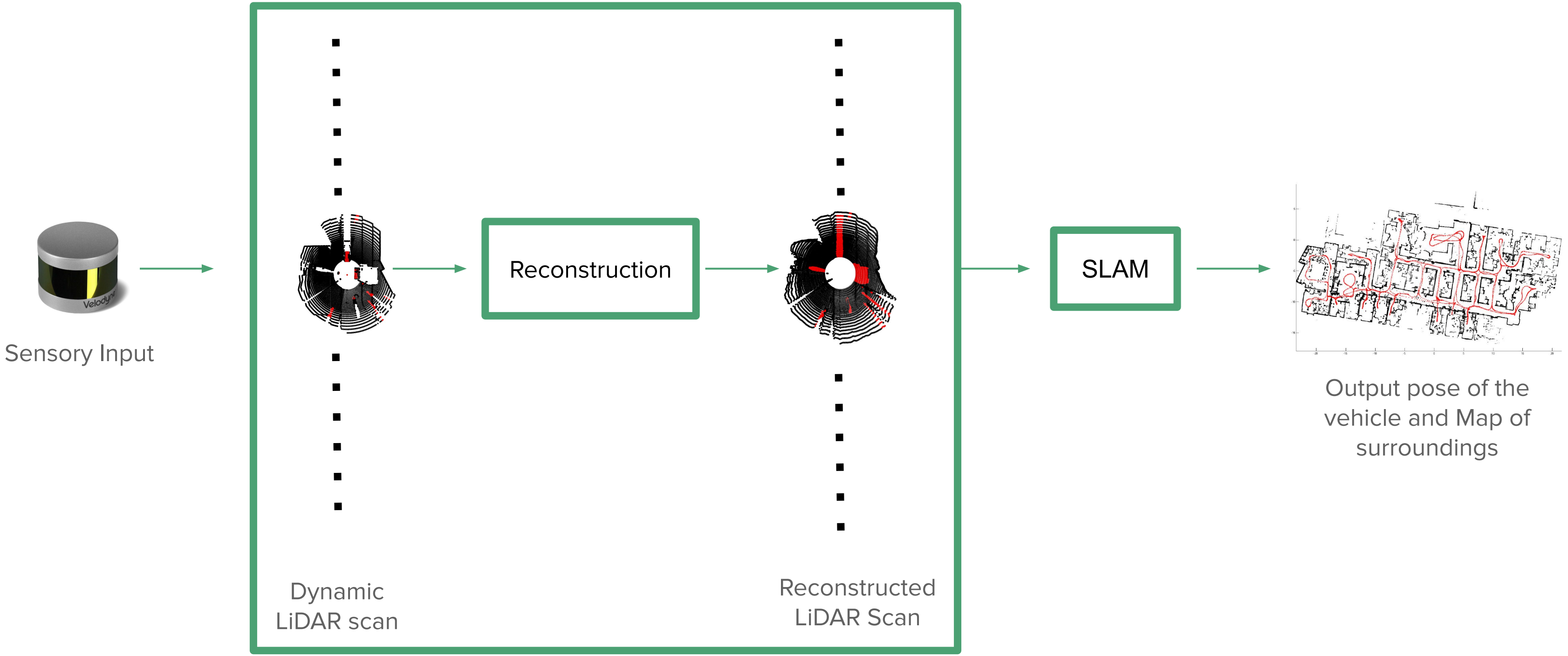


LiDAR scan

SLAM

Output Pose of the vehicle and Map of surroundings.

