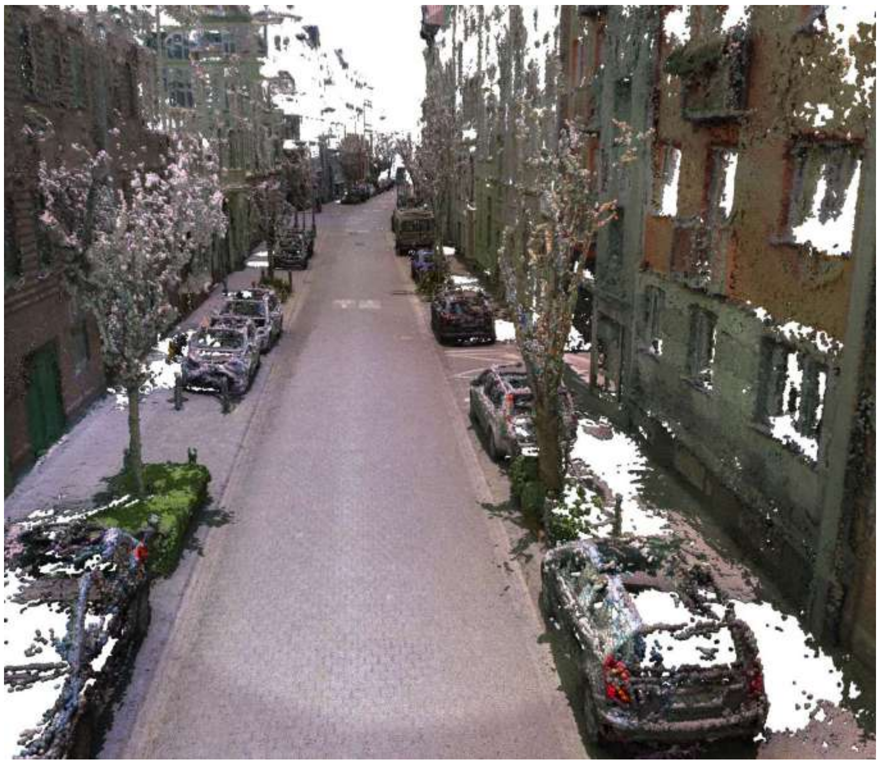
# Mapping and Semantic Perception in Urban Environments

**Jens Behley**

UNIVERSITÄT BONN

# Mapping the Environment



**Digital Twin**

**Autonomous Operation**

# Understanding the Environment

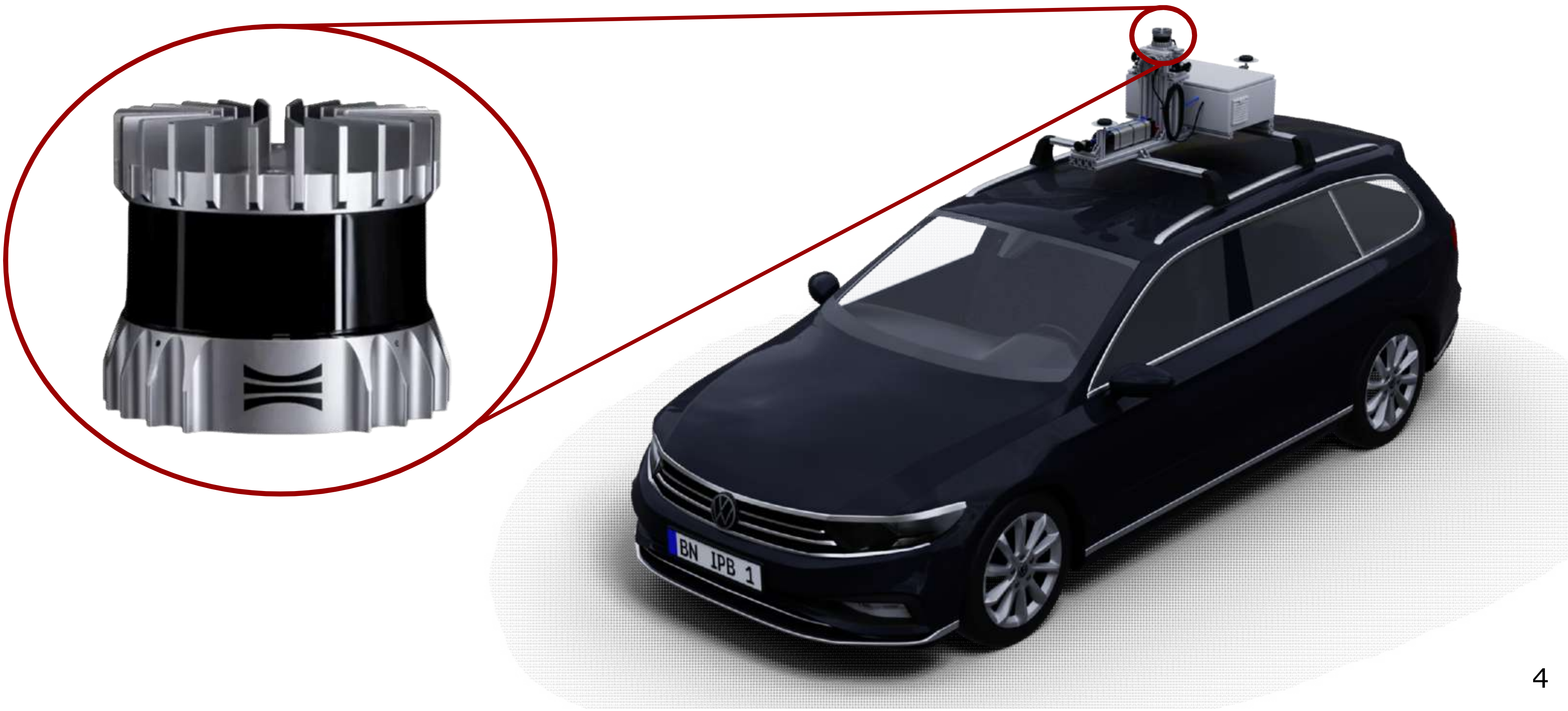
- Ego-pose of the vehicle
- Location and velocity of all traffic participants
- Scene Semantics: drivable areas, lanes, traffic signs, parking areas, …
- State of traffic lights, Police officers, …

**Scene interpretation**

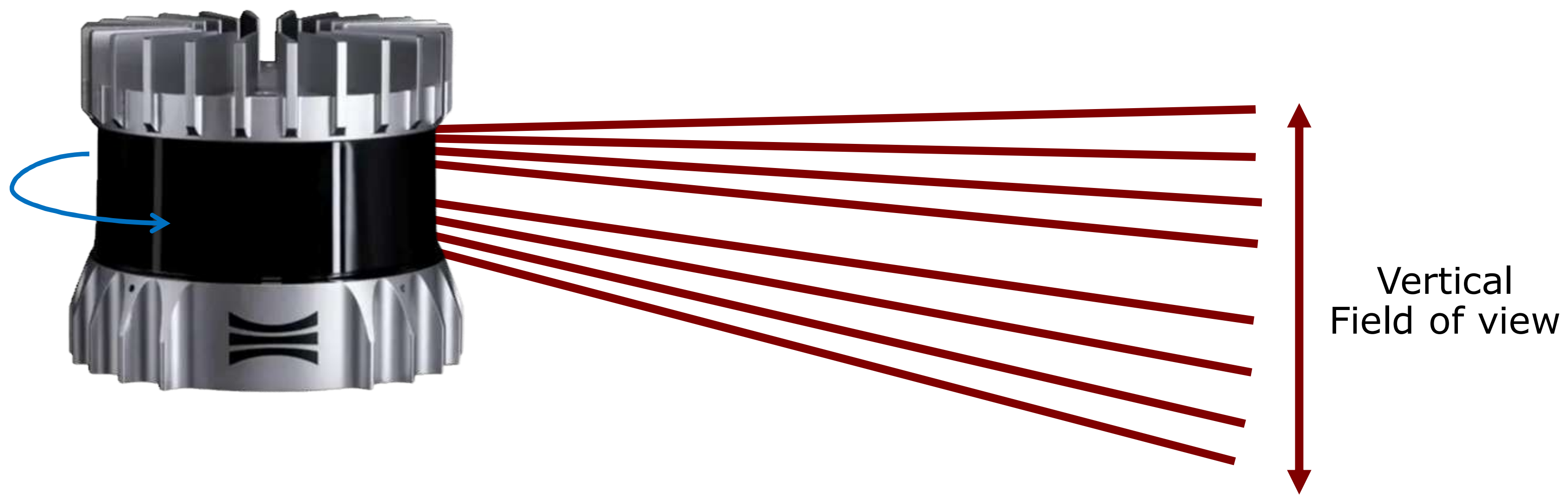
- Future locations of traffic participants
- Intend of traffic participants
- Traffic situation and interaction between traffic participants

**Scene understanding**

# 3D LiDAR Sensor

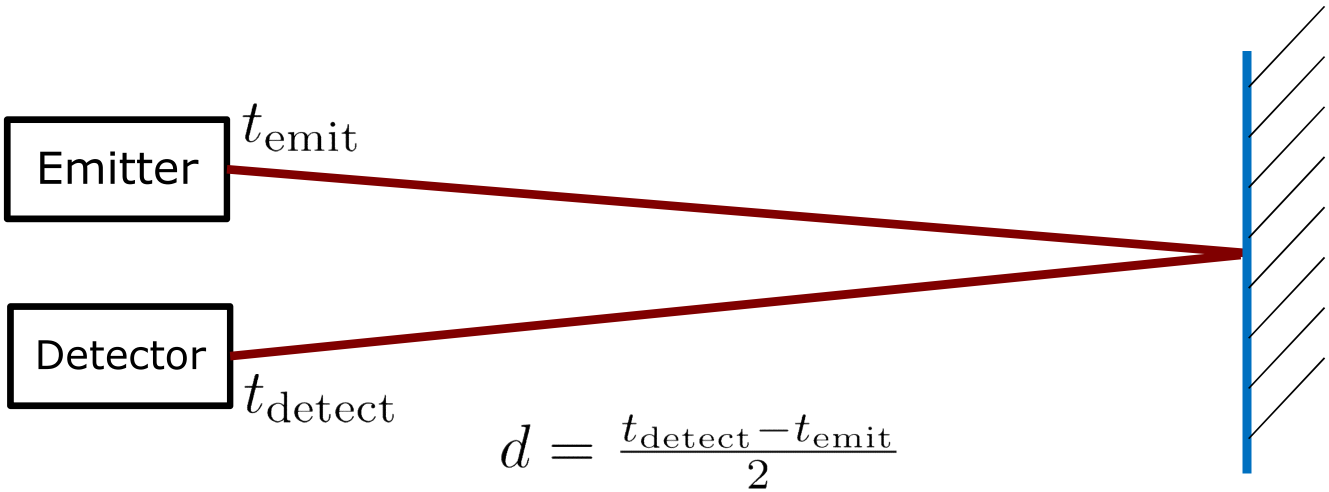


# 3D LiDAR Sensors



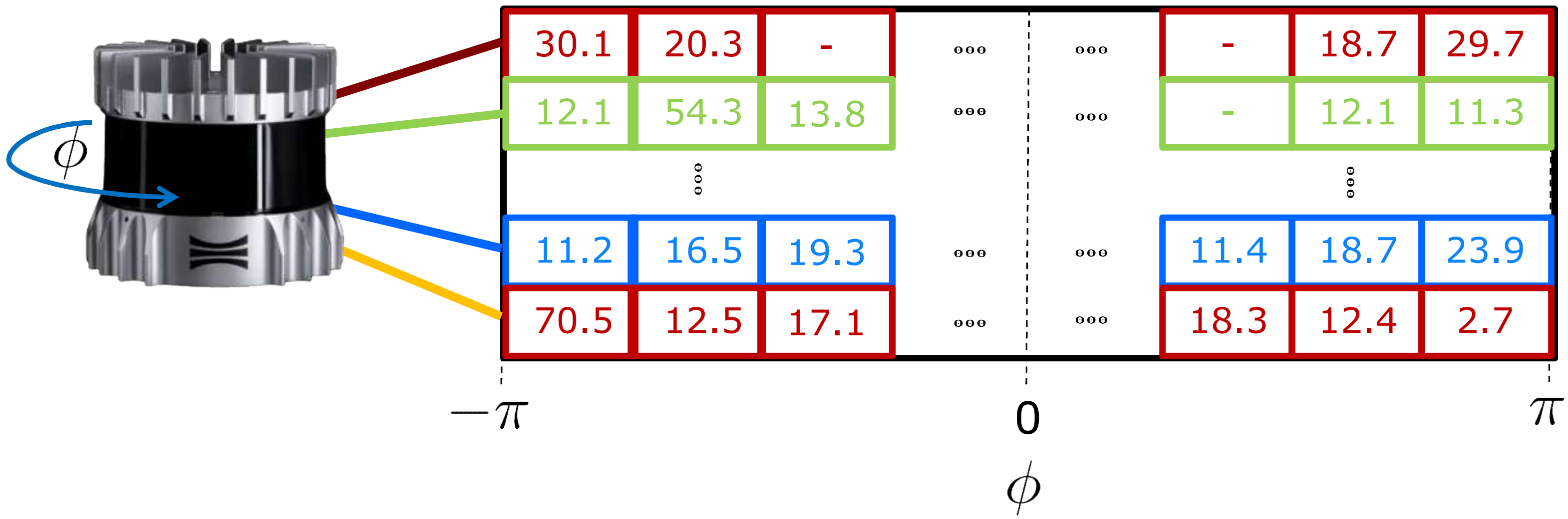

- Rotating multi-beam LiDAR sensors provides 360 degree horizontal field of view
- 10-20 Hz for complete sweep/turn

# Measurement Principle




- Distance $d$ = time of flight of photons (900 nm-1550 nm) from emittance to detection

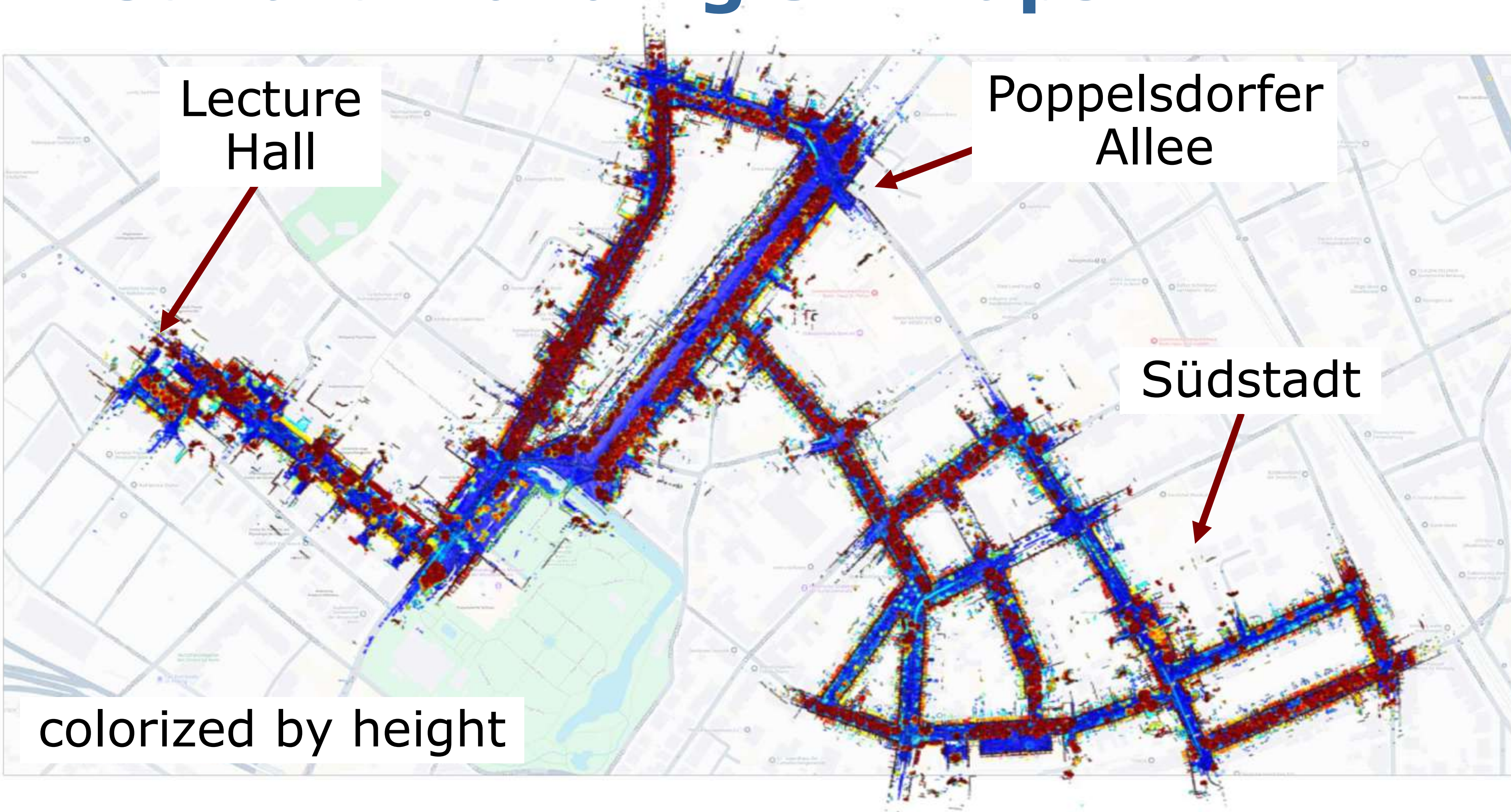
# Range Image



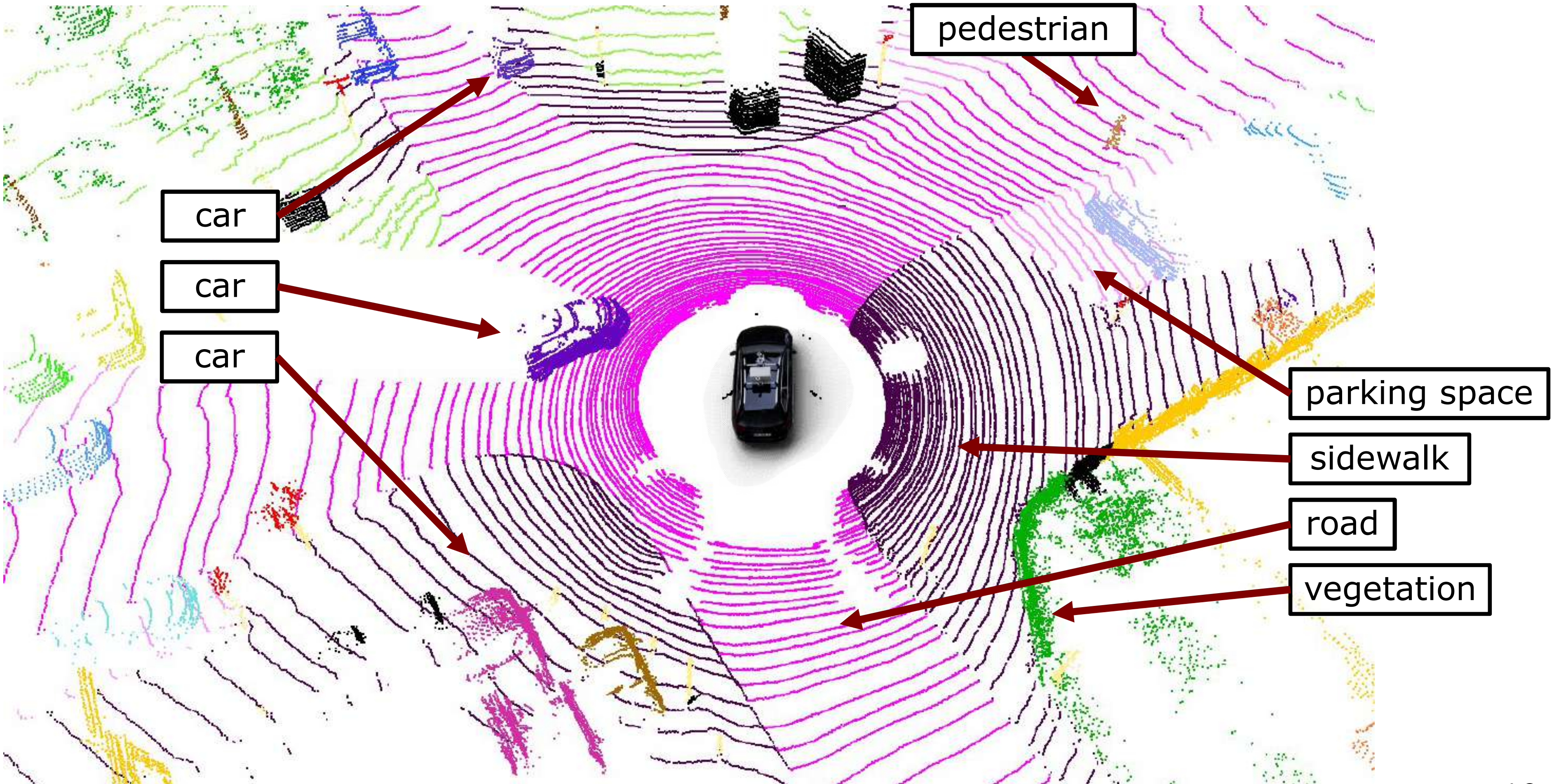

- Range image has ranges for each beam and turn angle $\phi$
- Common widths: 1024, 2048

# Obtaining 3D LiDAR Data

# First Part: Building 3D Maps

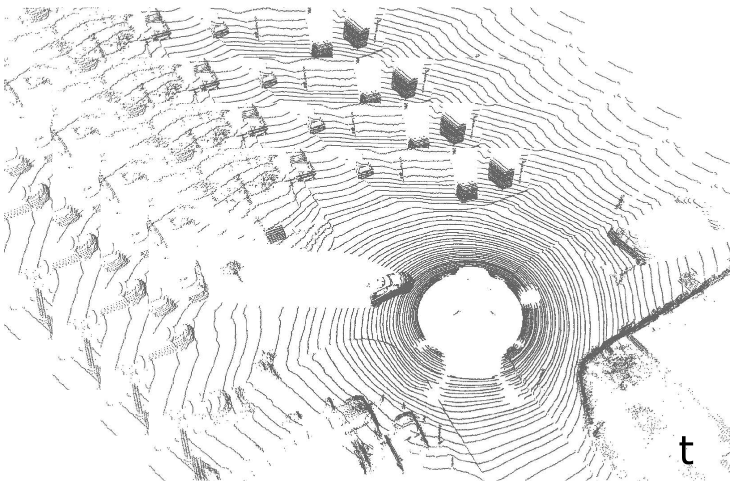


# Second Part: Semantic Interpretation



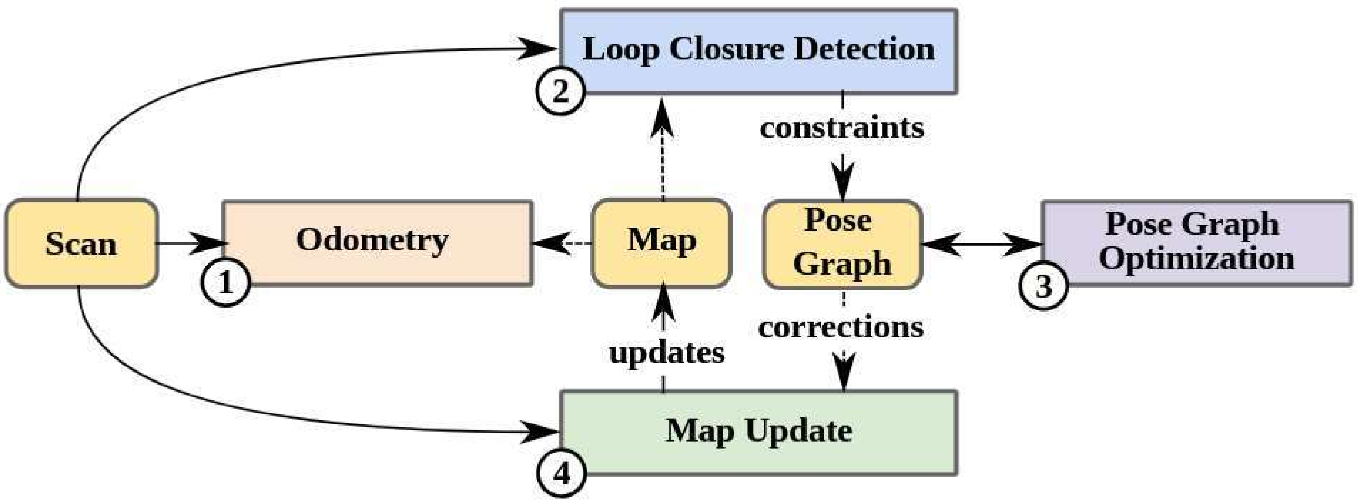
# Mapping the Environment

# Mapping the Environment




- **Input:** Sequence of LiDAR scans
- **Goal:** Build a globally-consistent (geo-referenced) map