Our Experimental Setup



Outline

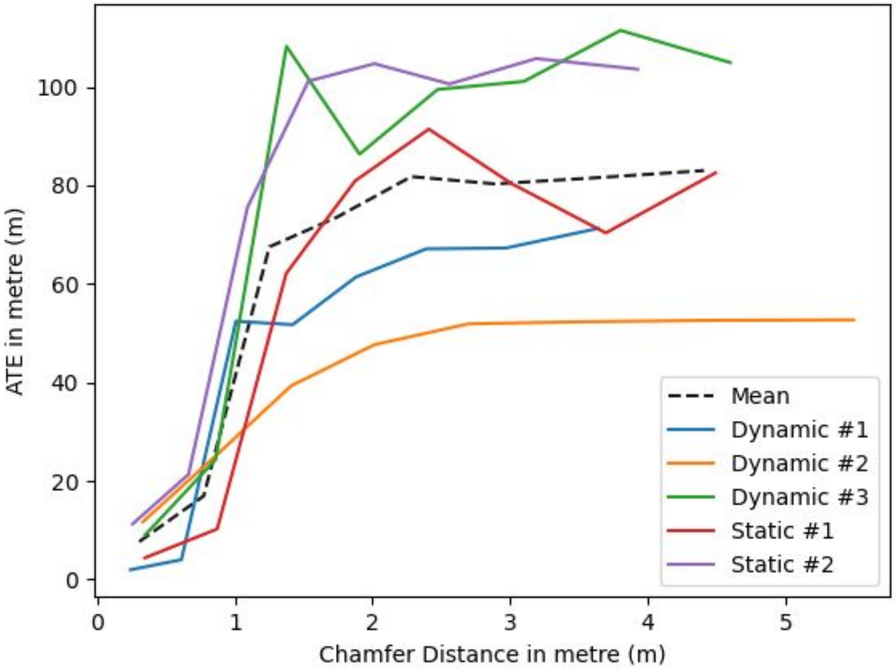
SLAM Reconstruction Threshold
or SRT
(to be practically feasible)

- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - DSLR-UDA (w/o paired static)
  - Lidar scan Quality Index
  - DSLR-Seg (with seg info)
  - Results

- Dataset Generation

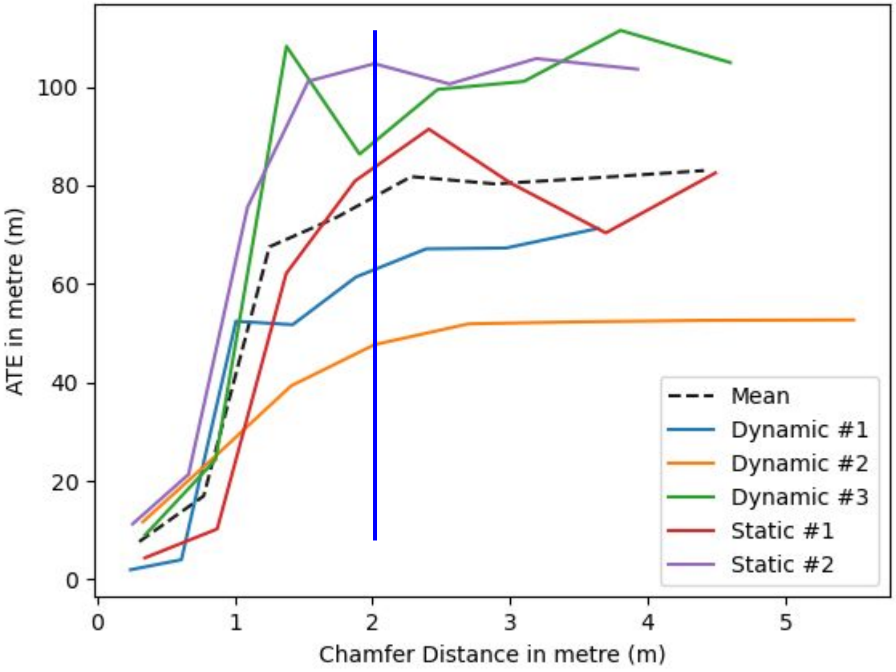
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - **SLAM Reconstruction Threshold**
  - Results