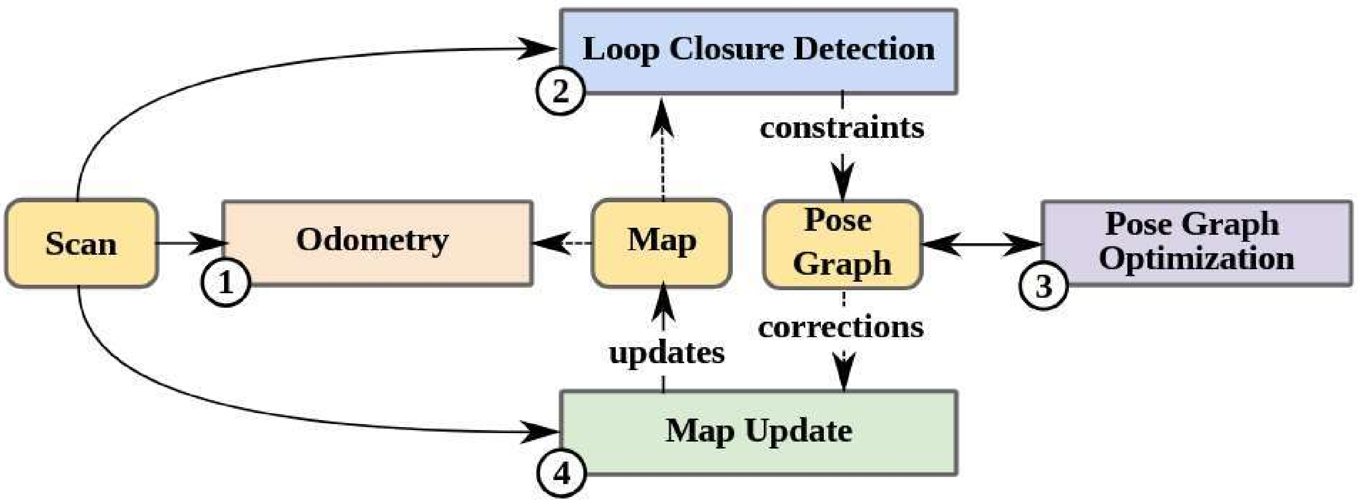
# LiDAR-based Simultaneous Localization and Mapping (SLAM)




- Sensor pose estimation requires a map (localization)
- Building a map requires sensor poses (mapping)

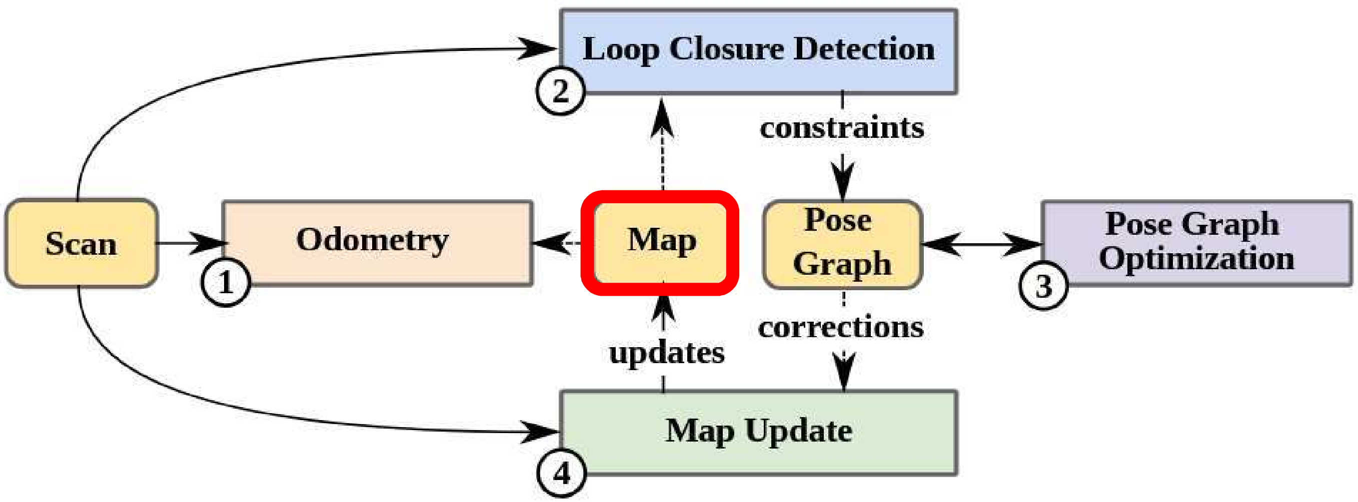
Behley et al. LiDAR SLAM in Carlone et al. SLAM Handbook: From Localization to Spatial Intelligence, Cambridge University Press, 2025. (in press)

# SLAM in a Nutshell



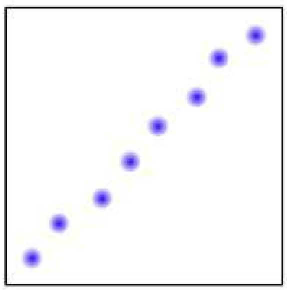

- Odometry = Relative poses estimation between scans
- Loop Closure Detection = Locations revisted?
- Pose Graph Optimization = Globally optimize poses based on relative pose estimates and loop closures

# Central Part of SLAM system: Map

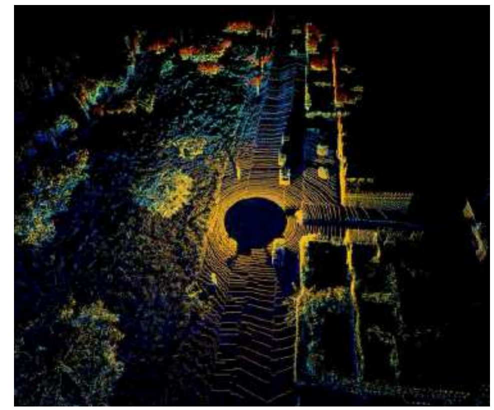



- Map plays a key role in a SLAM system
- Commonly: Geometric representations (points, surfaces, …)

# Point-Based Map



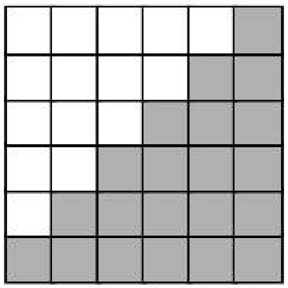
**Directly accumulated from measurements**



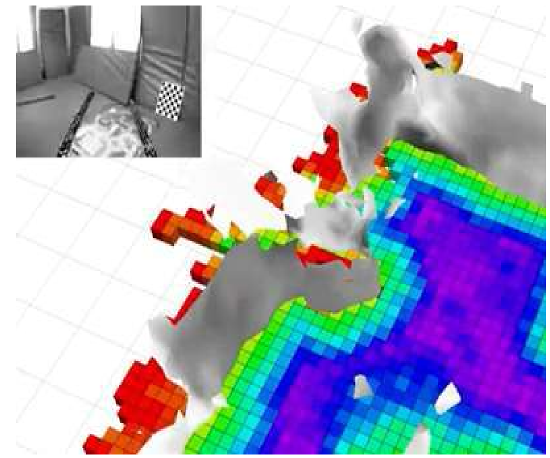
**+** Supports elastic map deformation

– Not suitable for planning

# Volumetric Map via Signed Distance



**Store occupancy probability or signed distance function (SDF) in voxels**

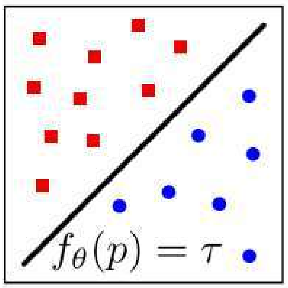


**+** Enable planning and mesh reconstruction

– Discrete and memory inefficient

– Not elastic

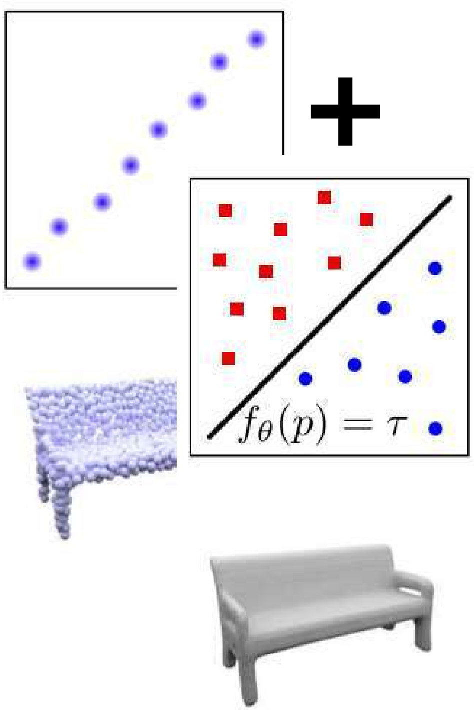
# Implicit Neural Map




**Model SDF with a neural network**

+ Enable planning and mesh reconstruction
+ Continuous, differentiable, and compact
– Not elastic

# The Best of Both Worlds: Point-Based Implicit Neural (PIN) Map

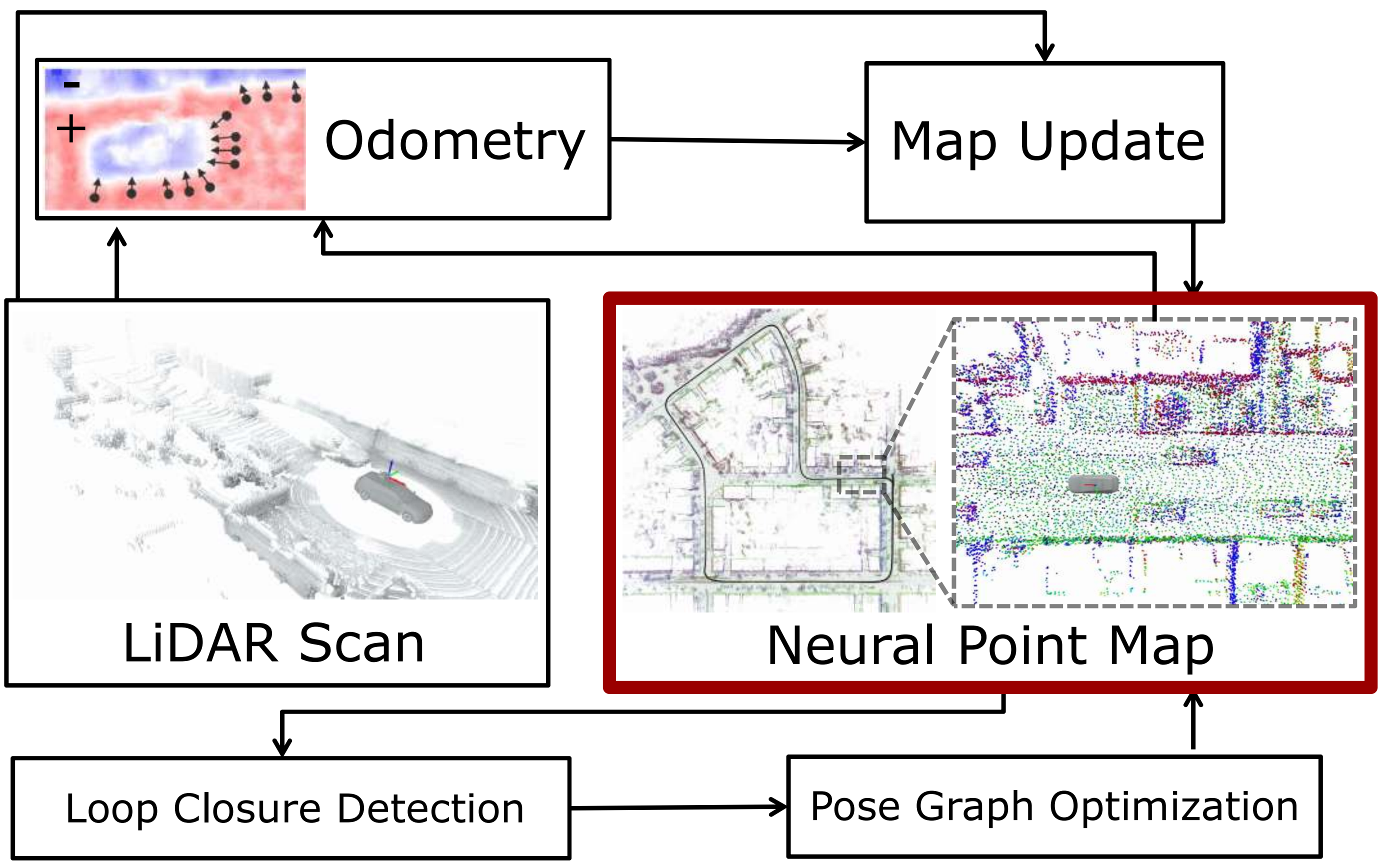



**Model SDF with locally defined neural point features and a globally shared MLP**

- + Enable planning and mesh reconstruction
- + Continuous, differentiable
- + Elastic to map deformation

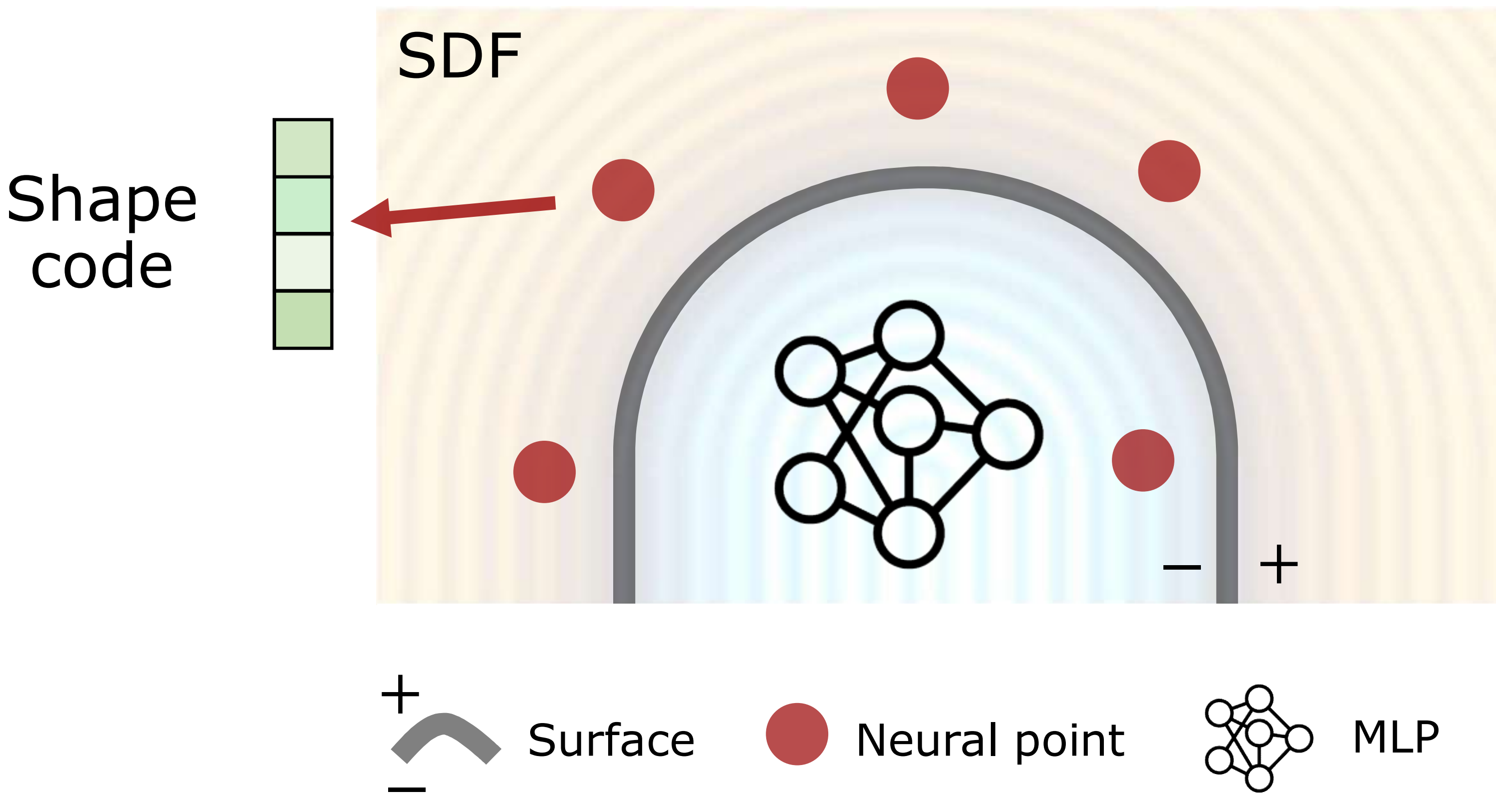
Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# PIN-SLAM Pipeline



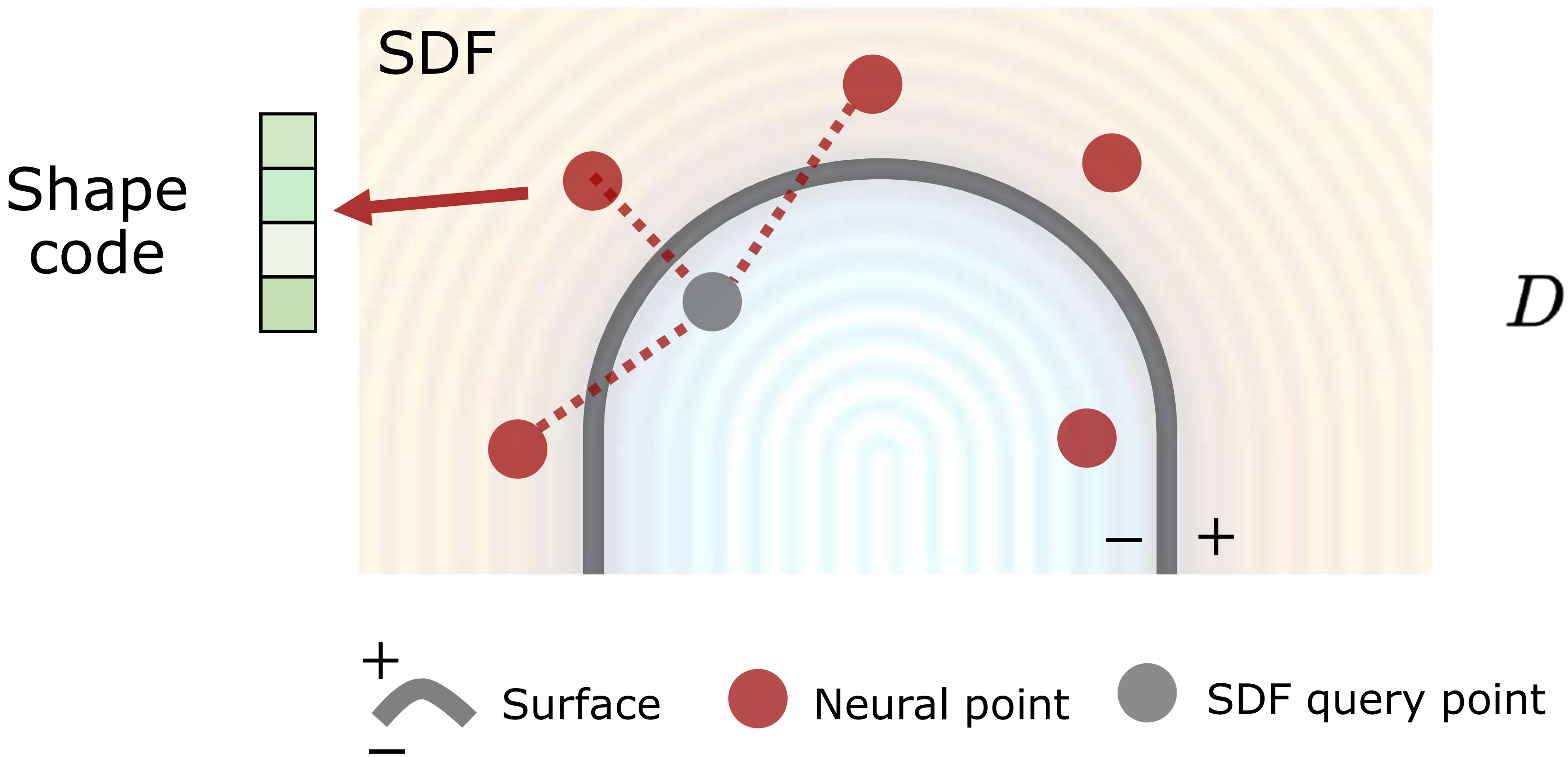

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Point-Based Neural Distance Field




Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

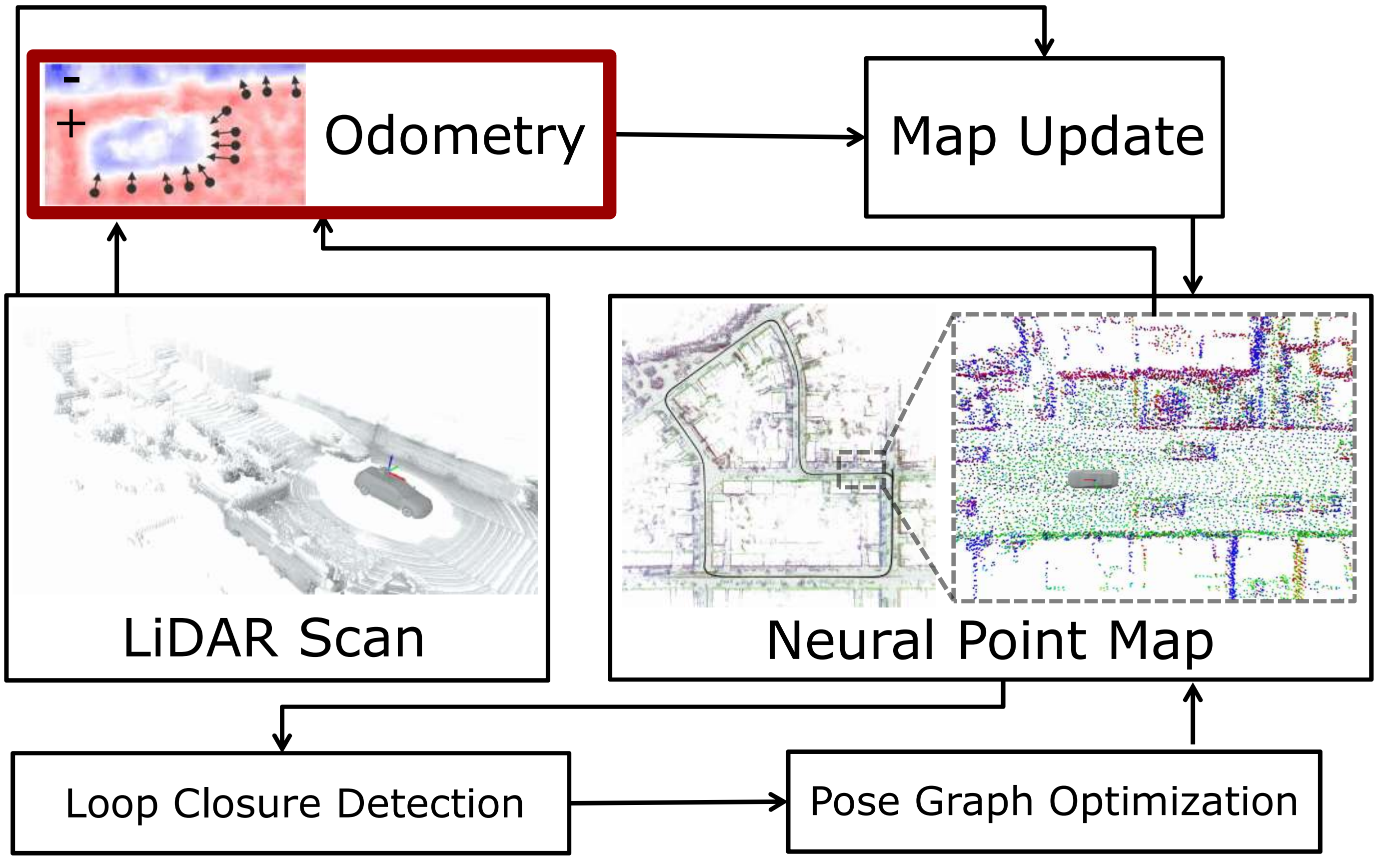
# Query SDF From Neural Points




$$D(x, y) = ?$$

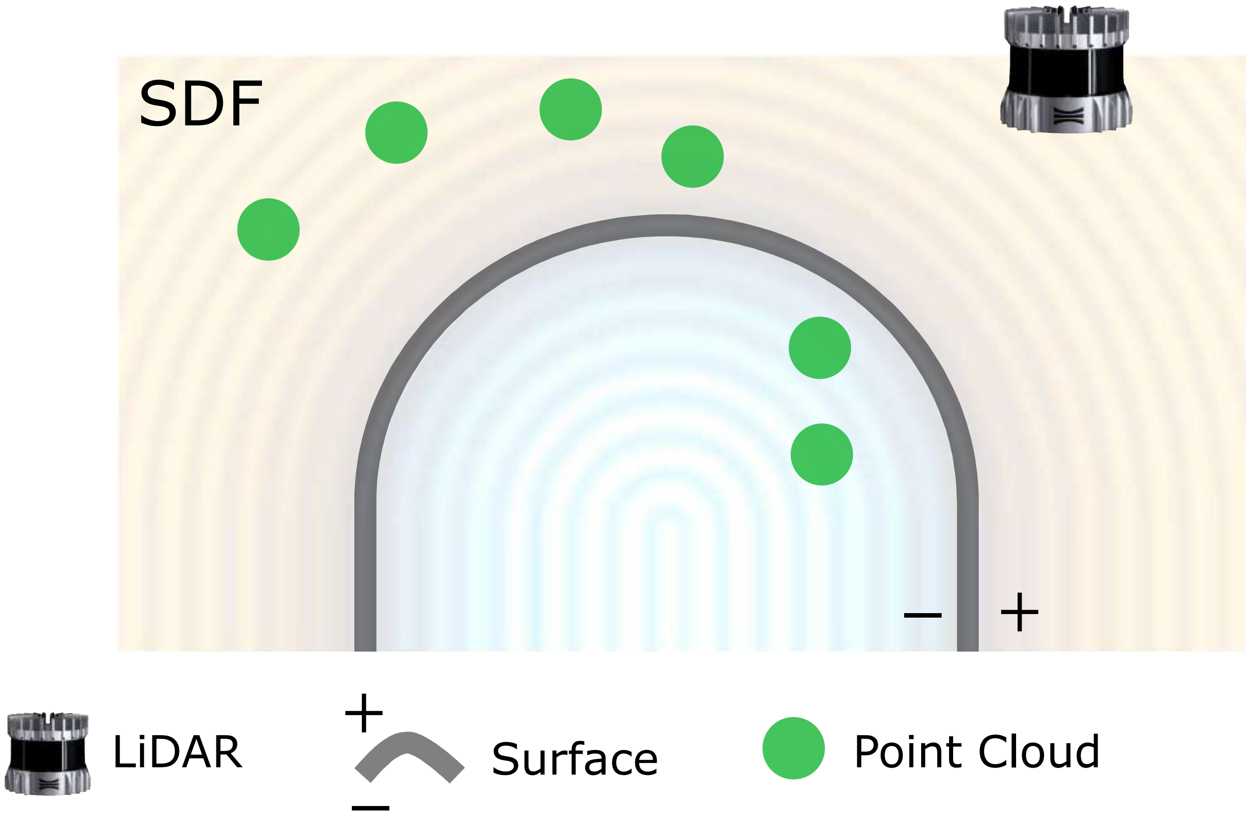
Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# PIN-SLAM Pipeline