SLAM Reconstruction Threshold



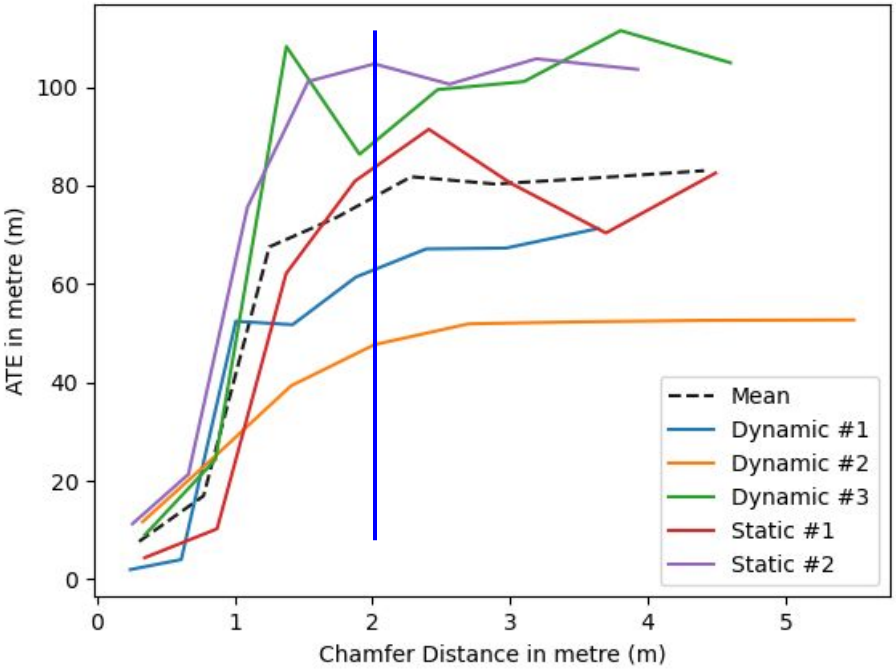
ATE is Absolute Trajectory Error

SLAM Reconstruction Threshold



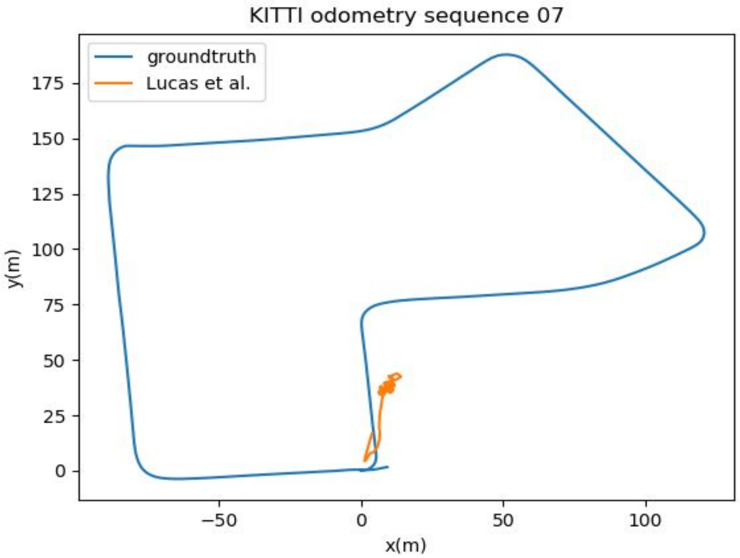
ATE is Absolute Trajectory Error

SLAM Reconstruction Threshold



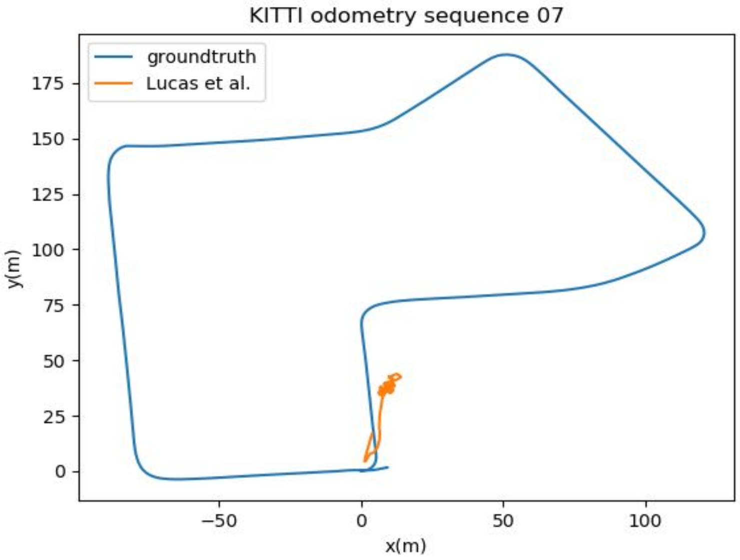
Model	CARLA-64
	Chamfer
AtlasNet	5109.98
ADMG	6.23
CHCP-VAE	9.58
CHCP-GAN	8.19
CHCP-AE	4.05
WCZC	478.12
EmptyCities	29.39
DSLR (Ours)	**1.00**
DSLR++ (Ours)	**0.49**
DSLR-Seg (Ours)	**0.02**
DSLR-UDA(Ours)	-