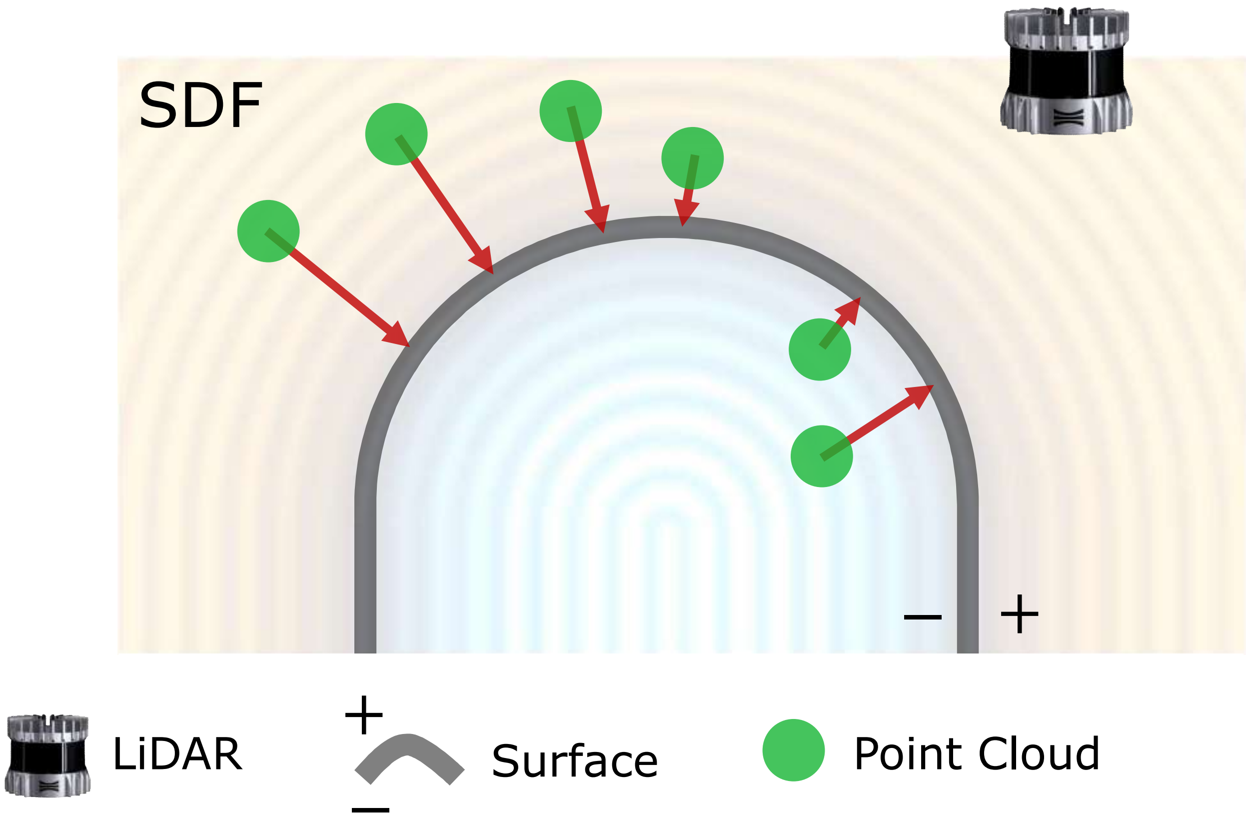

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Localization in Neural Distance Field



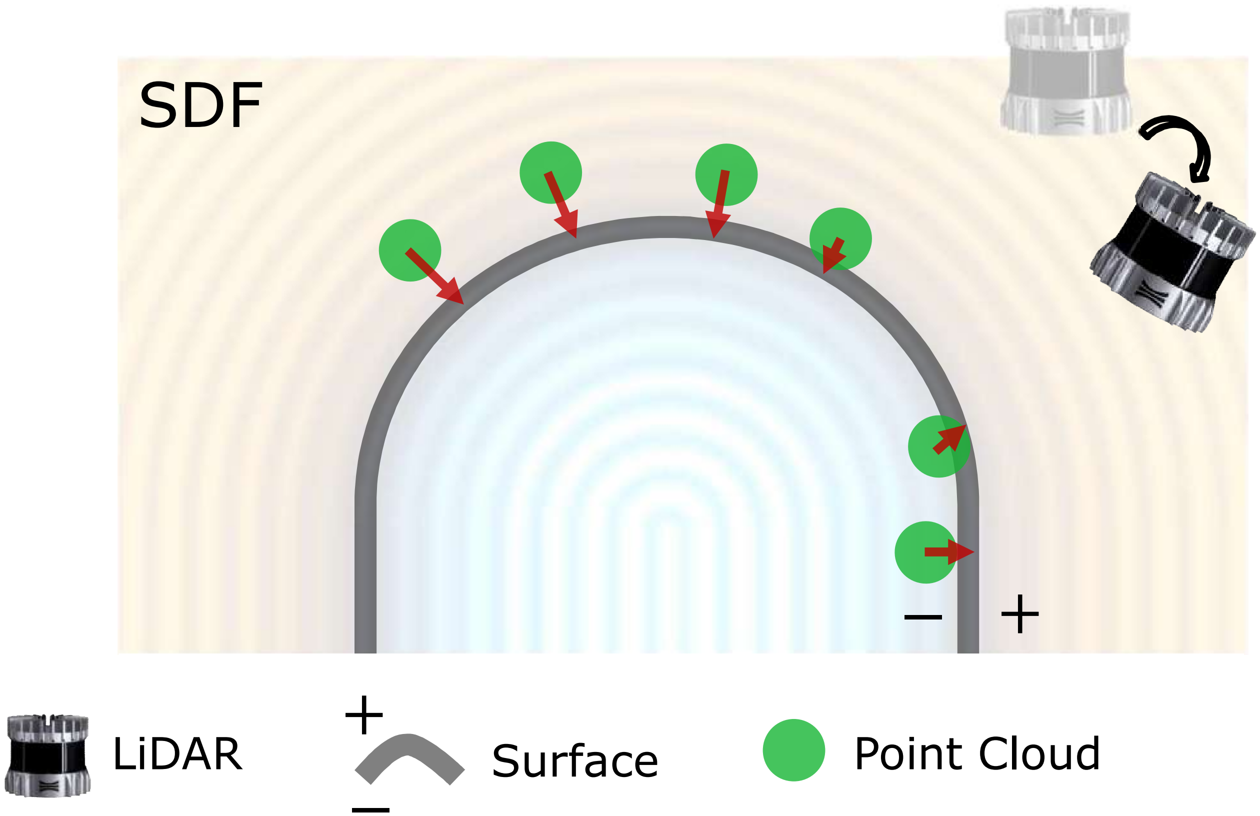

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Localization in Neural Distance Field



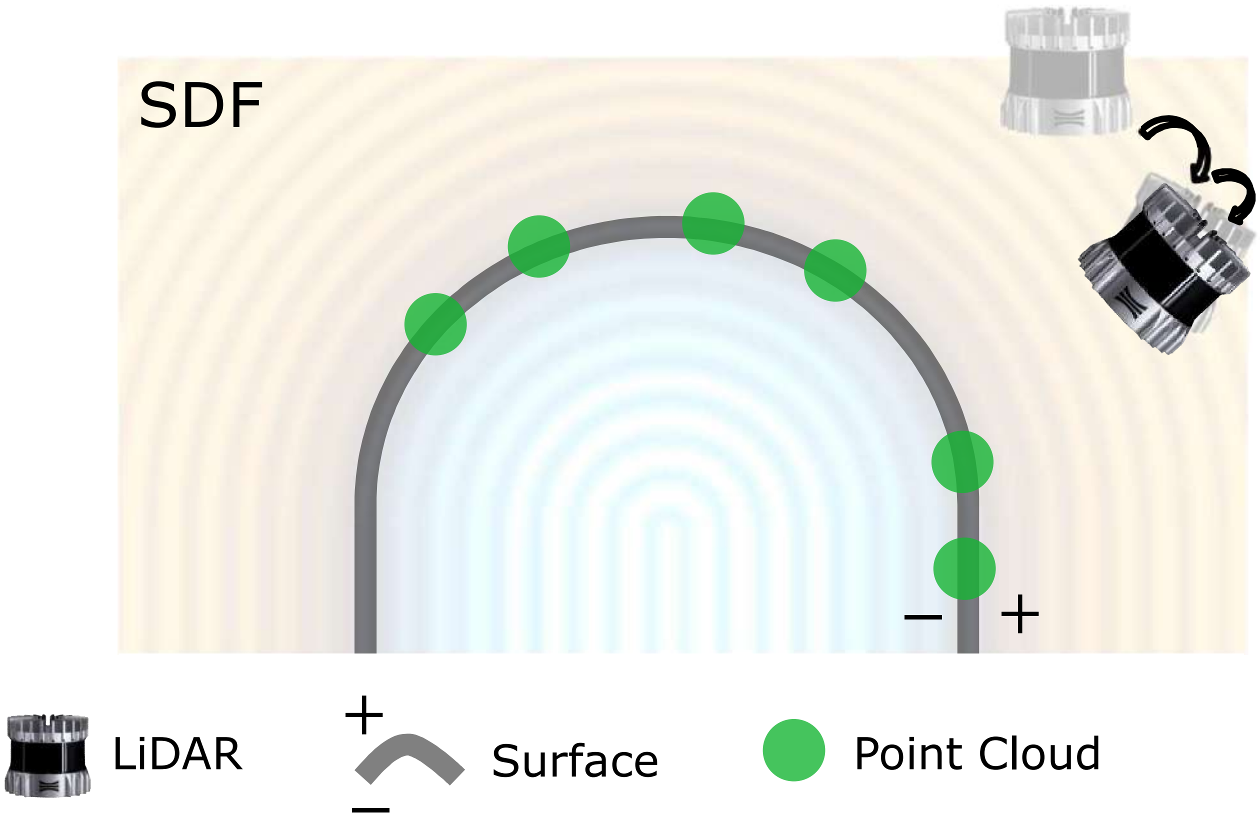

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Localization in Neural Distance Field



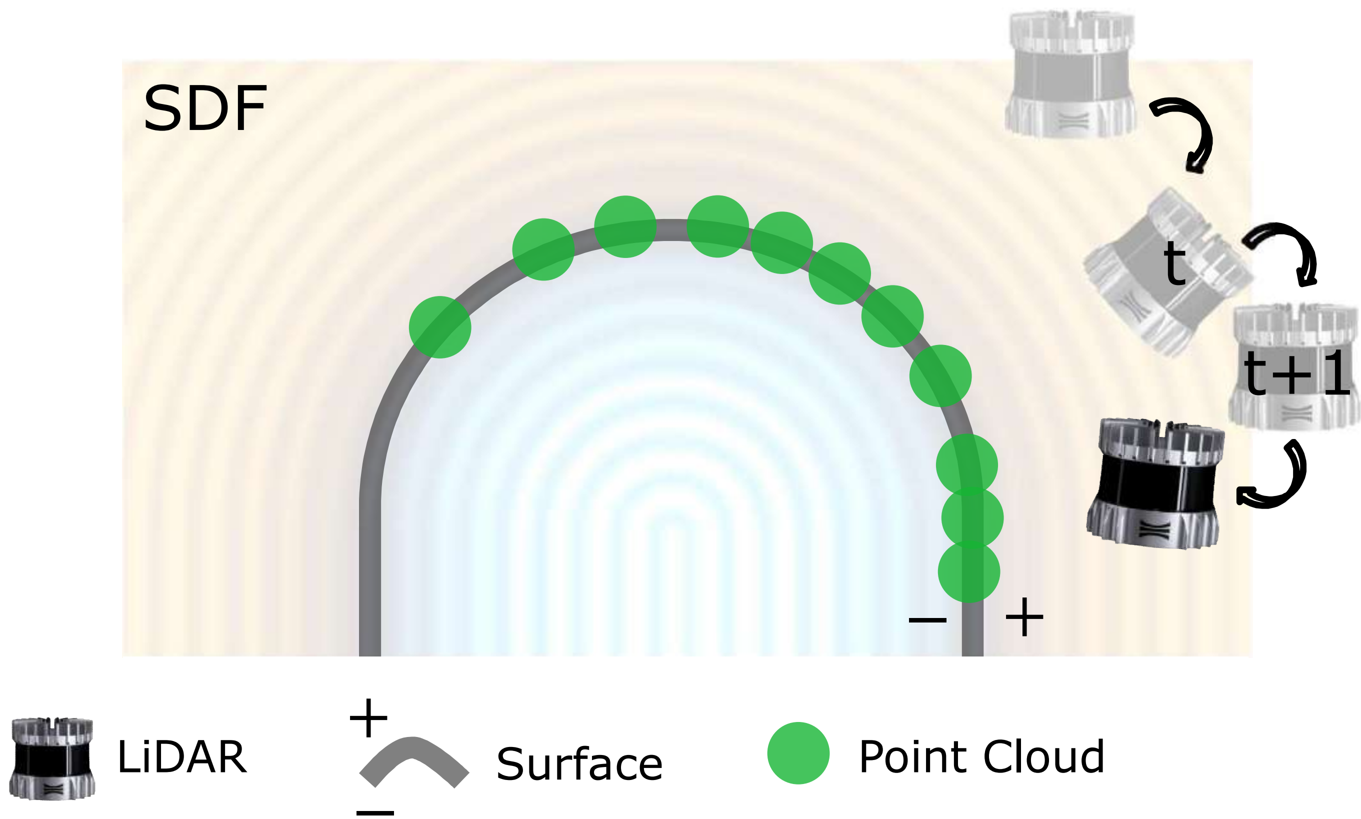

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Localization in Neural Distance Field



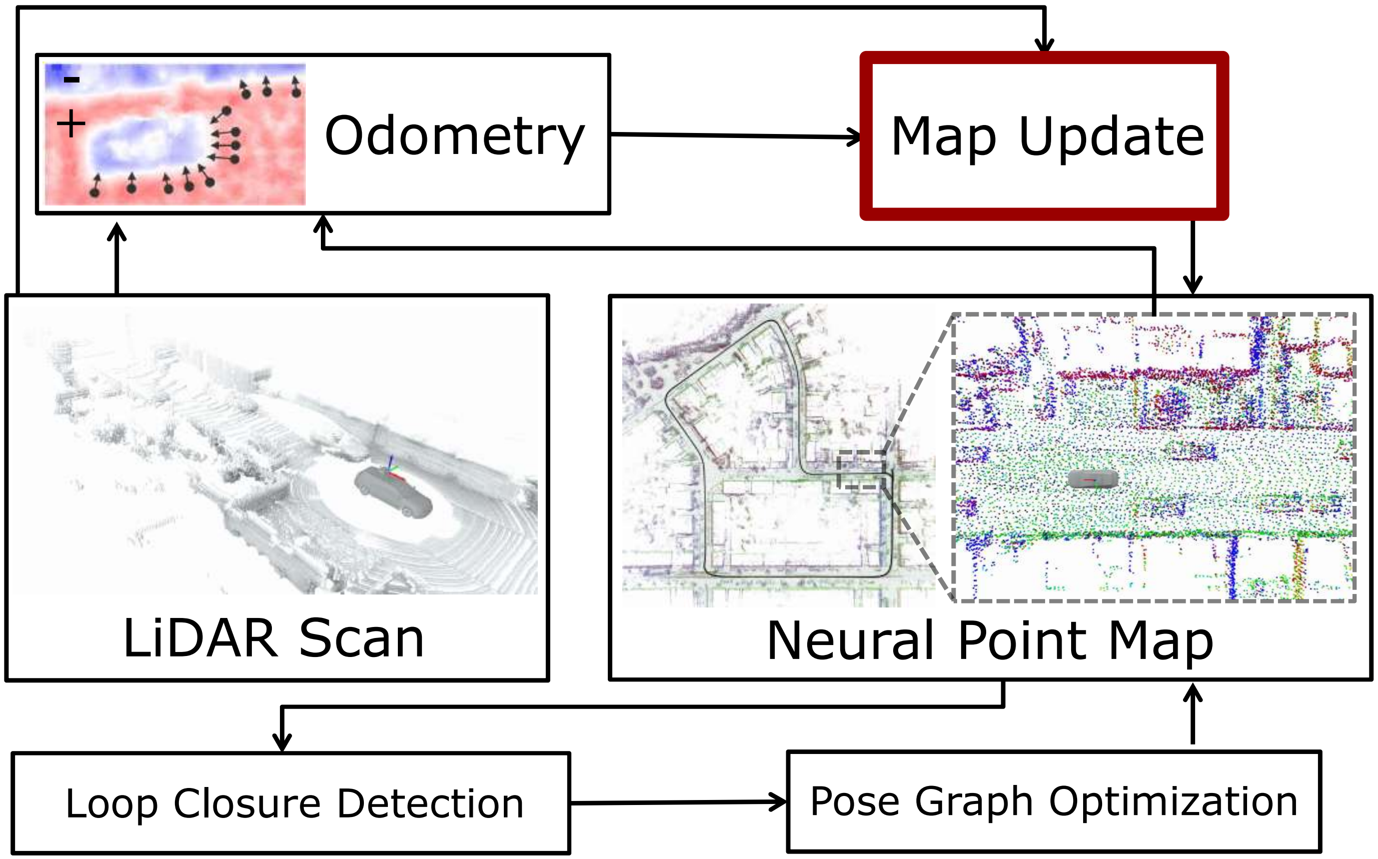

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# LiDAR Odometry Using Distance Field



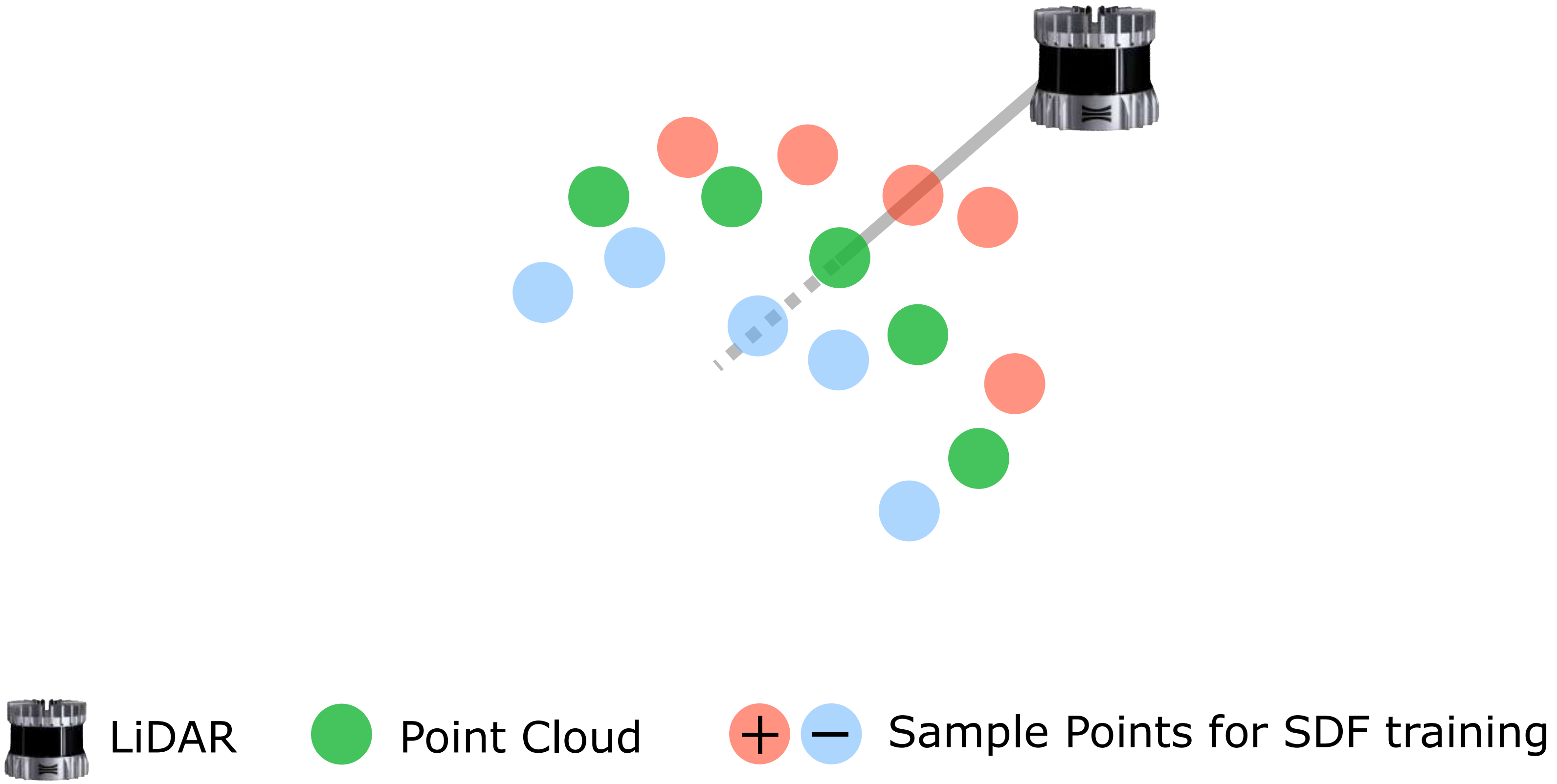

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# PIN-SLAM Pipeline



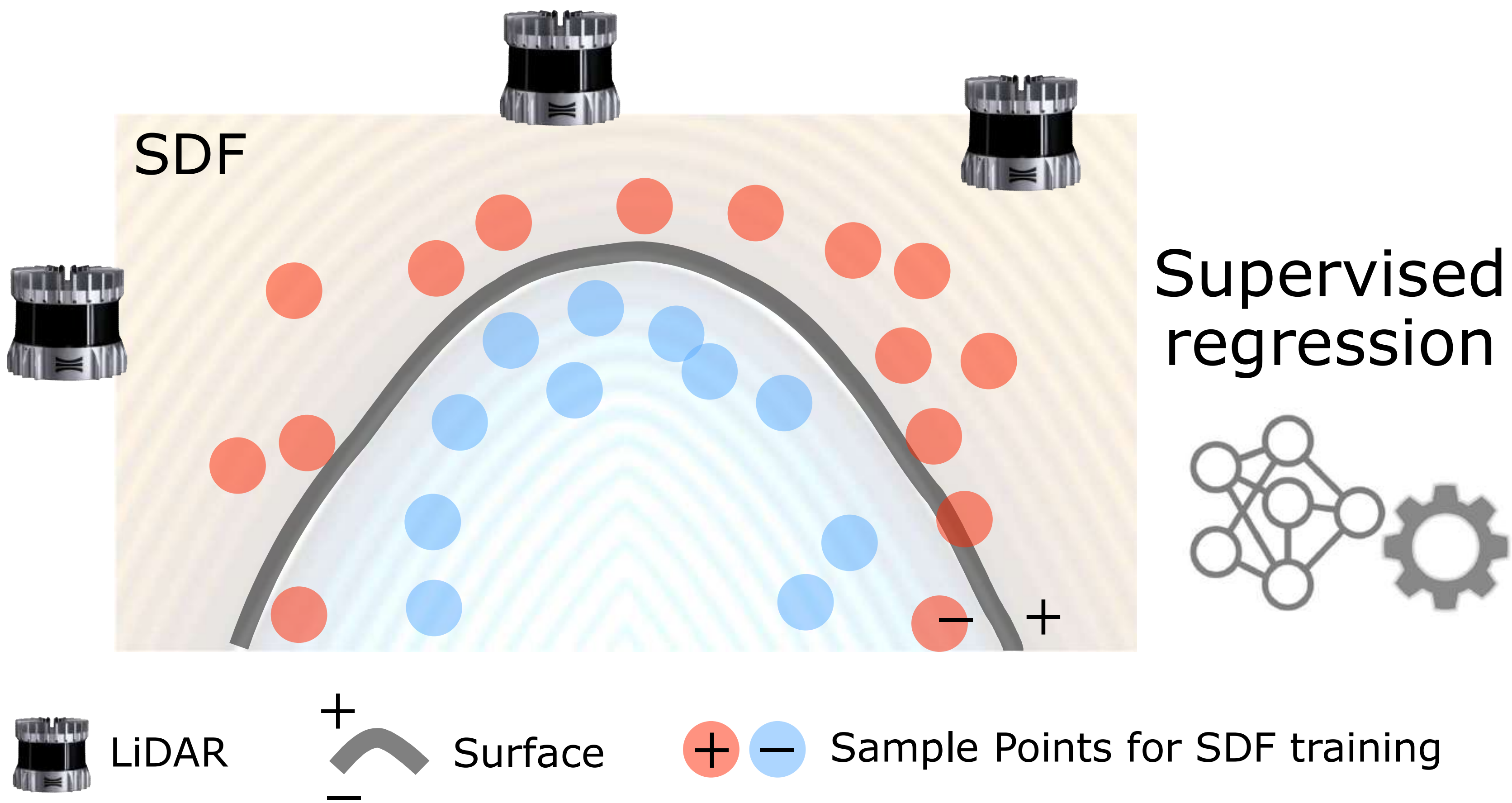

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Learning a Neural Distance Field



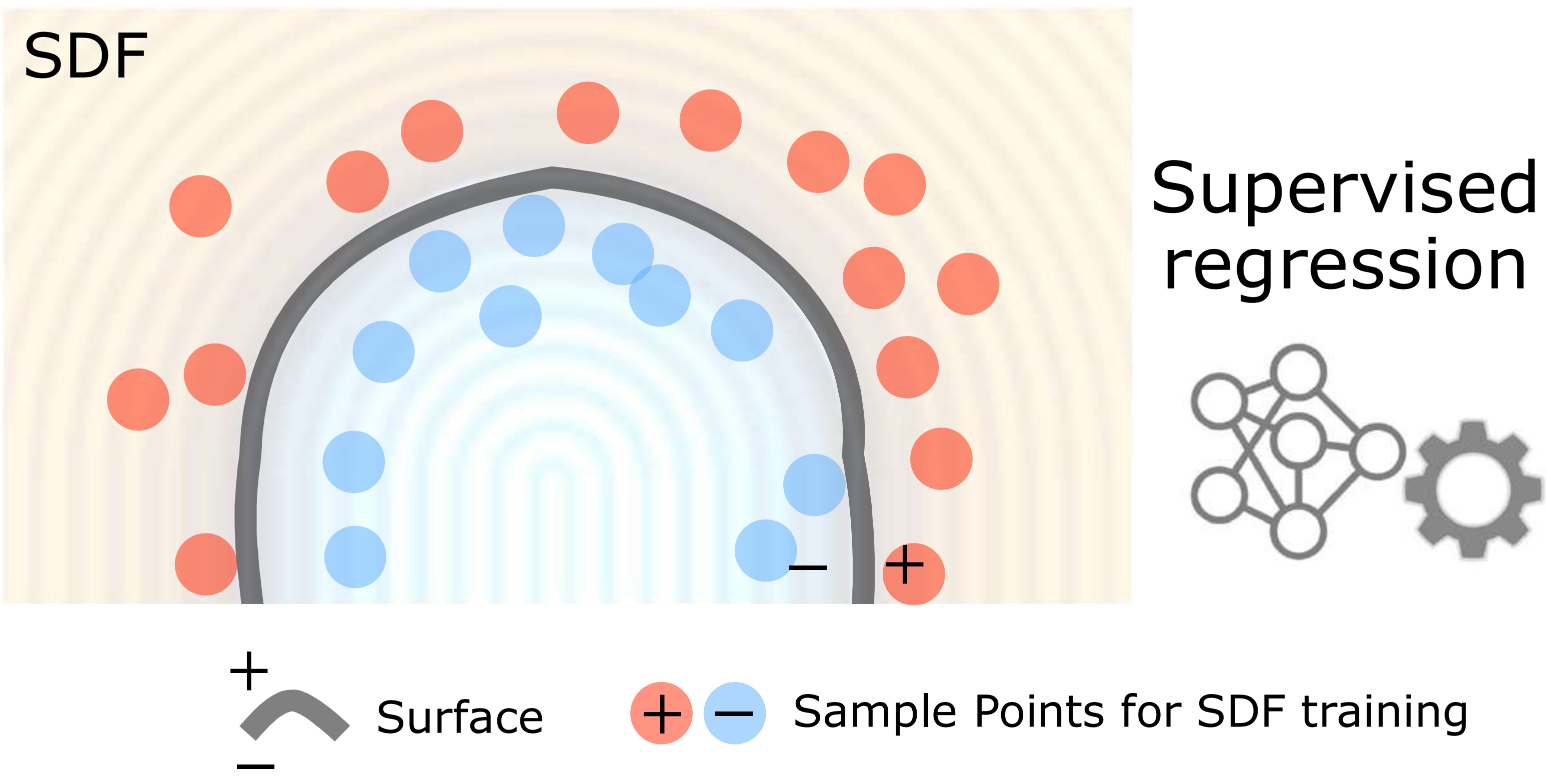

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Learning a Neural Distance Field



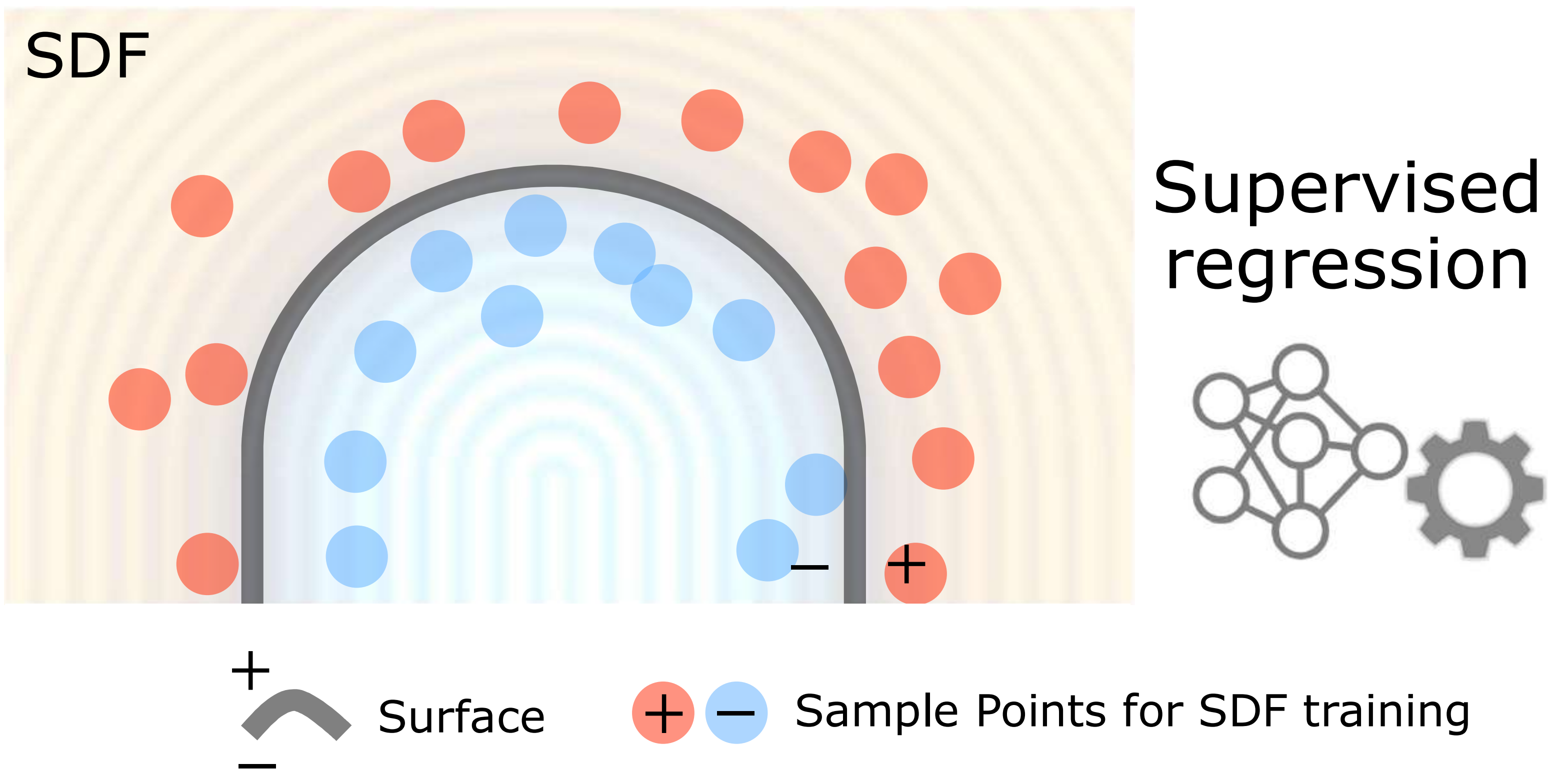

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Learning a Neural Distance Field



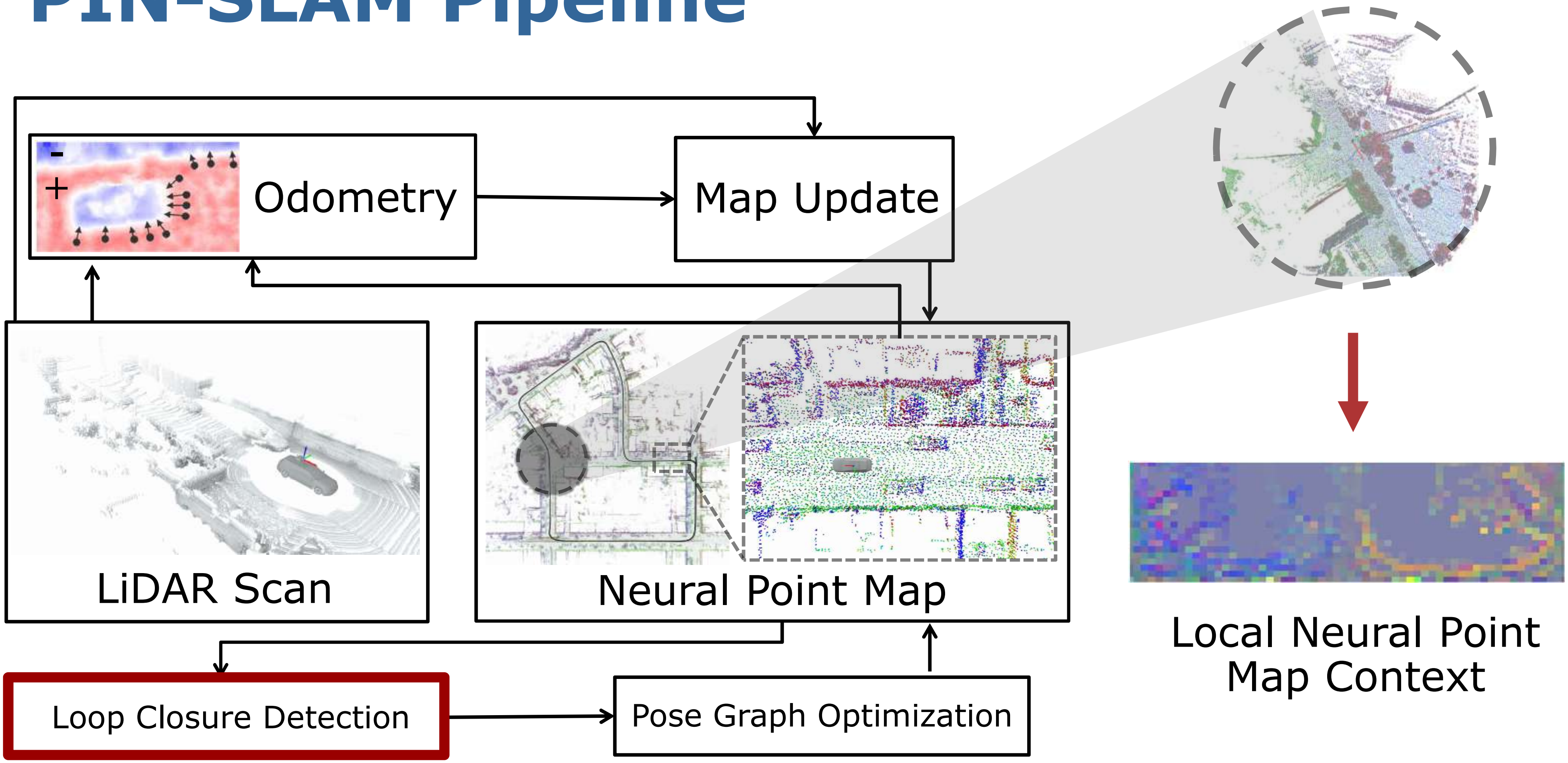

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# Learning a Neural Distance Field



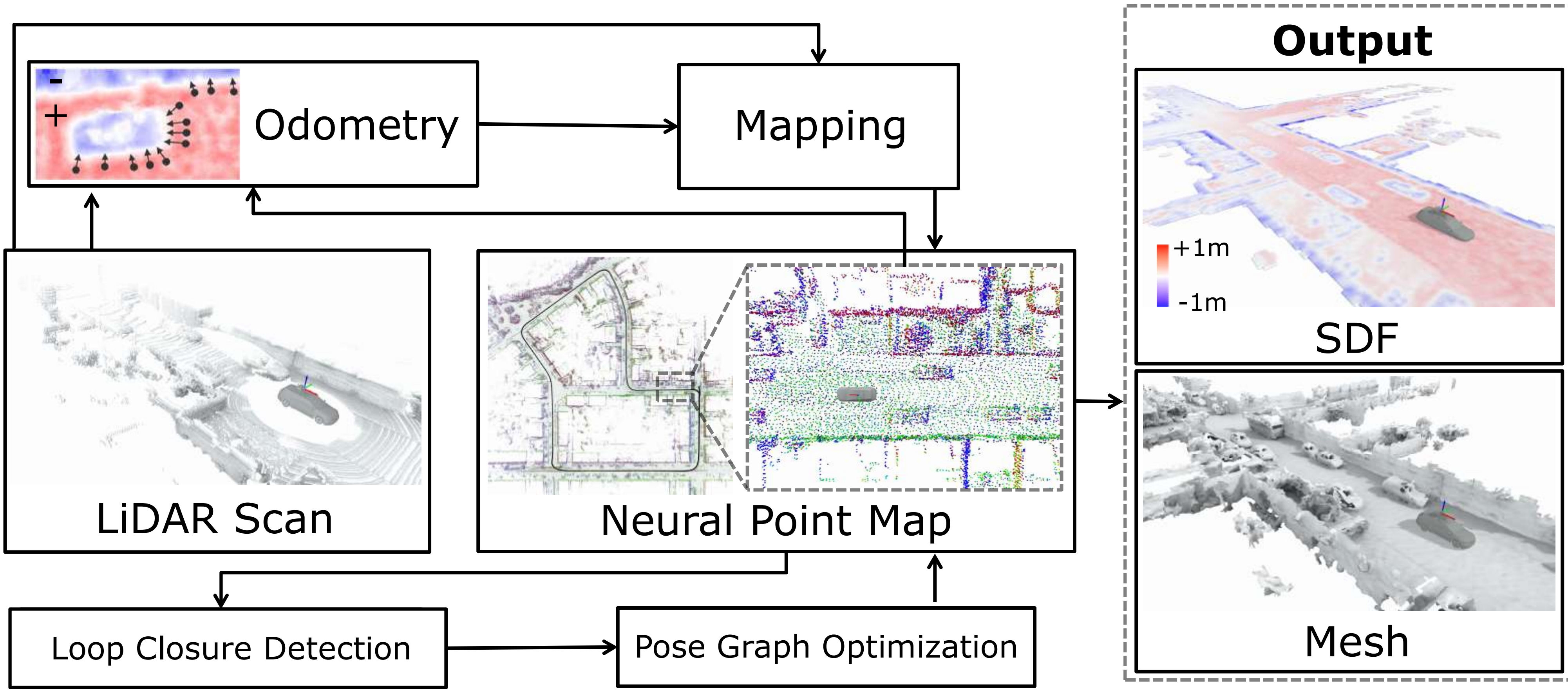

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# PIN-SLAM Pipeline



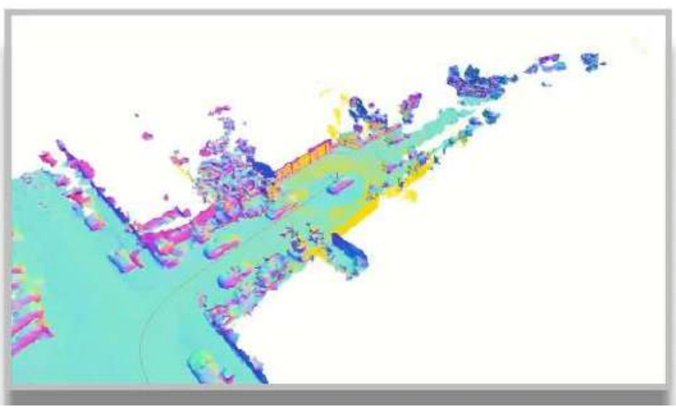

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

# PIN-SLAM Output



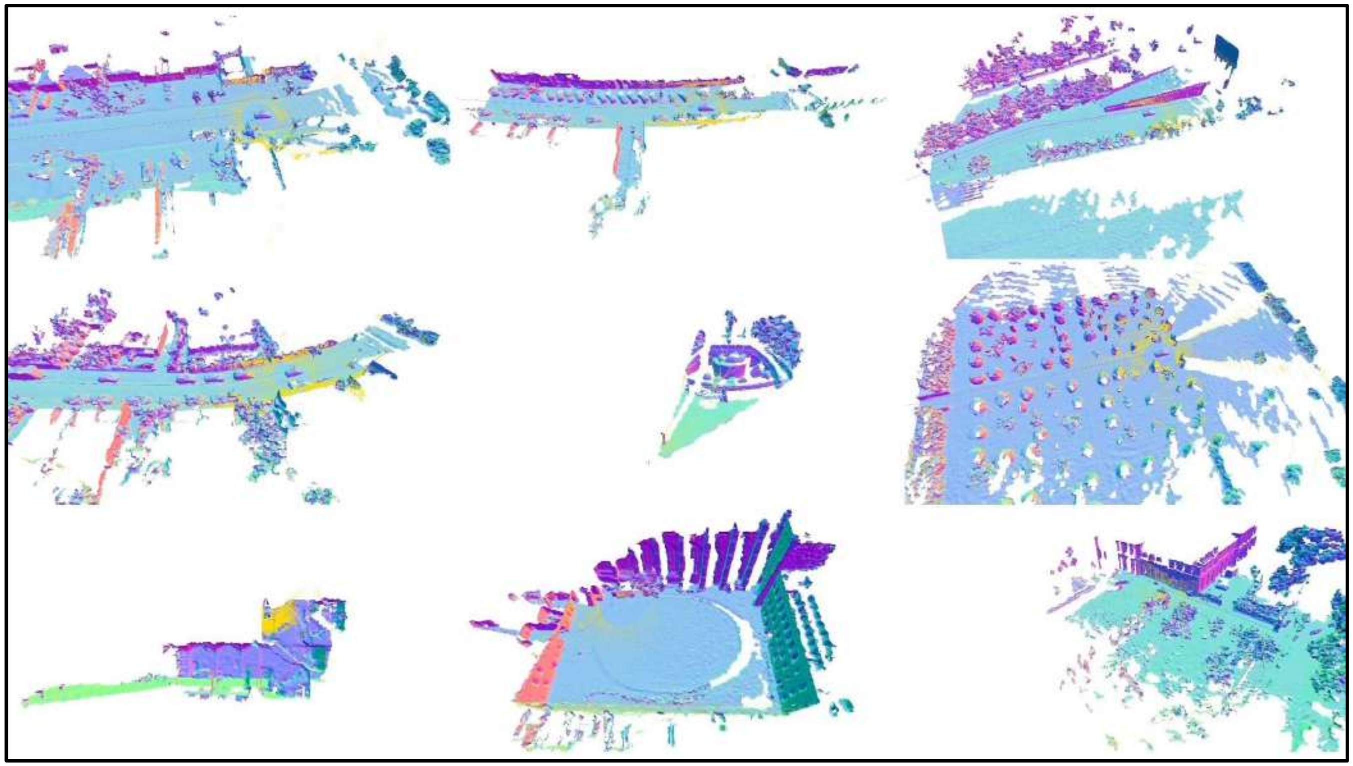

Pan et al. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. T-RO, 2024.