Localization plots

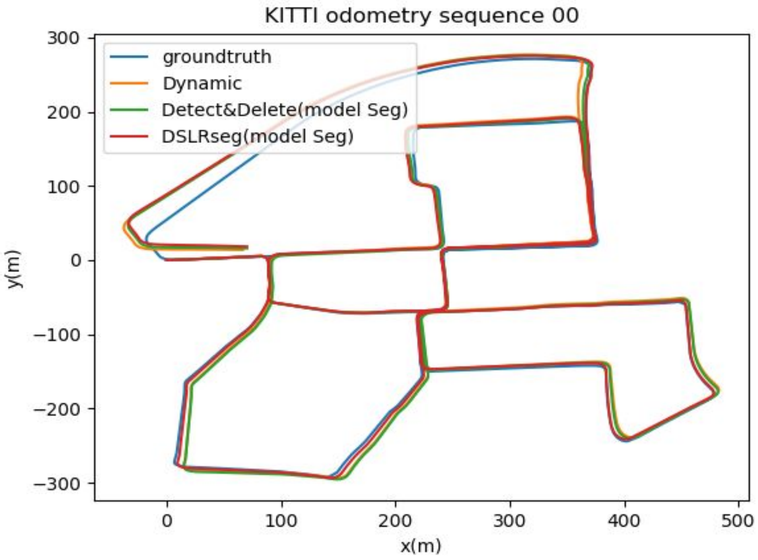


Degradation in SLAM output below SRT.

Localization plots



Degradation in SLAM output below SRT.



Reconstruction required from our model is required for performing competitively against existing Dynamic SLAM methods.

Outline

DST for LIDAR based SLAM
Results

- Solve DST for LiDAR
  - DSLR (w/o seg info)
  - DSLR-UDA (w/o paired static)
  - Lidar scan Quality Index
  - DSLR-Seg (with seg info)
  - Results
- Dataset Generation
- Solve DST for LIDAR based SLAM
  - SLAM Setup
  - SLAM Reconstruction Threshold
  - **Results**