KITTI seq. 00

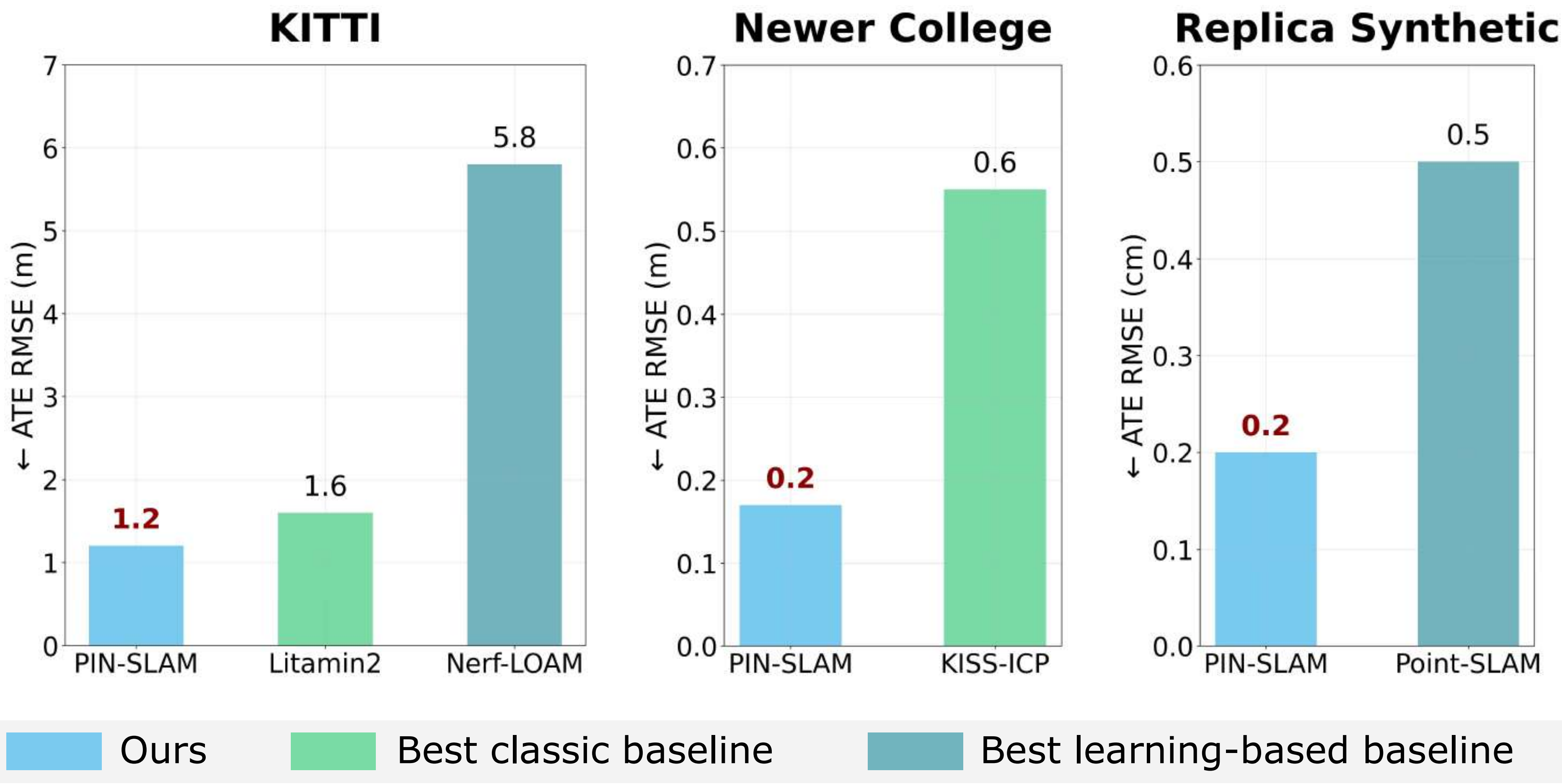


▶ 2×

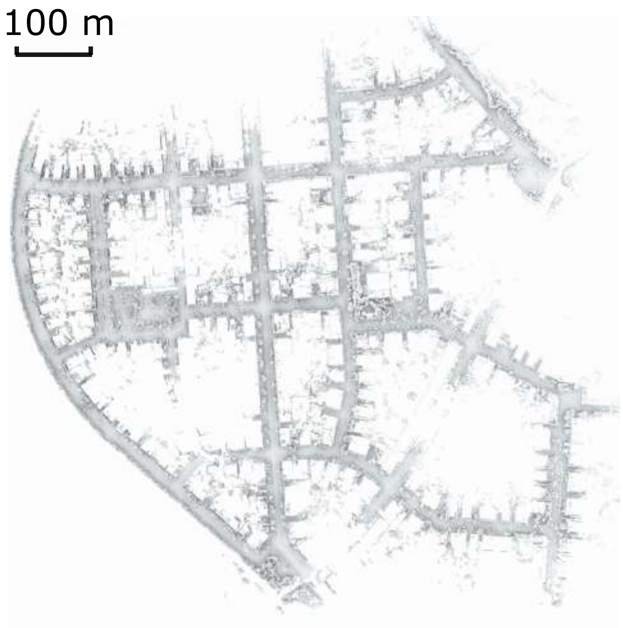
# PIN-SLAM Works with Various LiDARs

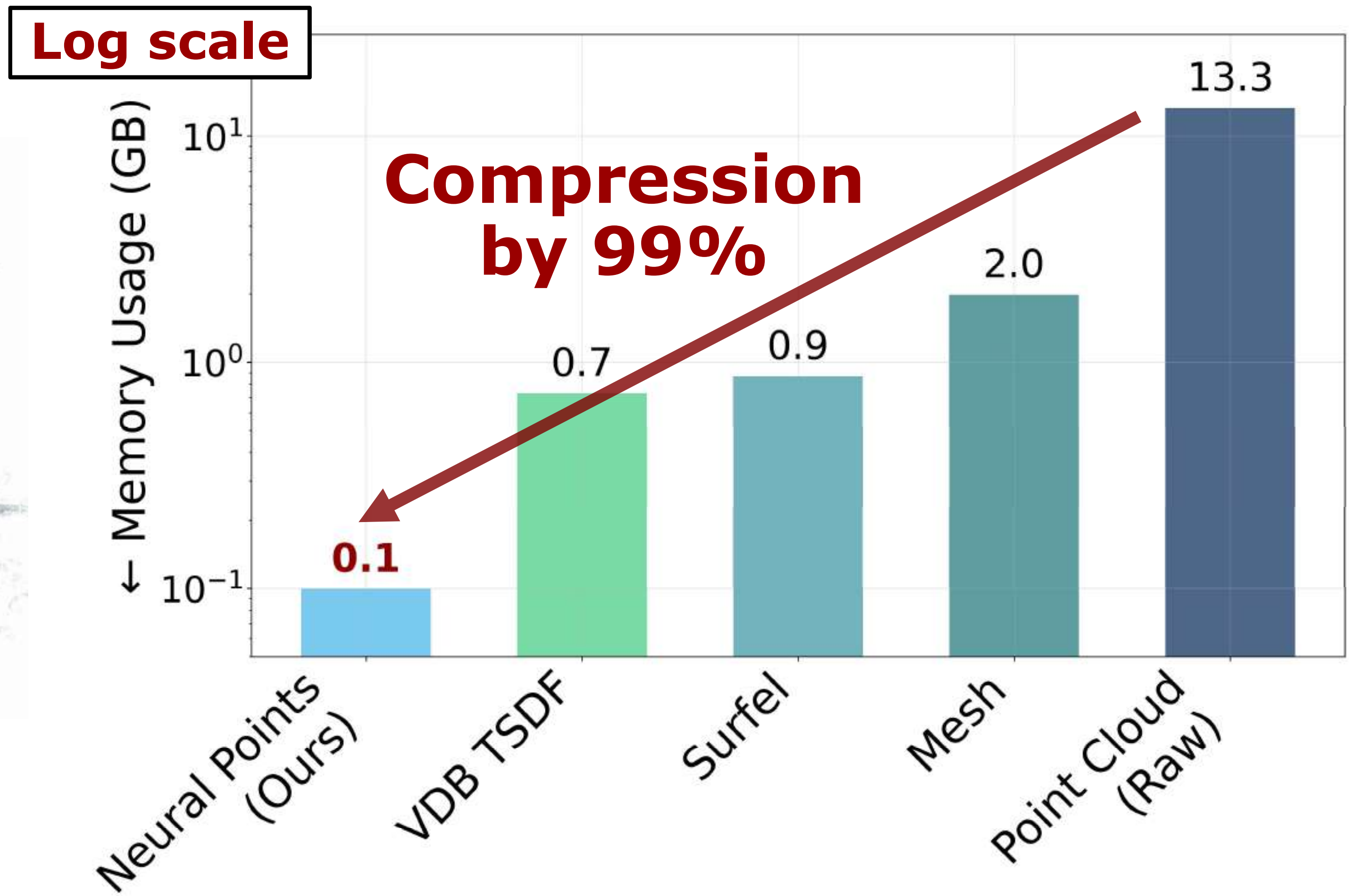
# PIN-SLAM Achieves Top Pose Accuracy



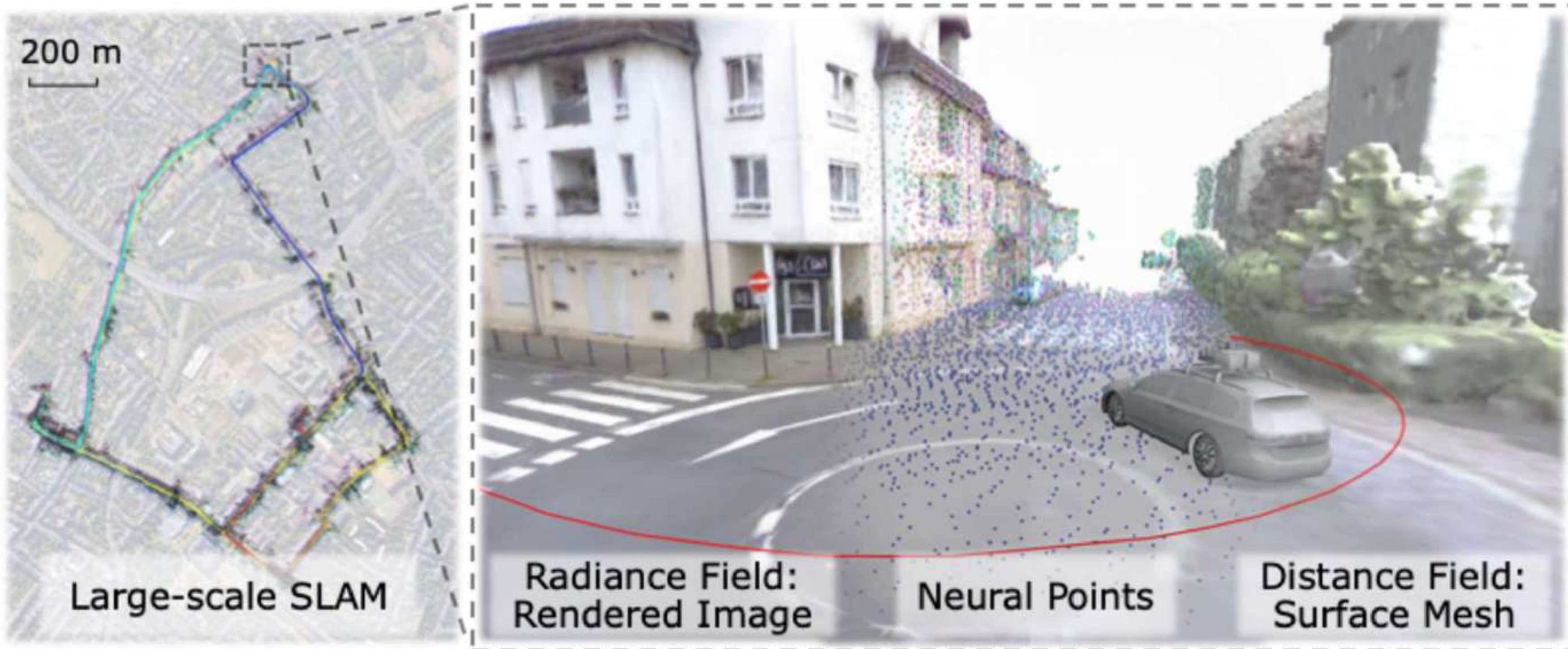
# Our Map is Compact




KITTI seq. 00



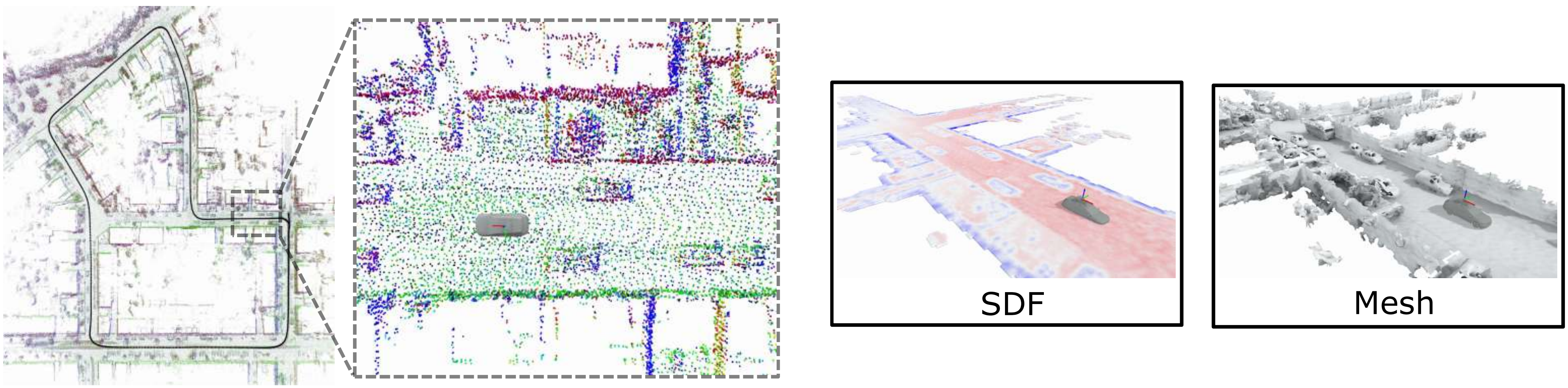
# Neural Points with Gaussian Splatting




- Neural points quite flexible: also represent Gaussian splats
- Allows to additionally represent visual appearance

Pan et al. PINGS: Gaussian Splatting Meets Distance Fields within a Point-Based Implicit Neural Map. RSS, 2025.

# Summary: Point Implicit Neural Maps




- Point-based map representation with per-point features
- Can be optimized to represent SDF and radiance field (color)
- Loop closures provide via pose graph optimization globally consistent trajectory

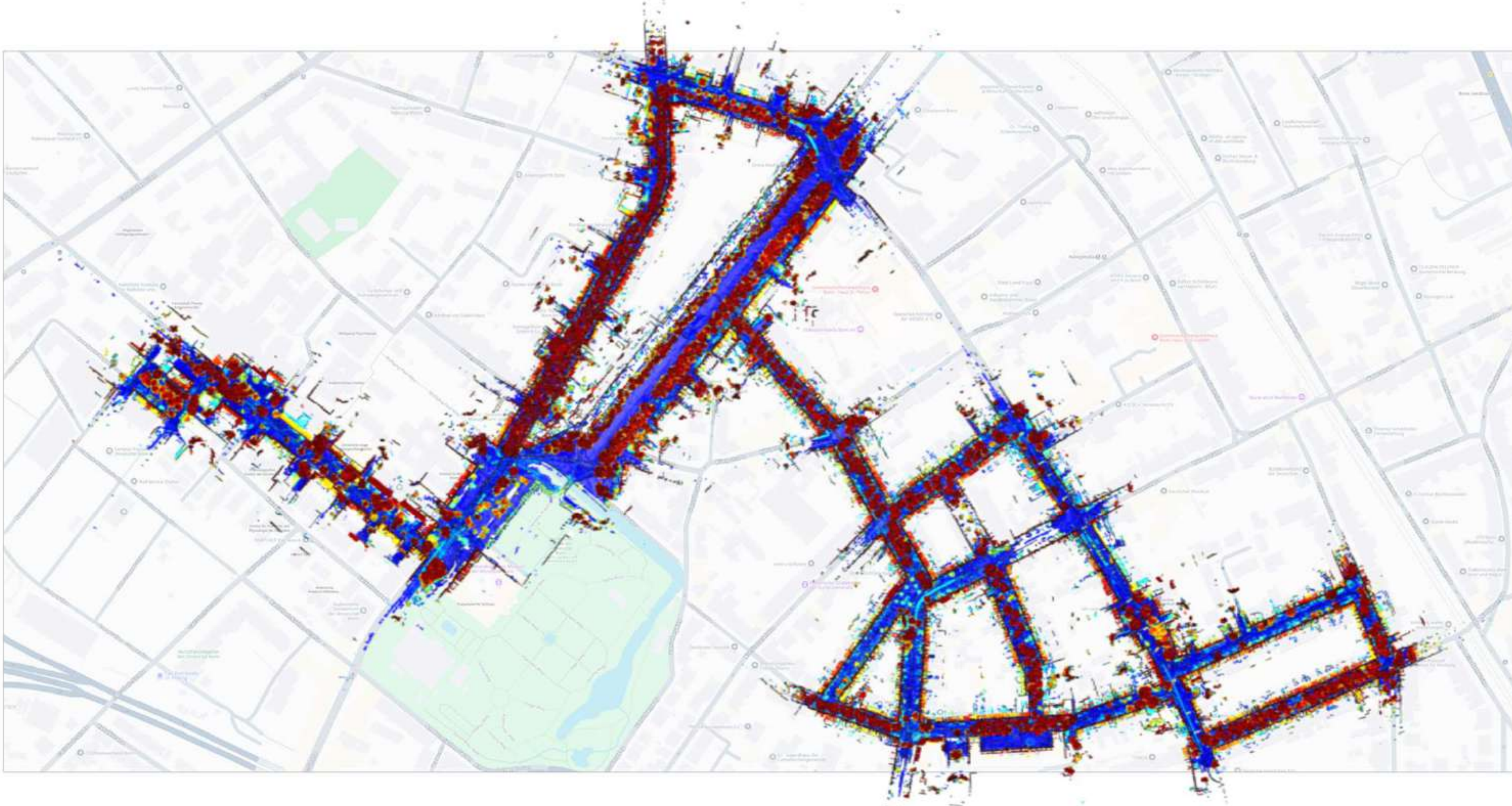
# Georeferenced Map

- LiDAR SLAM for local consistency
- GNSS for geo-referencing and global consistency

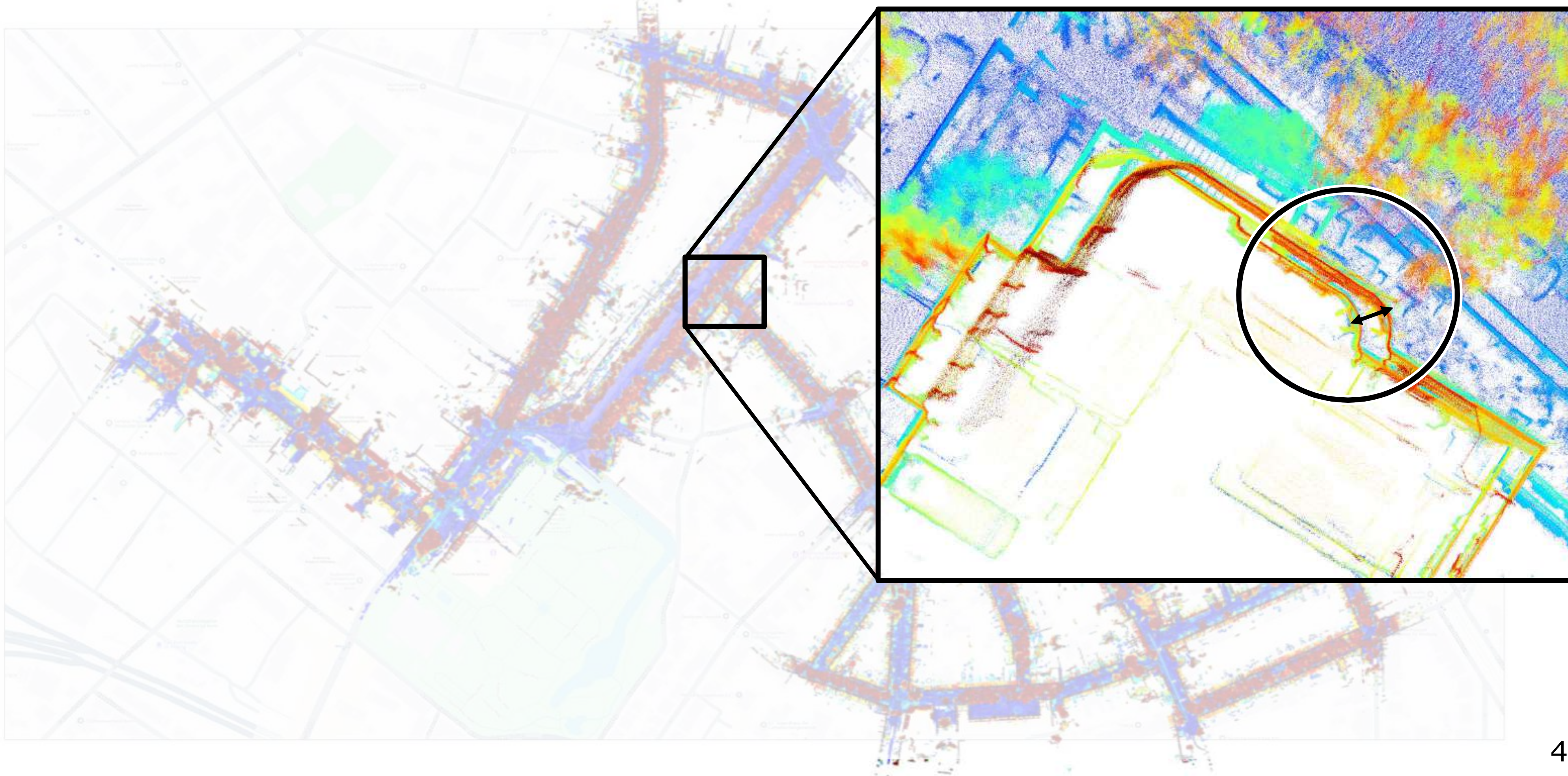
LiDAR poses

GNSS

# Standard: LiDAR SLAM + GNSS



# However: Local Misalignment



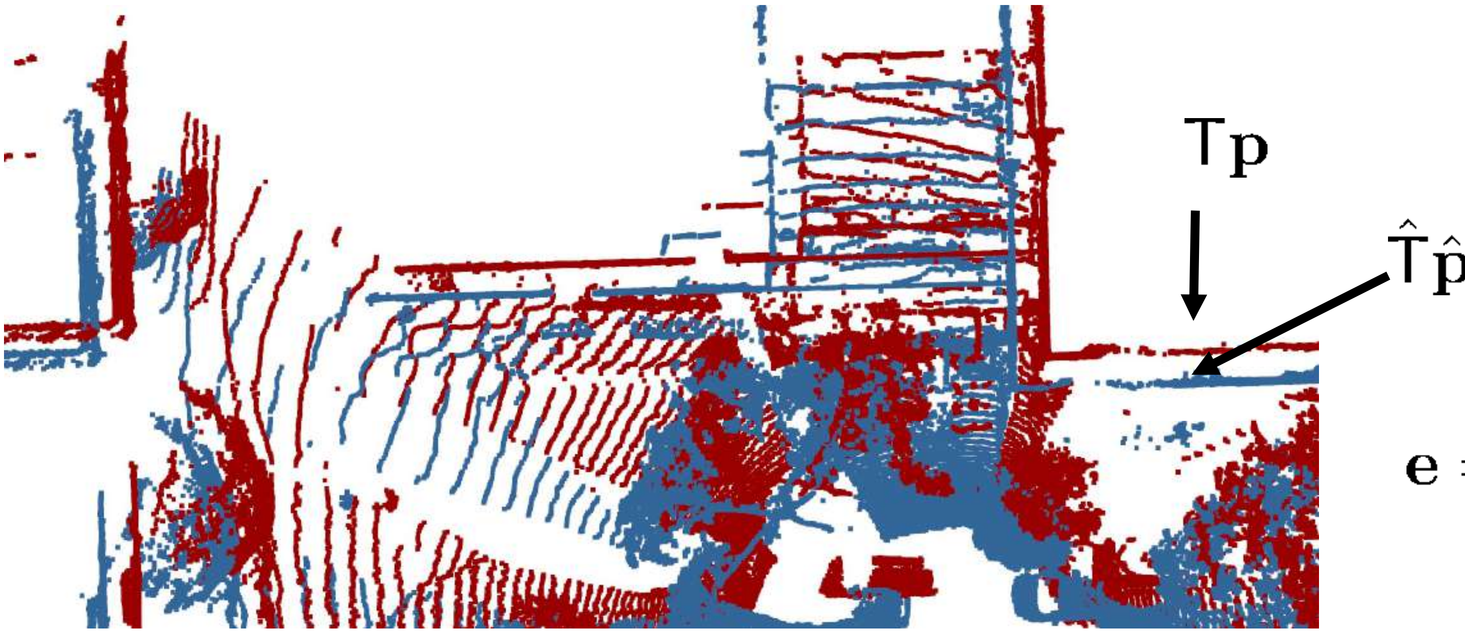
# LiDAR Bundle Adjustment

- Starting from LiDAR SLAM + GNSS estimate
- **Goal:** fix local misalignments
- **Idea:** align each scan to each other!

- **Problem:** Computationally challenging to jointly optimize all scan poses!

Wiesmann et al. Efficient LiDAR Bundle Adjustment for Multi-Scan Alignment Utilizing Continuous-Time Trajectories, arXiv, 2025.

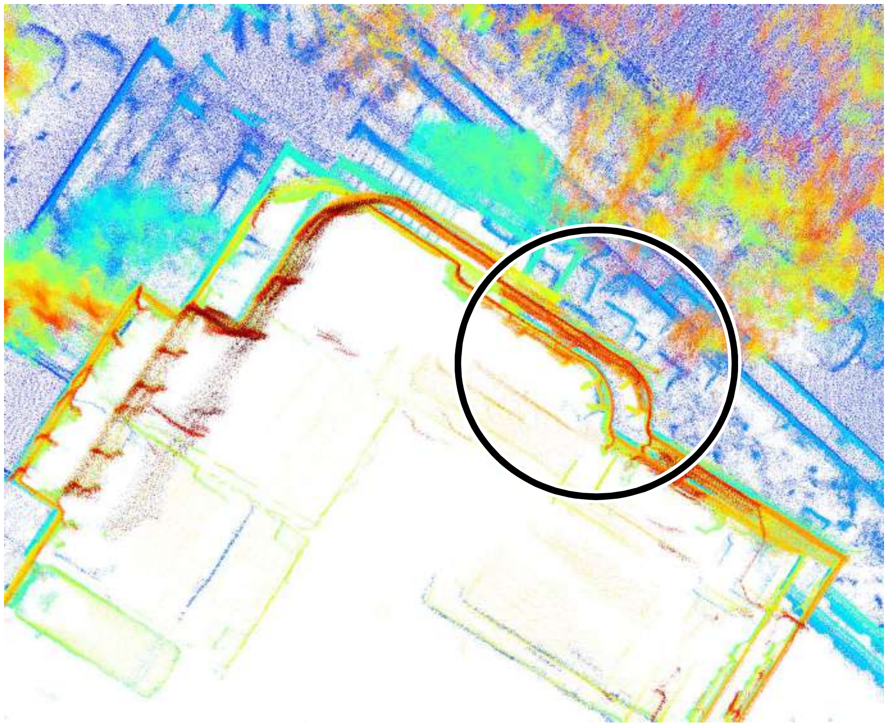
# Our Solution to Joint Alignment



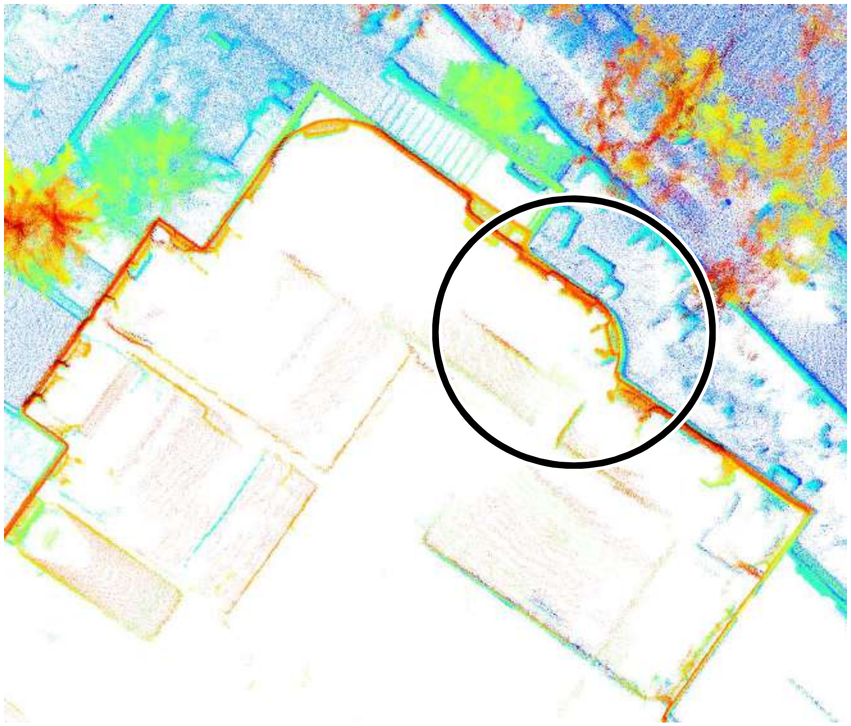

$$\mathbf{e} = \mathrm{T}\mathbf{p} - \hat{\mathrm{T}}\hat{\mathbf{p}}$$

- **Assumption:** Initialization near to **global optimum**
- Sample nearby LiDAR scans to make it tractable
- **Continuous** trajectory to account for moving LiDAR sensor

# Qualitative Results: Local



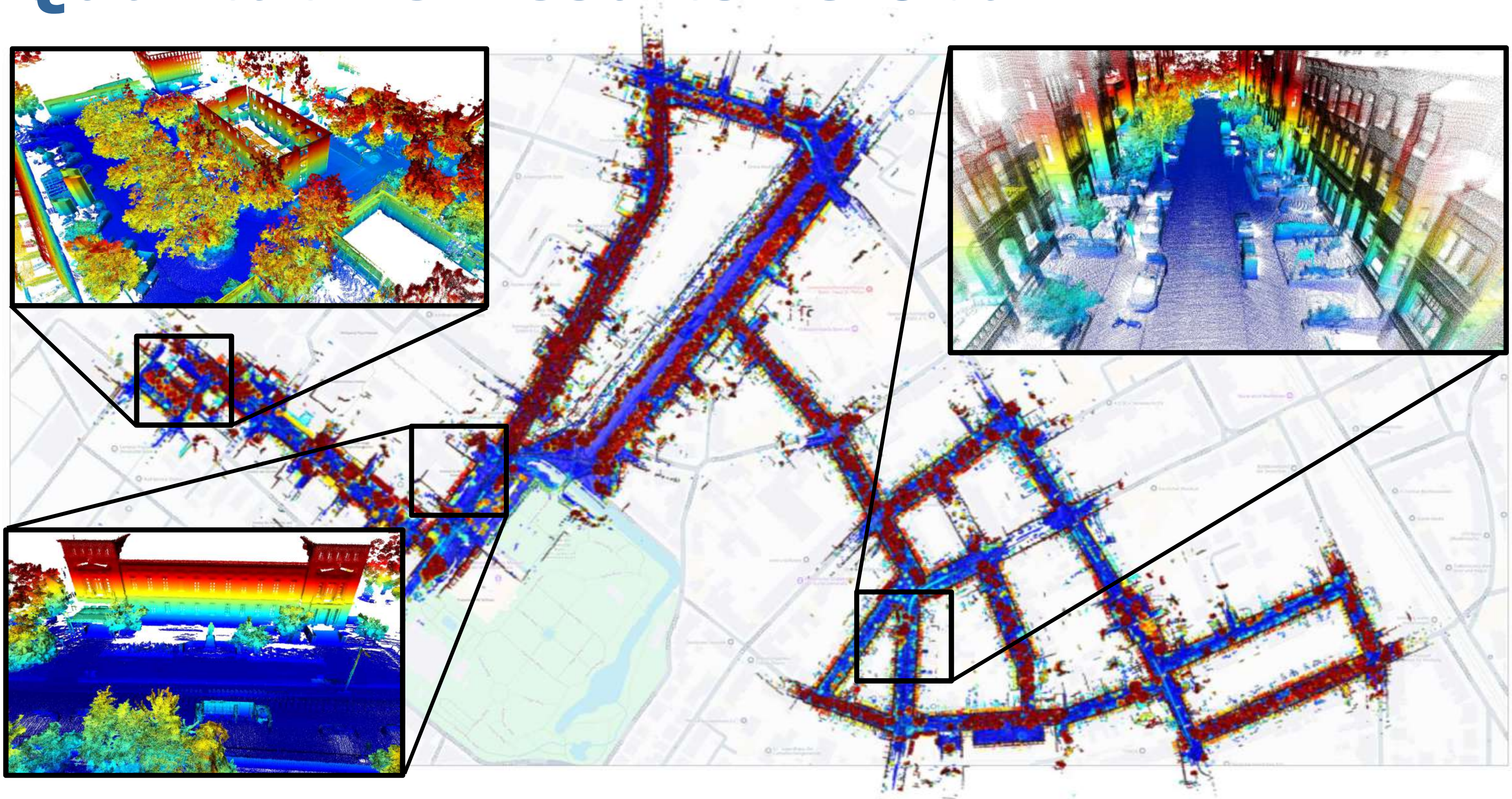
Initial Estimate



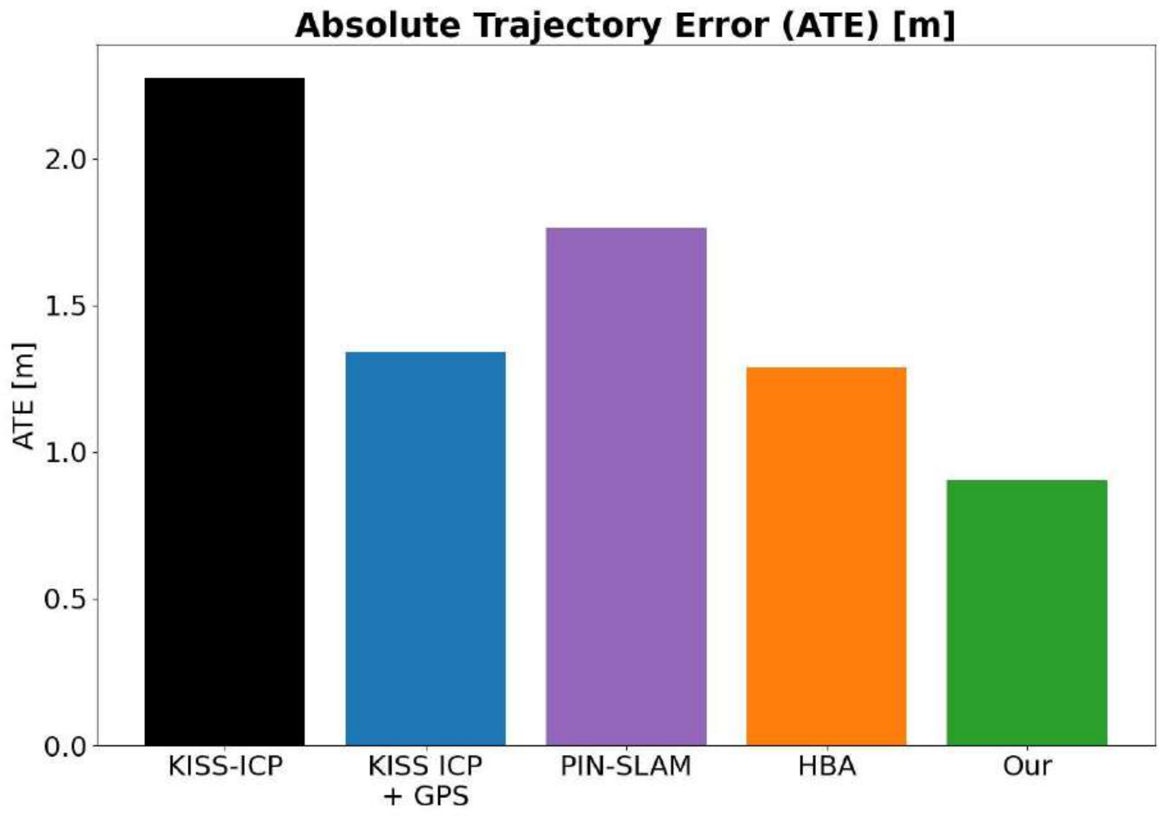
Our Refinement

Wiesmann et al. Efficient LiDAR Bundle Adjustment for Multi-Scan Alignment Utilizing Continuous-Time Trajectories, arXiv, 2025.

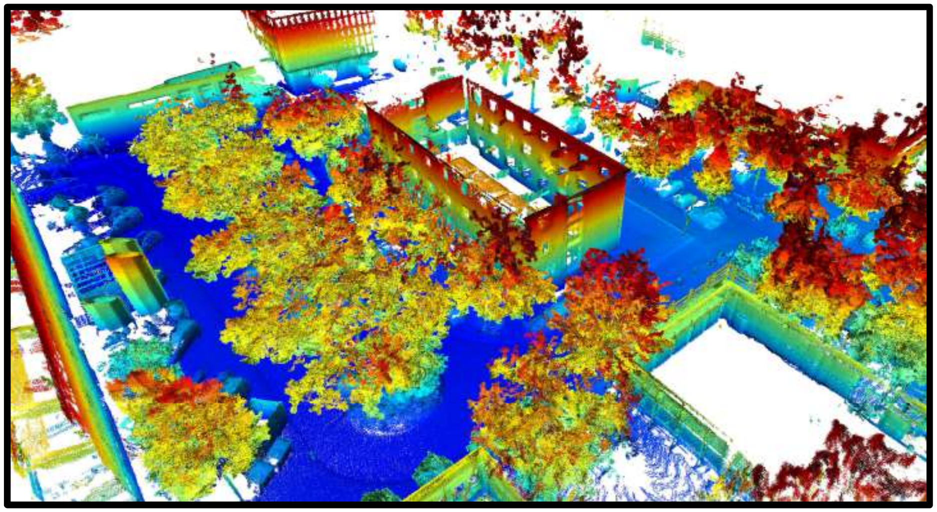
# Qualitative Results: Global

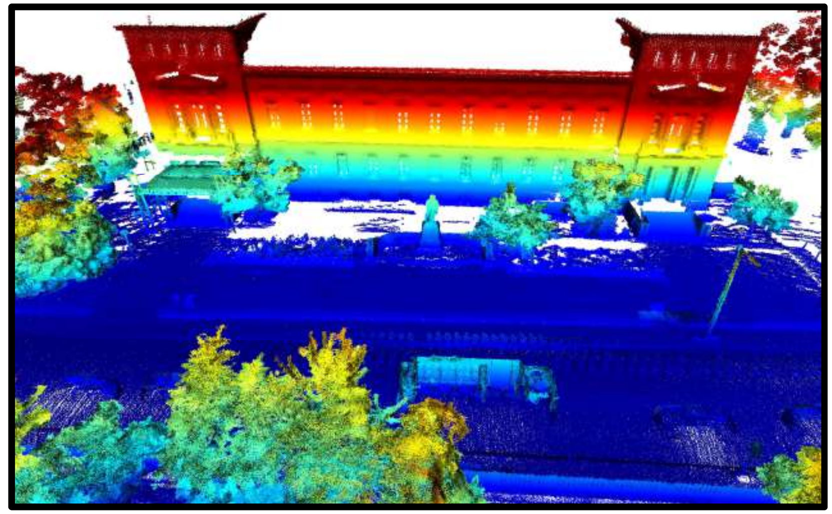


# Quantitative Results



# Summary: LiDAR Bundle Adjustment

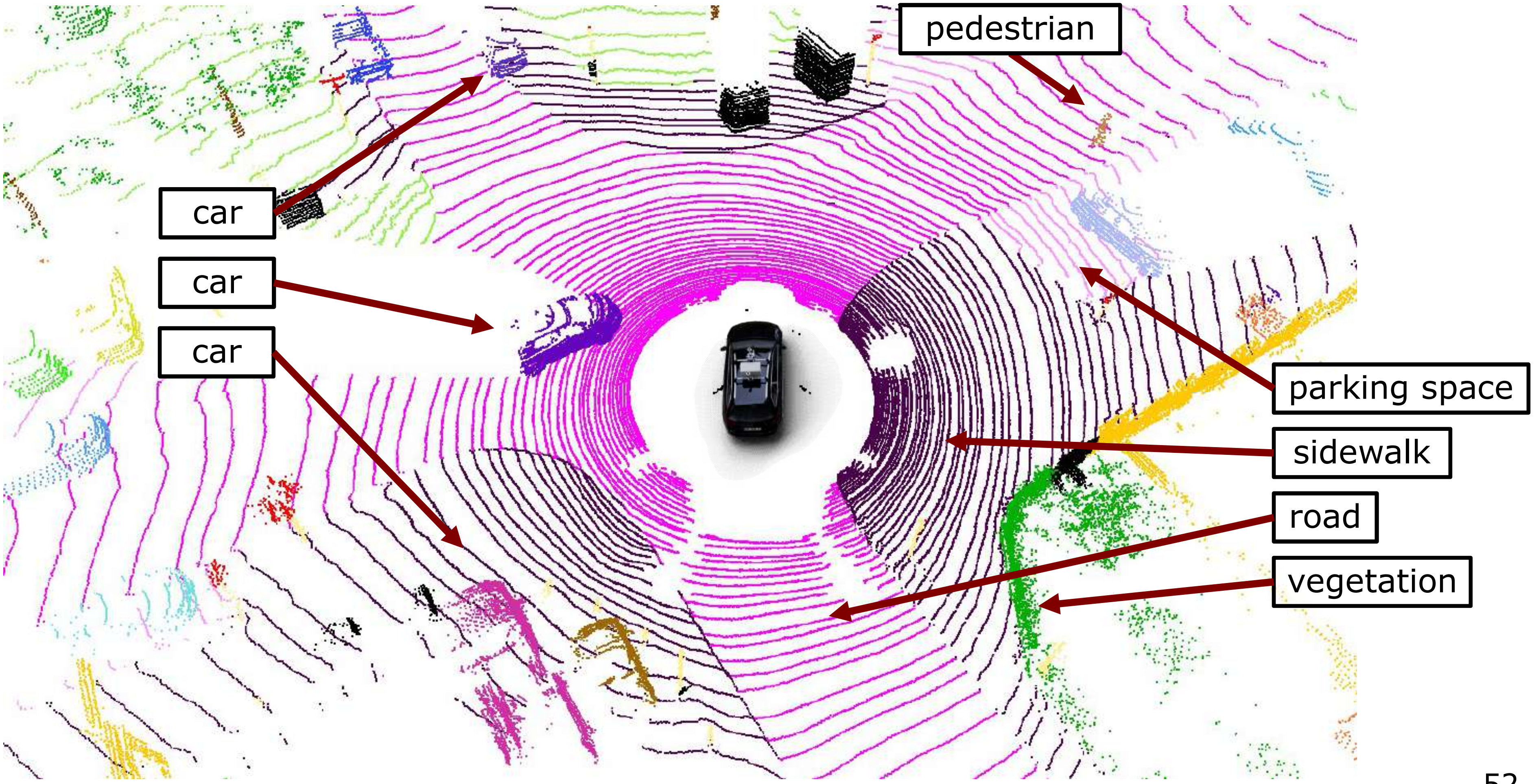




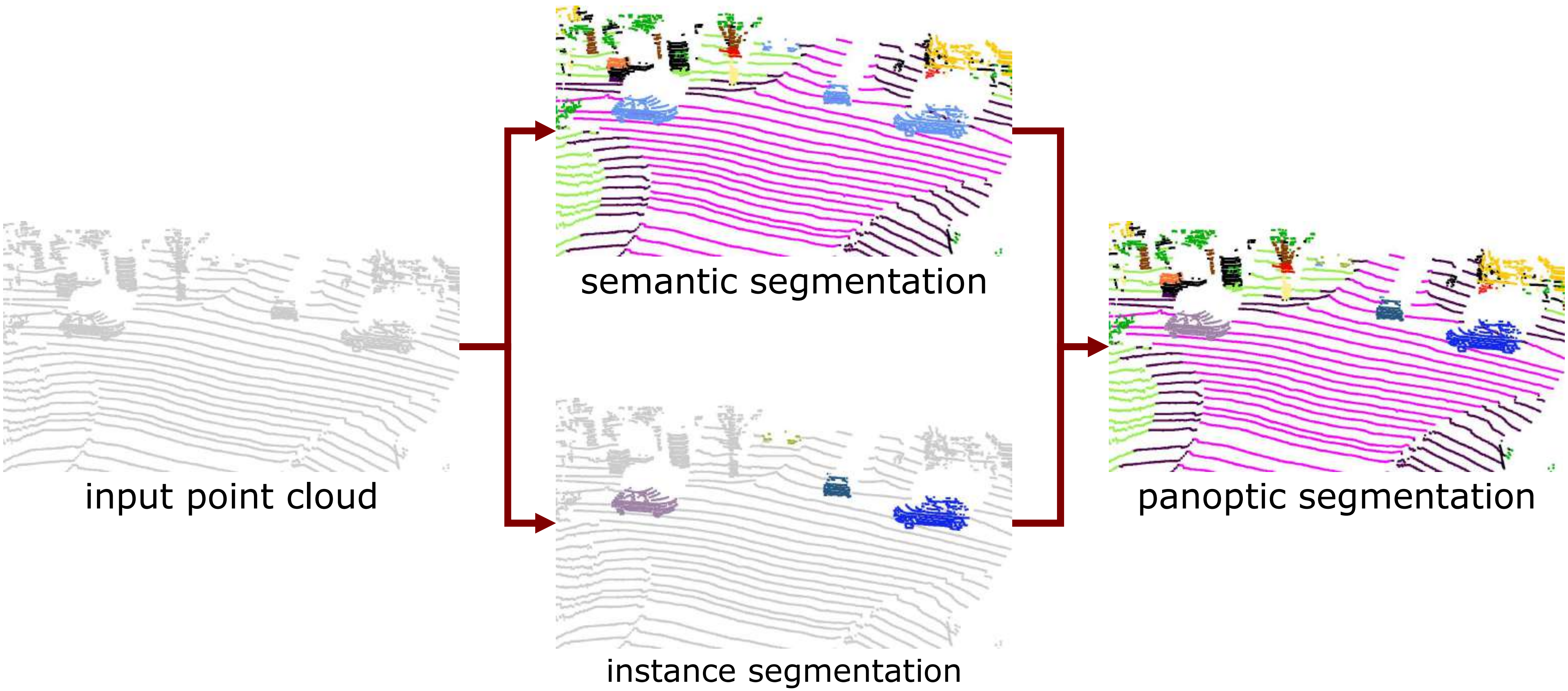
- Refine initial pose estimate
- **Joint alignment** of all scans
- Special: estimate **continuous trajectory**
- Allows **centimeter-accurate** mapping