Quantitative SLAM Results

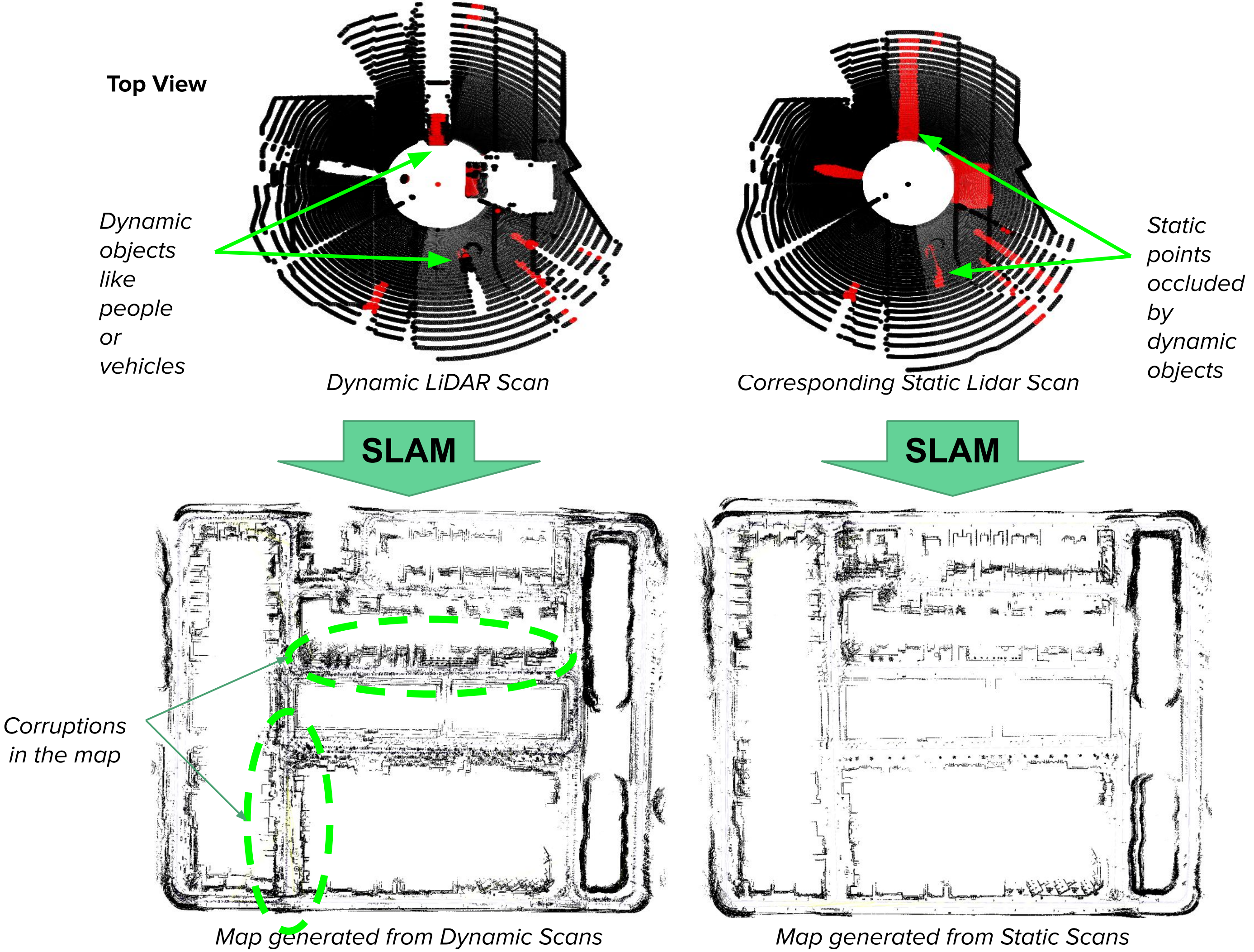
MS depicts the usage of segmentation masks have been generated by a segmentation model

GTS depicts the usage of ground truth segmentation masks

Model	ATE	Drift	RPE	
			Trans	Rot
KITTI-64 Dataset				
Pure-Dynamic	11.809	14.970	1.620	1.290°
Detect & Delete (MS)	12.490	**13.599**	1.623	1.290°
Detect & Delete (GTS)	**11.458**	22.856	1.630	1.336°
DSLR-Seg (MS)	99.642	34.372	**1.610**	**1.290°**
DSLR-Seg (GTS)	11.699	19.67	1.620	1.295°
CARLA-64 Dataset				
Pure-Dynamic	10.360	18.580	0.056	0.403°
Detect & Delete (MS)	10.430	18.870	0.060	0.430°
Detect & Delete (GTS)	11.650	24.430	0.063	0.401°
DSLR-Seg (MS)	10.760	17.430	0.050	0.390°
DSLR-Seg (GTS)	**7.330**	**13.63**	**0.050**	**0.340°**
ARD-16 Dataset				
Pure-Dynamic	1.701	–	0.036	**0.613°**
DSLR (Ours)	**1.680**	–	**0.035**	0.614°

Thank you

Project page : https://dslrproject.github.io/dslr/


Top View
Dynamic objects like people or vehicles
Dynamic LiDAR Scan
Corresponding Static Lidar Scan
Static points occluded by dynamic objects
SLAM
SLAM
Corruptions in the map
Map generated from Dynamic Scans
Map generated from Static Scans

CARLA-64 Reconstruction Results