# Understanding the Environment

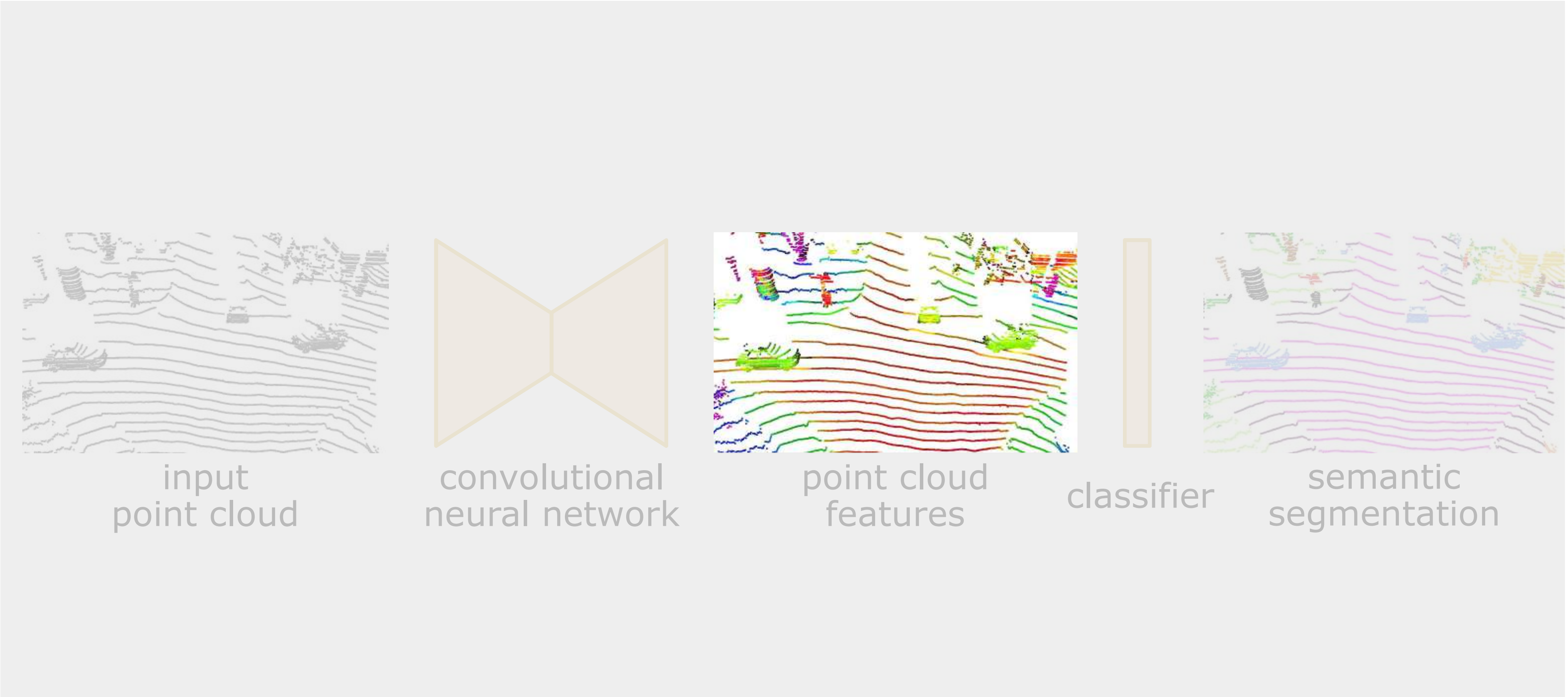
# Reminder: Semantic Interpretation



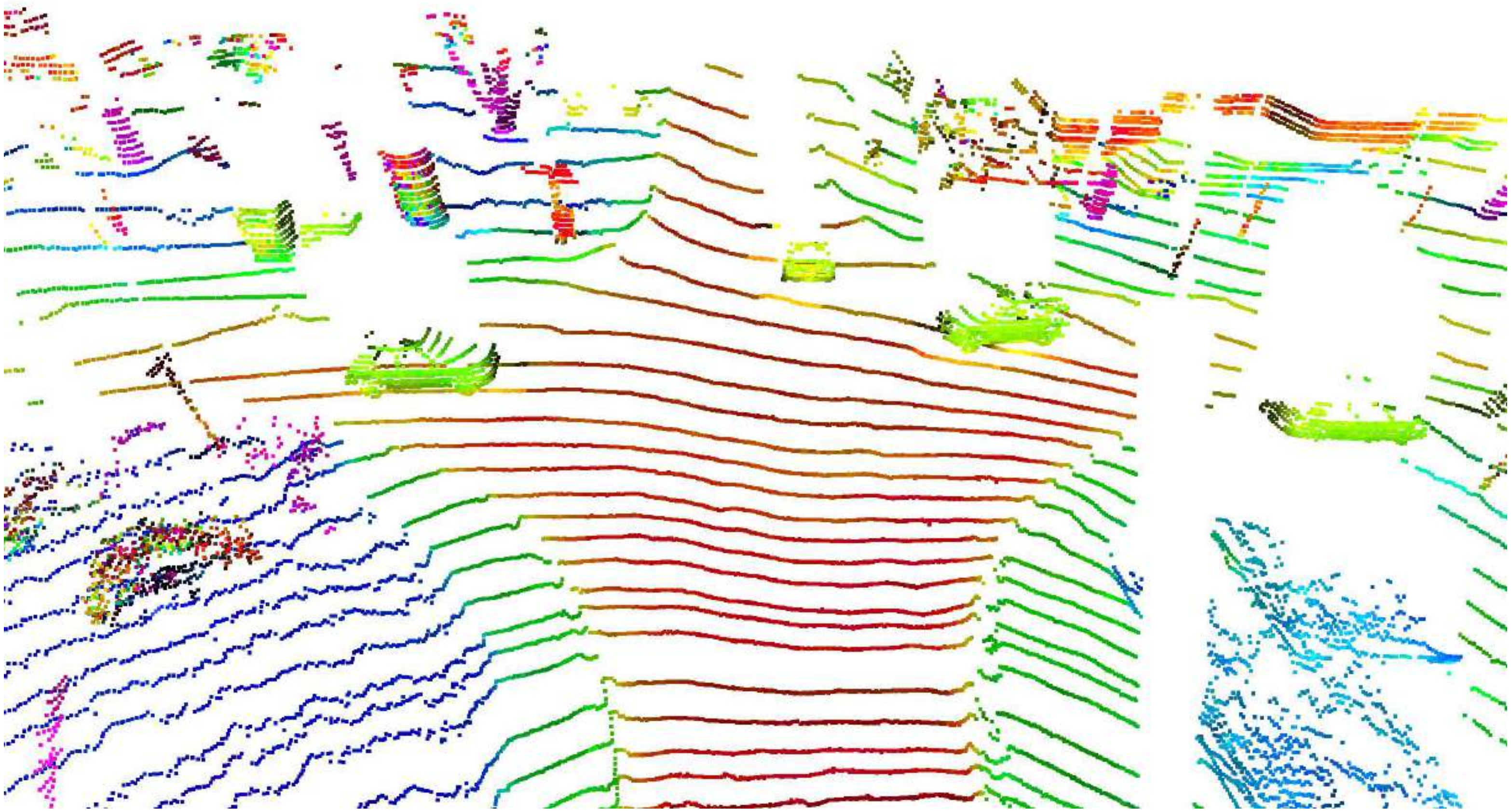
# Traditional Panoptic Segmentation



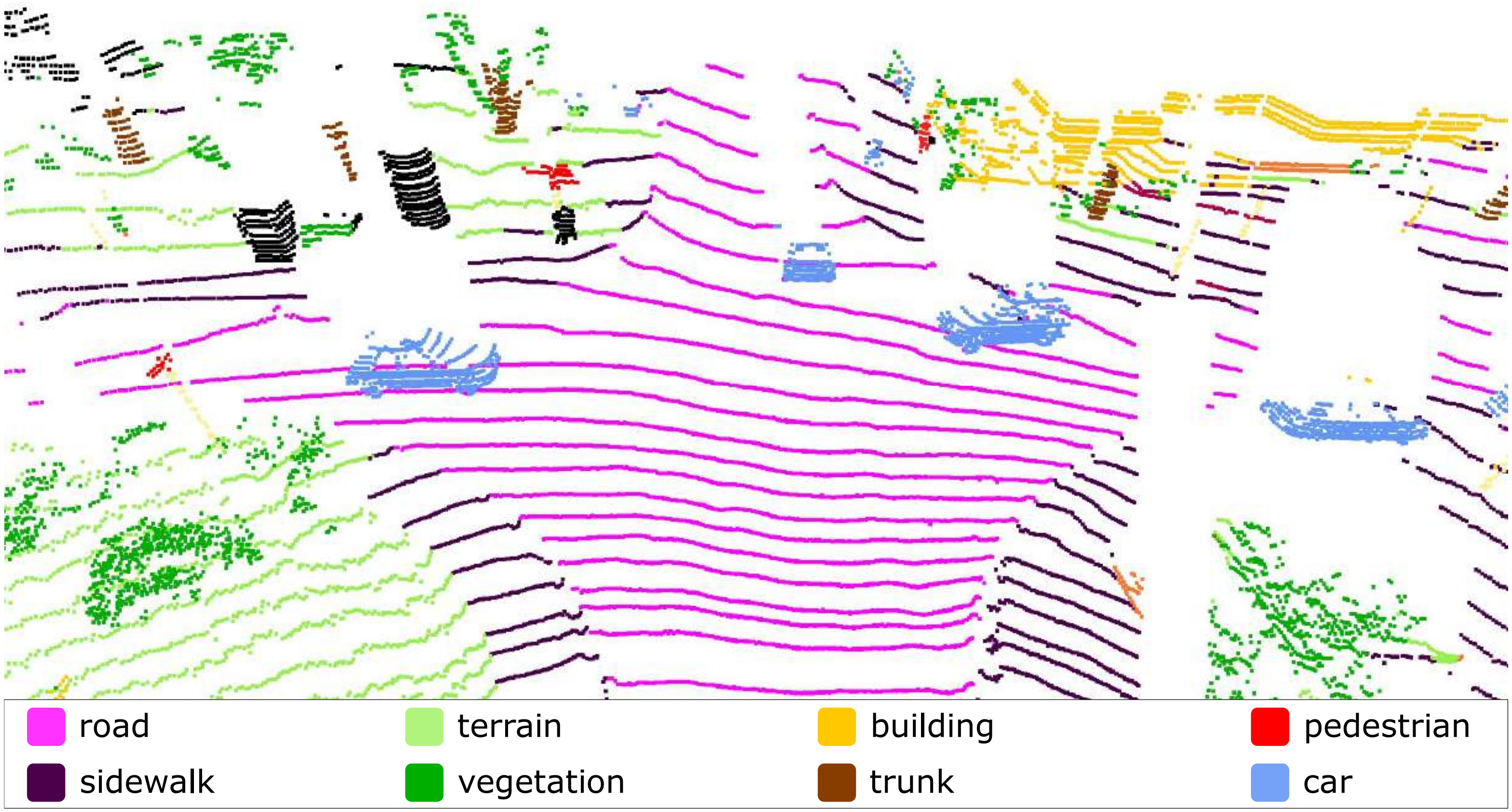
# Feature Extraction



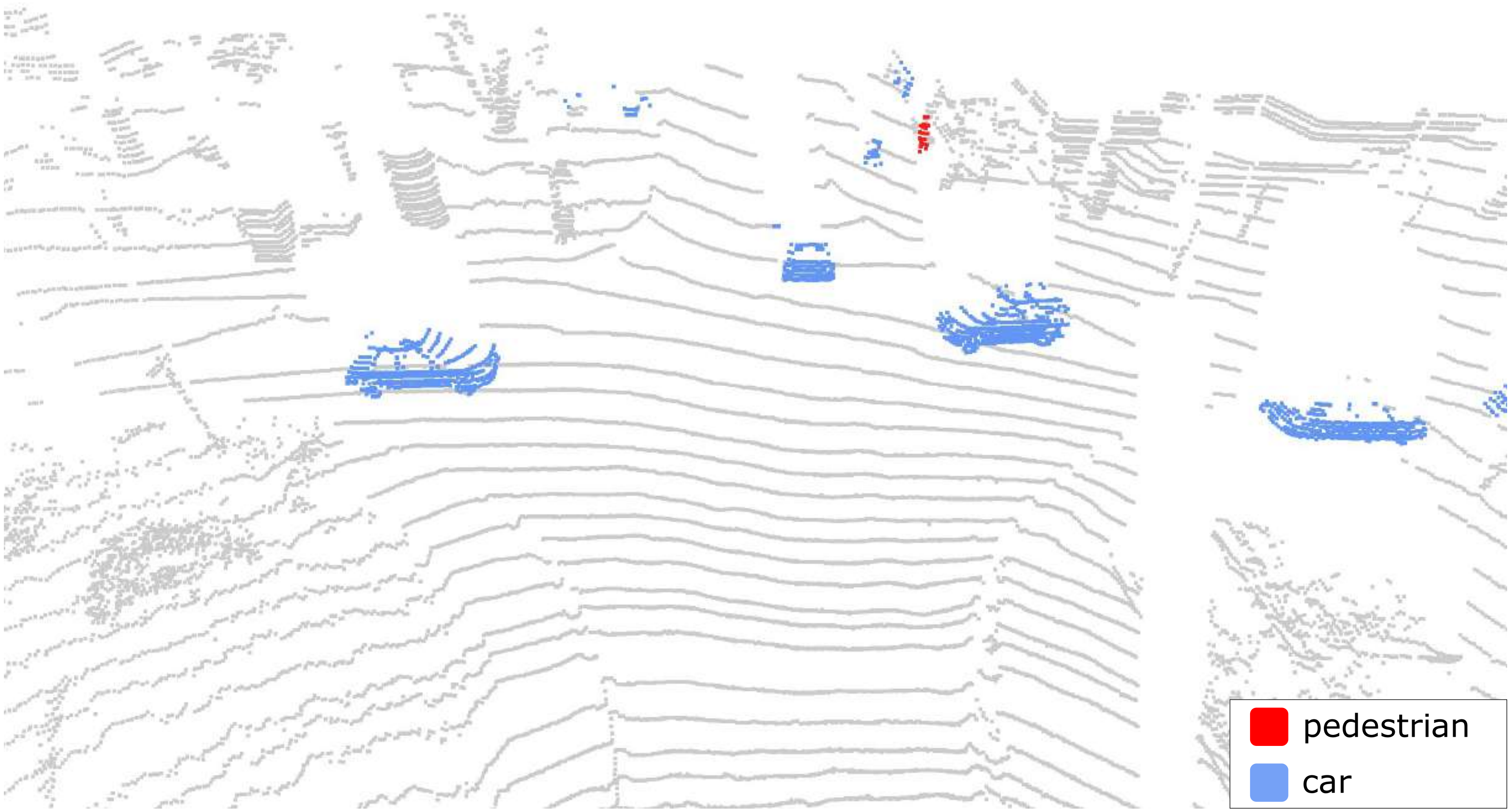
# Point Cloud Features (PCA)



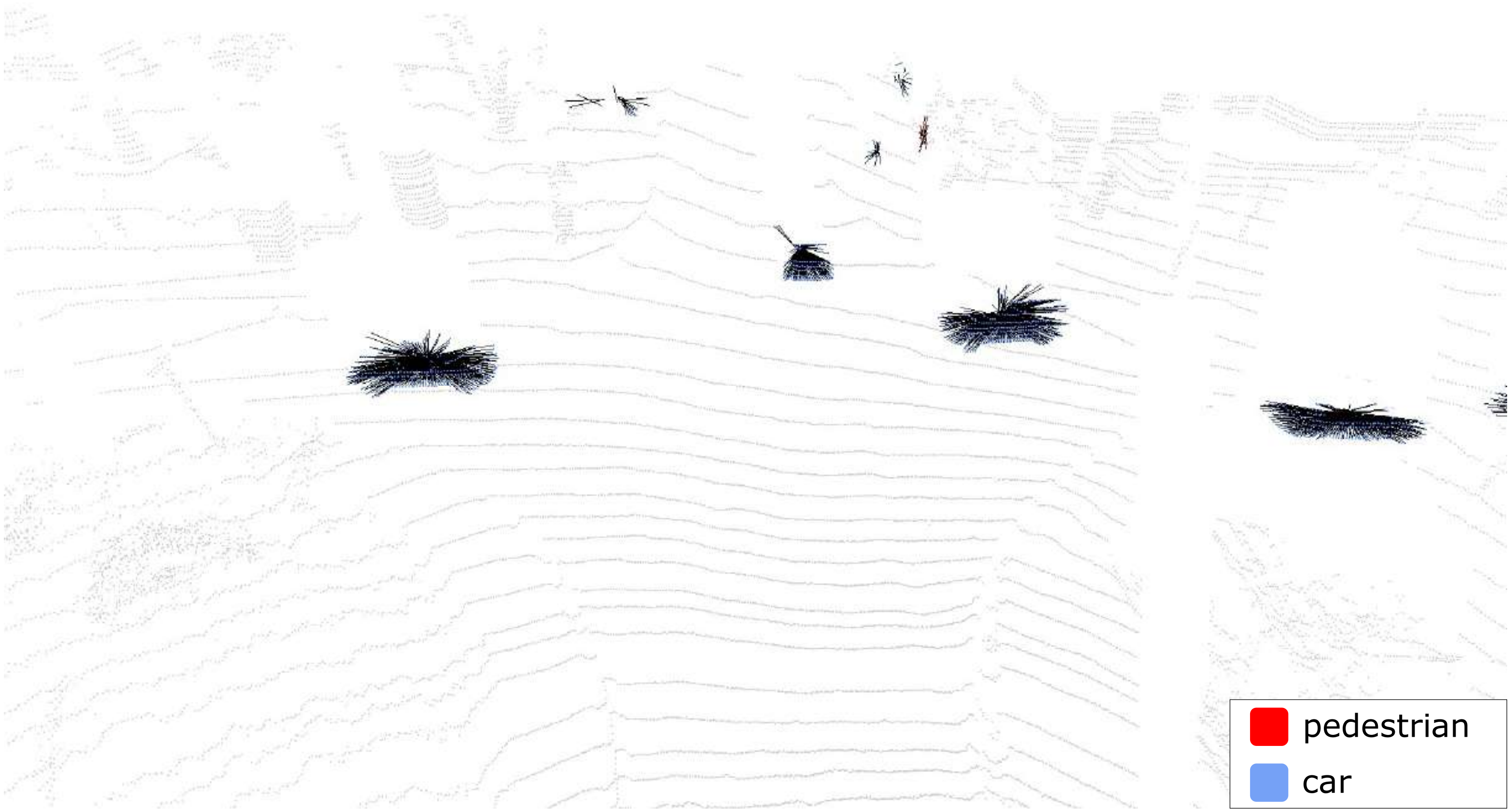
# 3D Semantic Segmentation



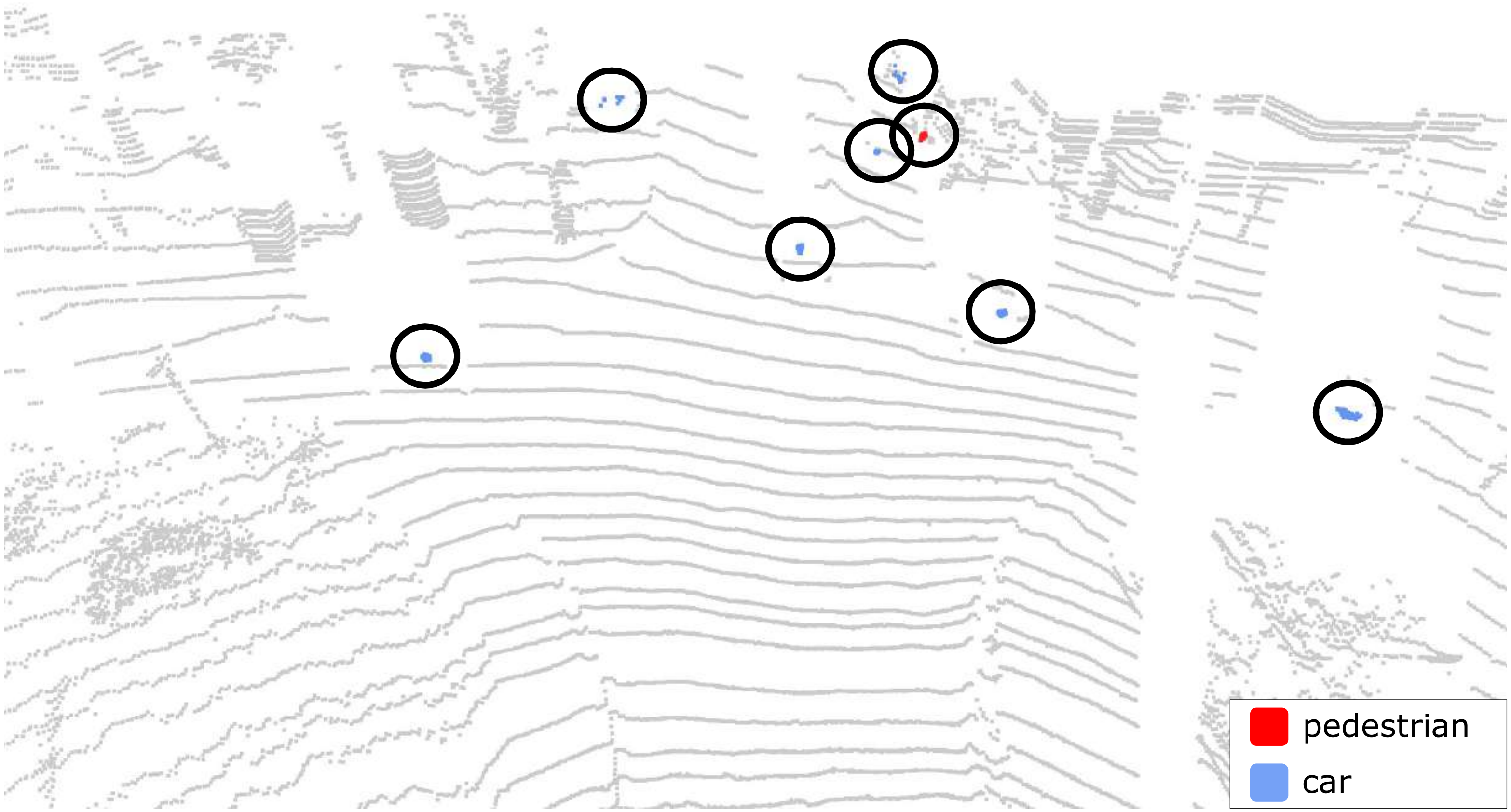
# Filtering Stuff Classes



# Predicting Offsets to Instance Centers

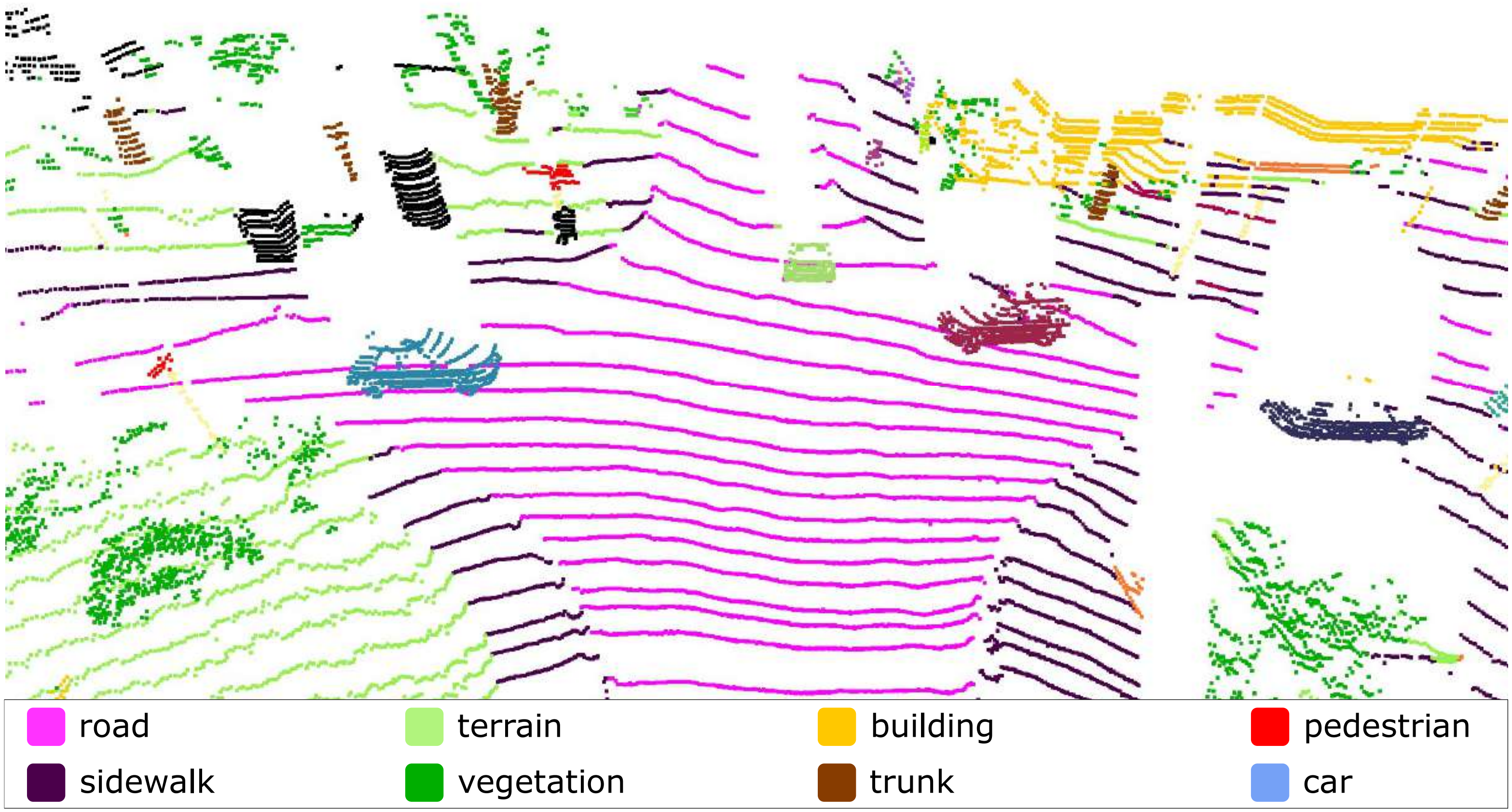


# Using Offsets and Clustering

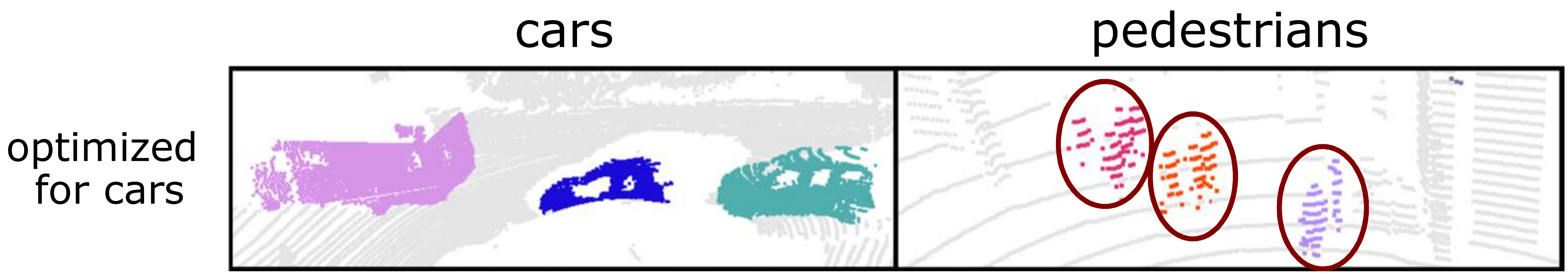


# 3D Instance Segmentation

# 3D Panoptic Segmentation



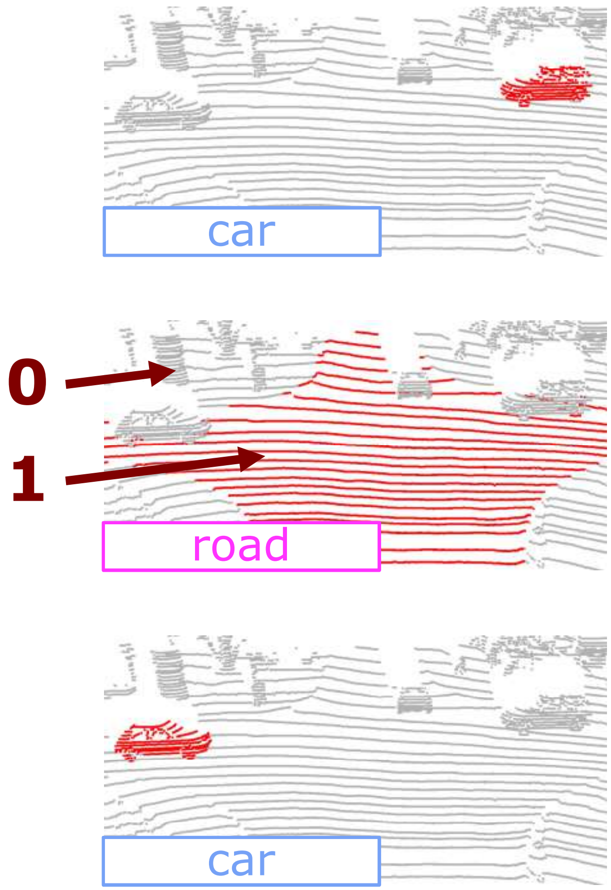
# Limitations of Clustering




- Clustering requires hand-tuning parameters
- Cannot learn parameters from data
- Optimizing for proxy task: offset prediction

# How can we avoid clustering?

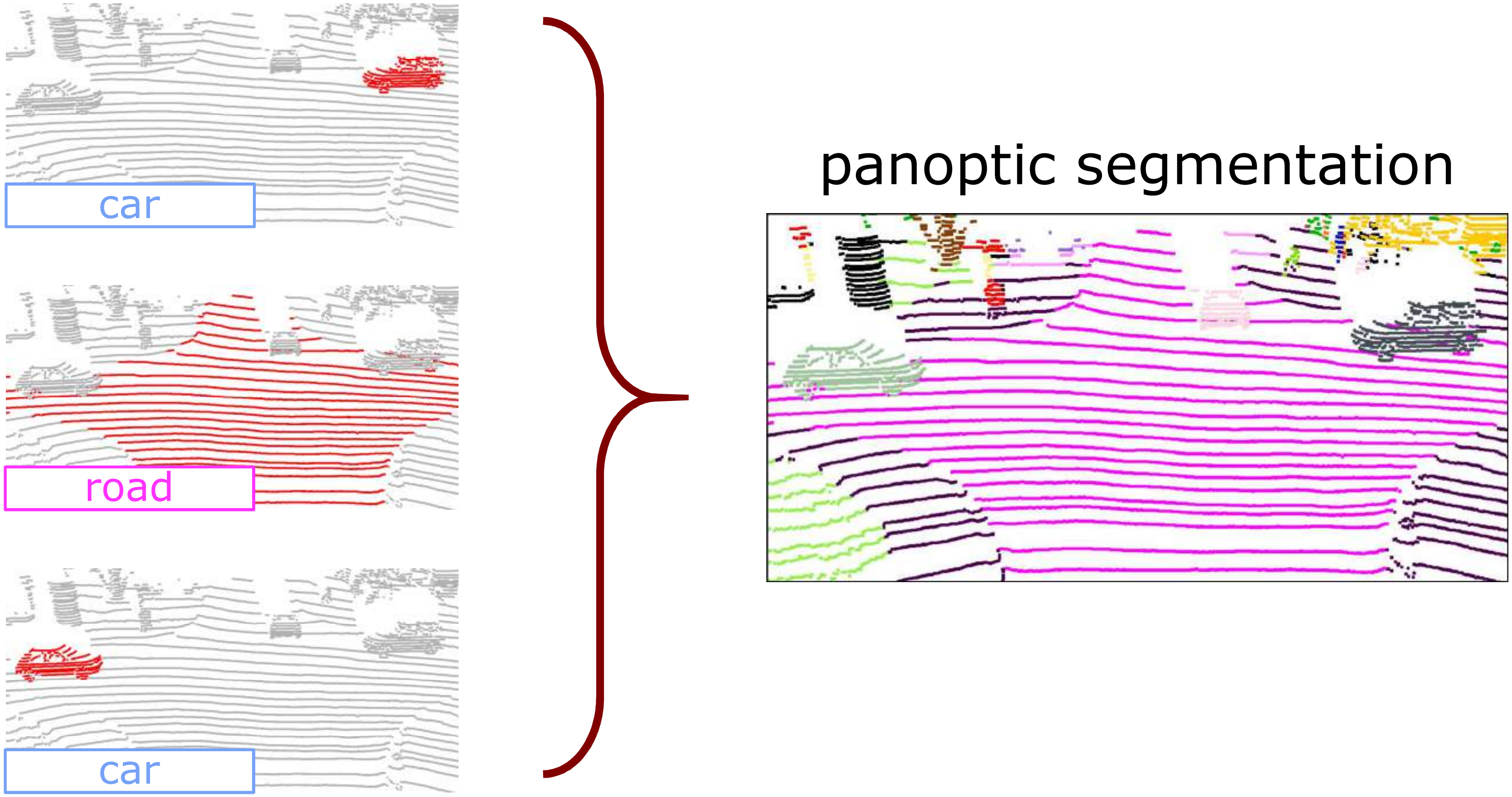
# Panoptic Segmentation Using Masks




- Binary mask + semantic class
- "Things" instance or "stuff" class
- Mask with class and instance ID

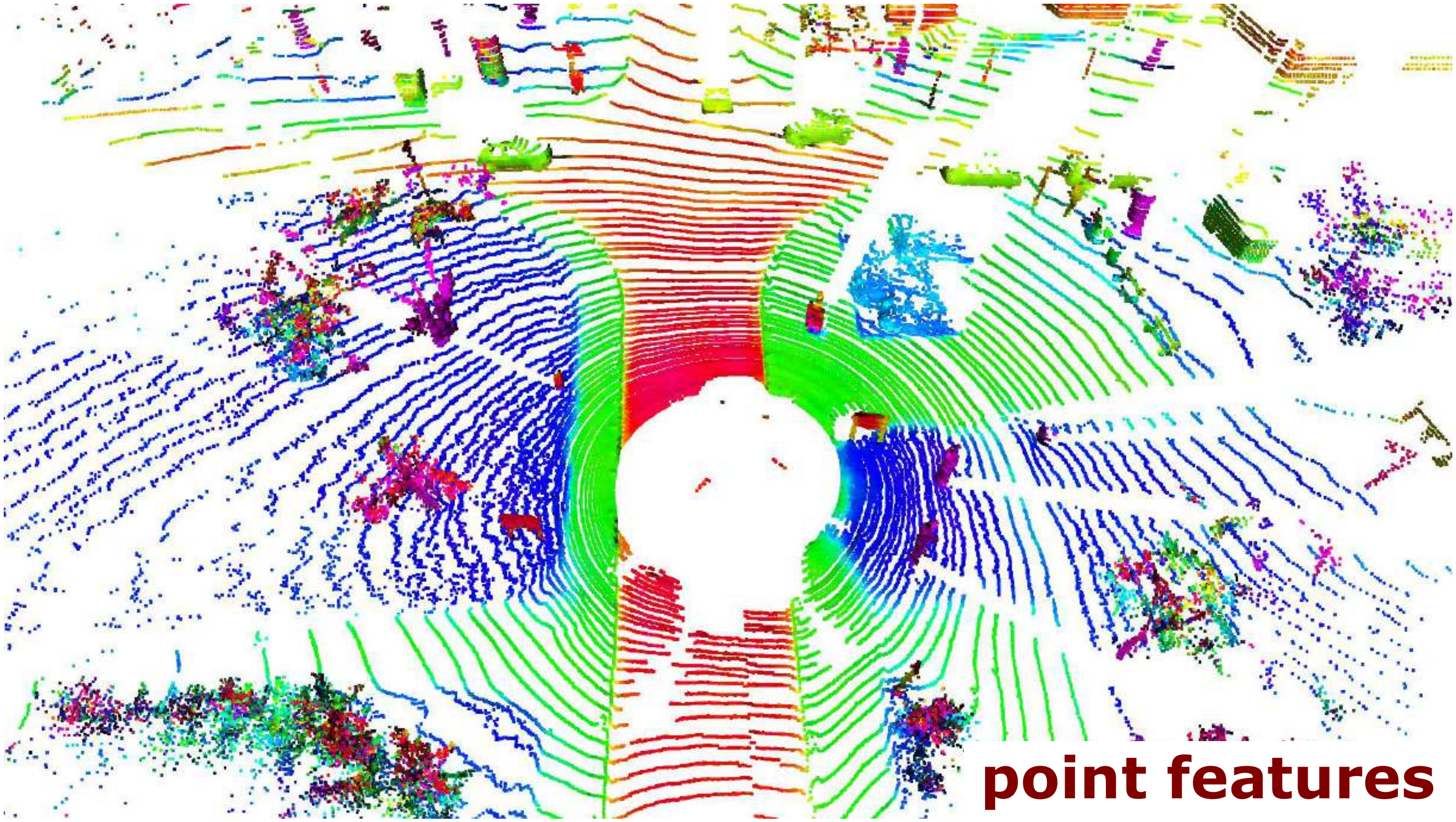
Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Panoptic Segmentation Using Masks



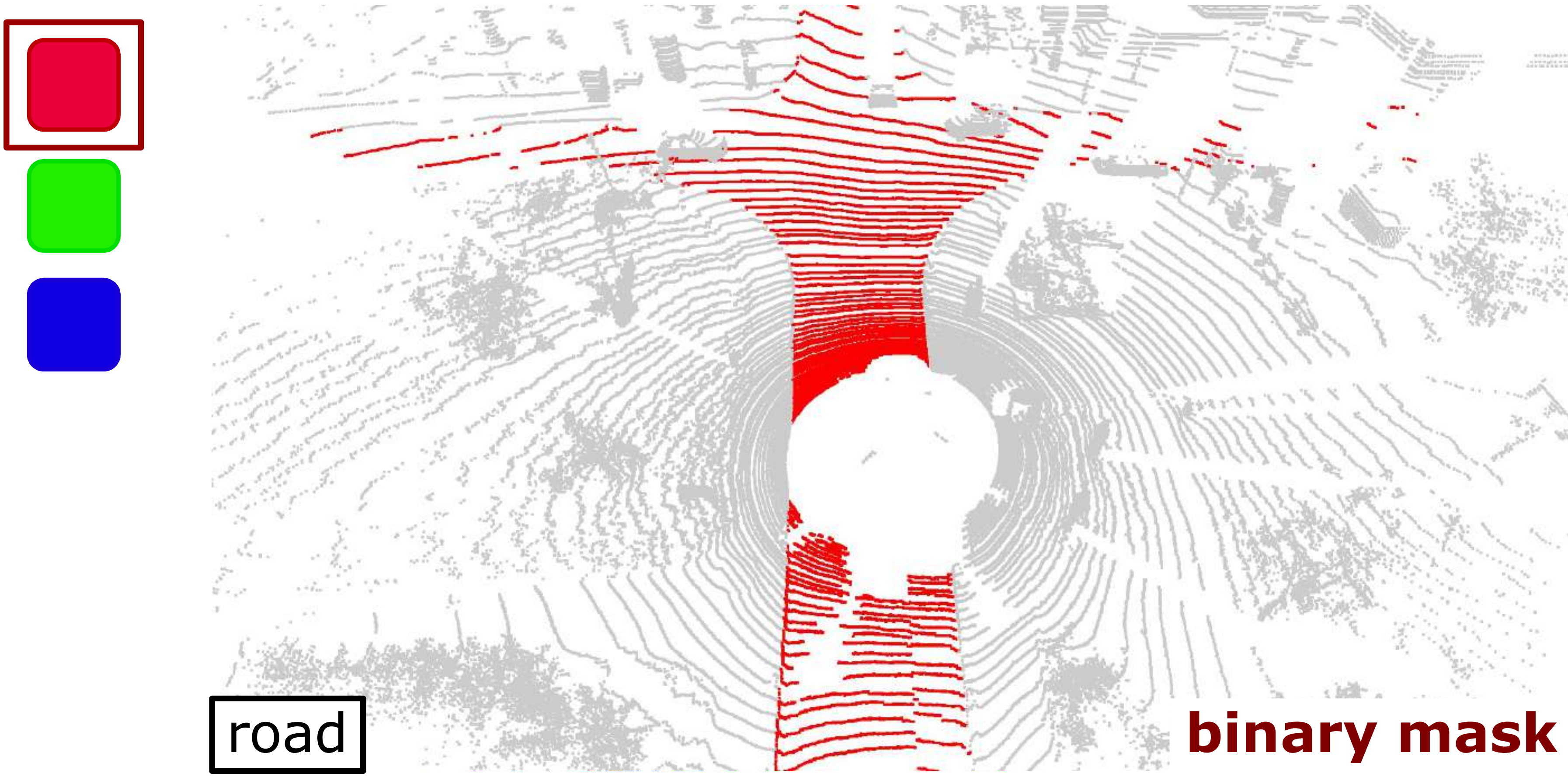

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Feature Extraction



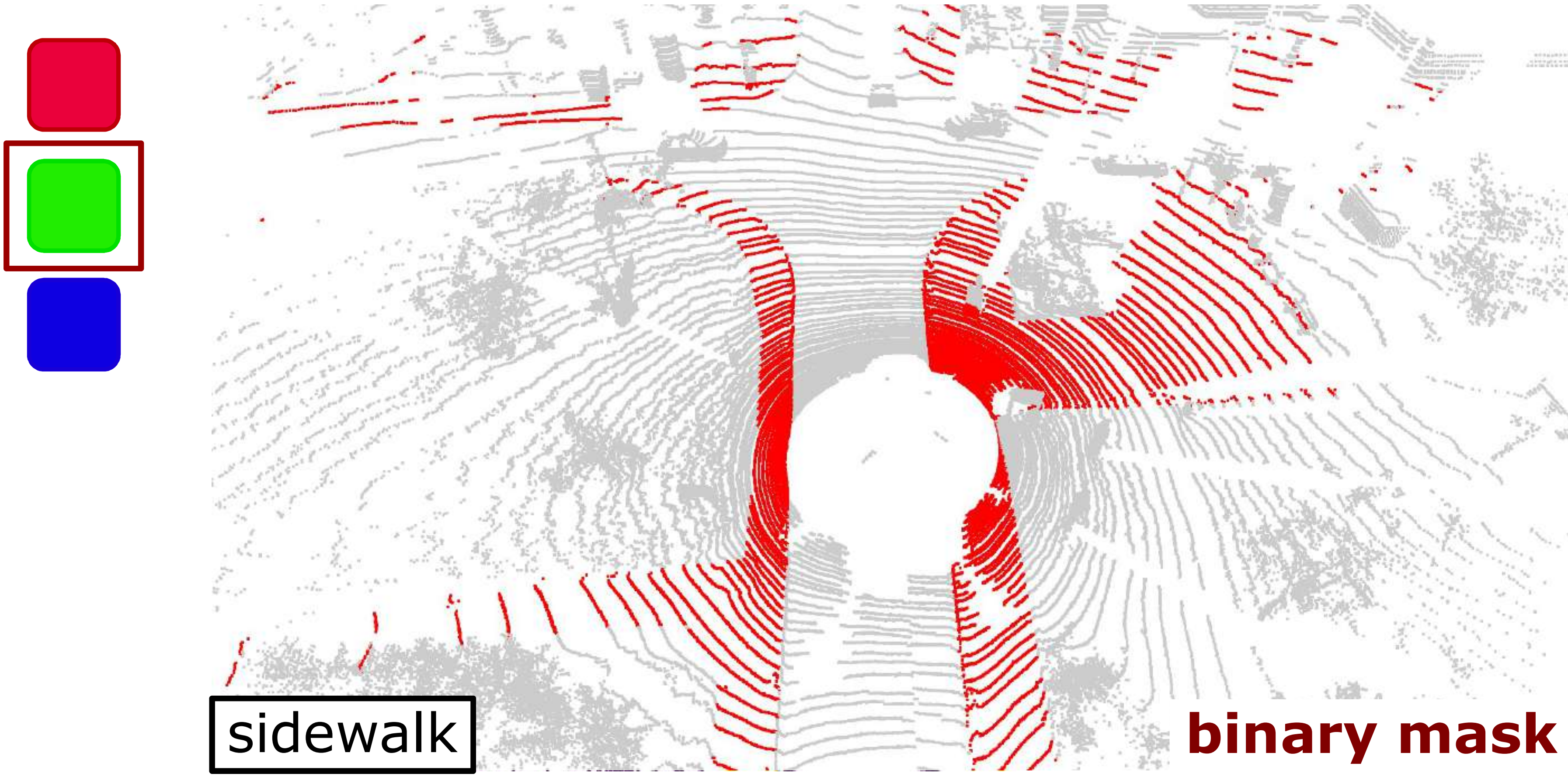

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Queries as Mask Proposals



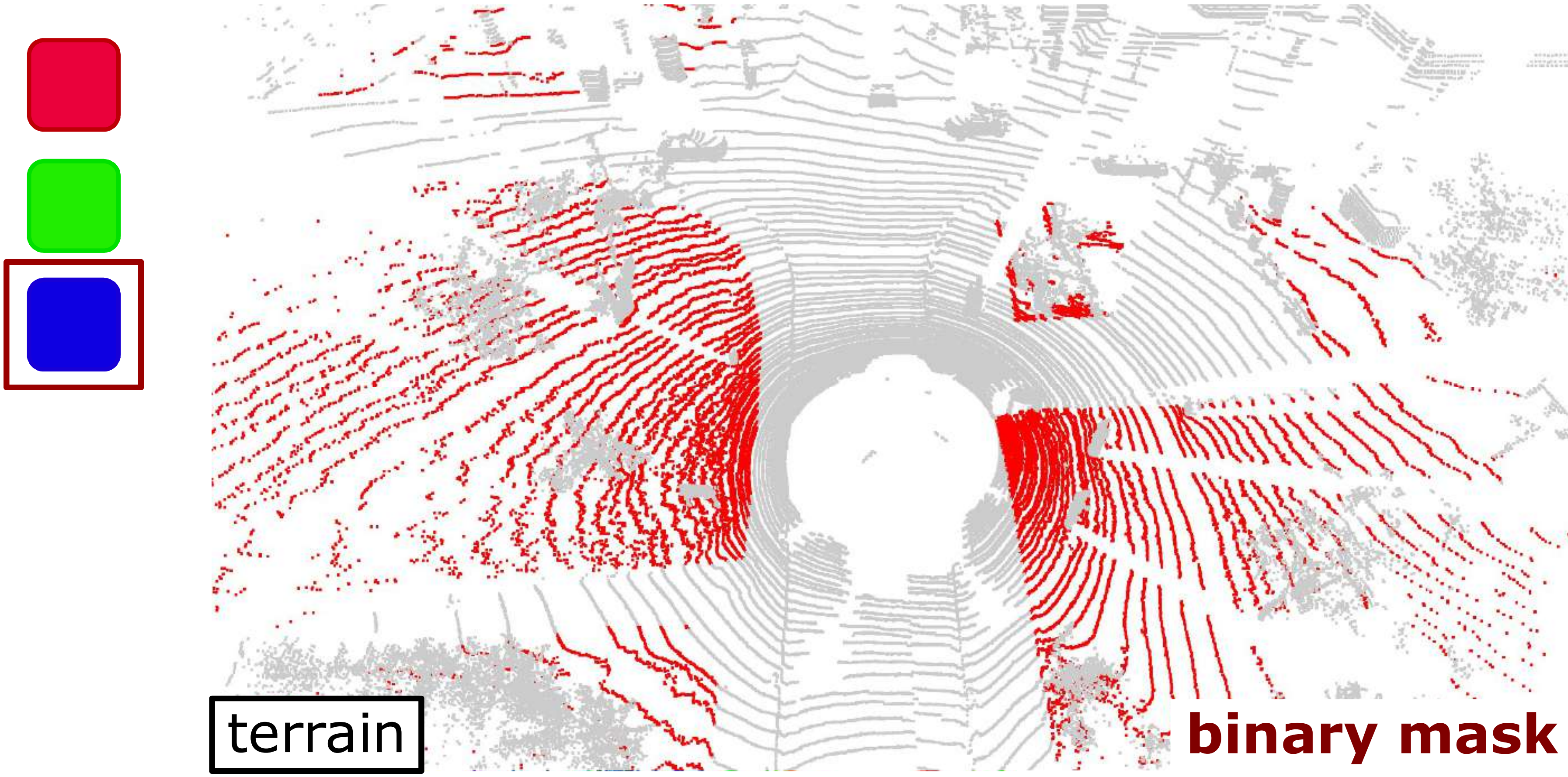

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Queries as Mask Proposals



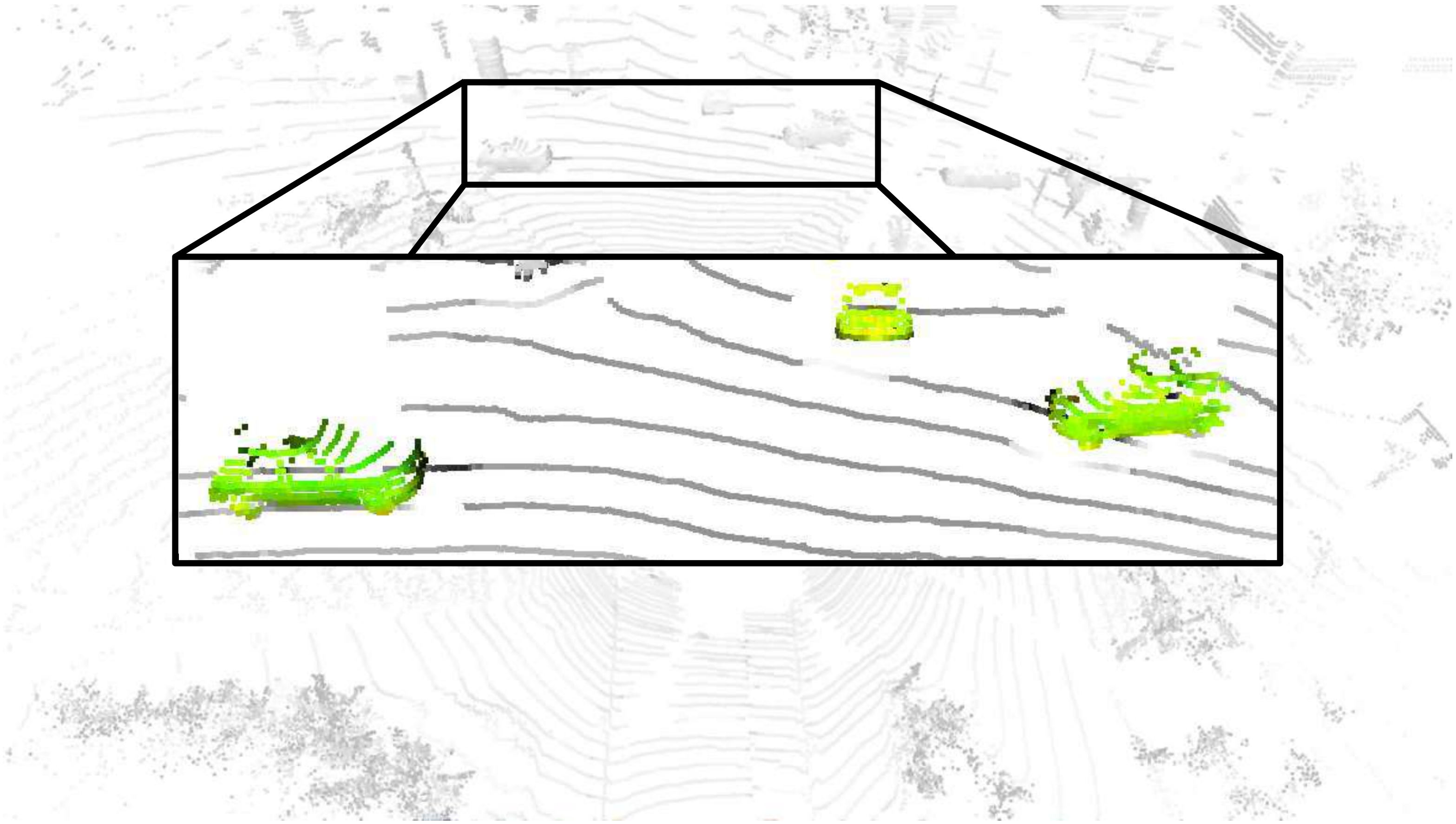

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Queries as Mask Proposals



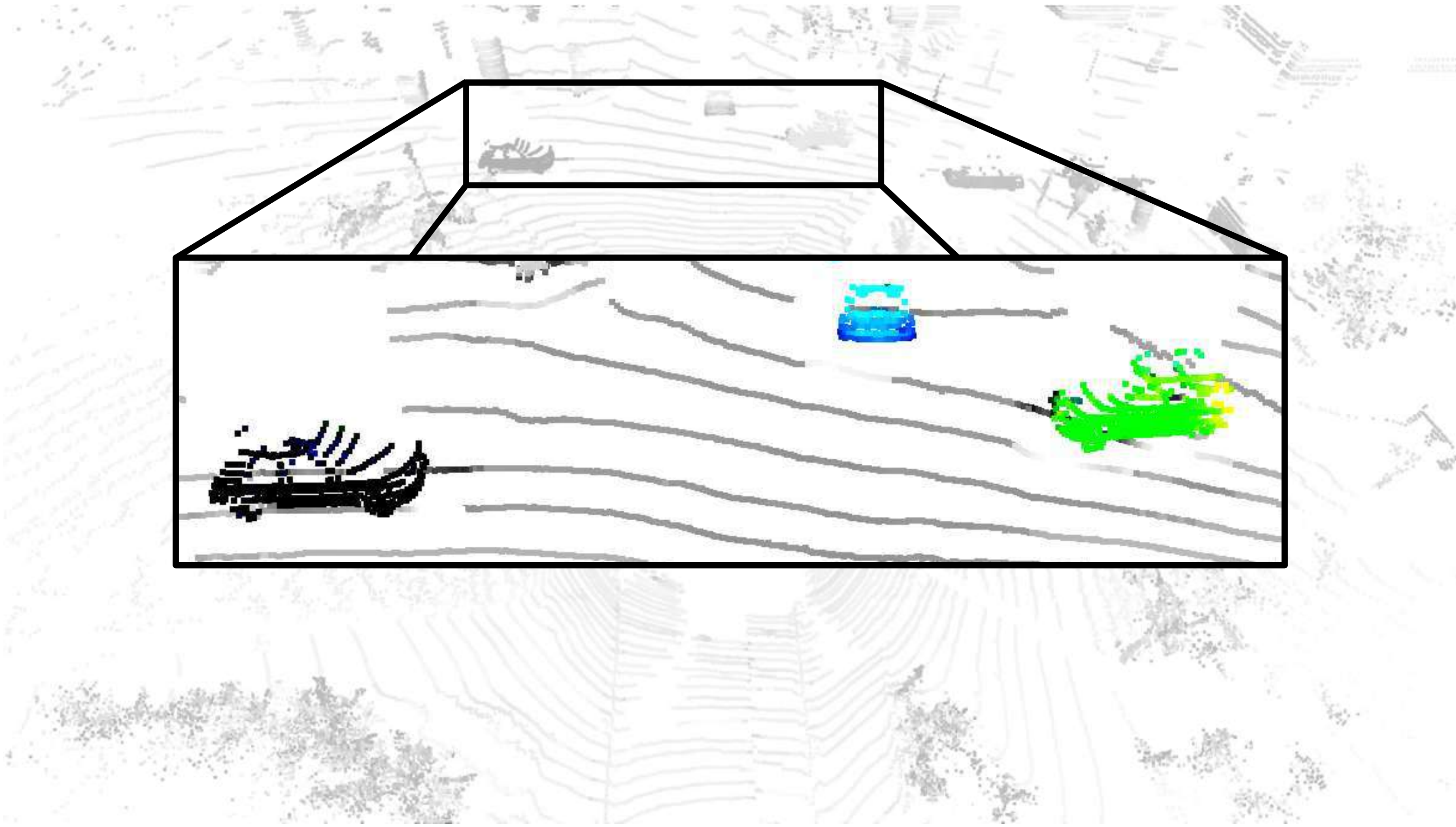

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Features for Things



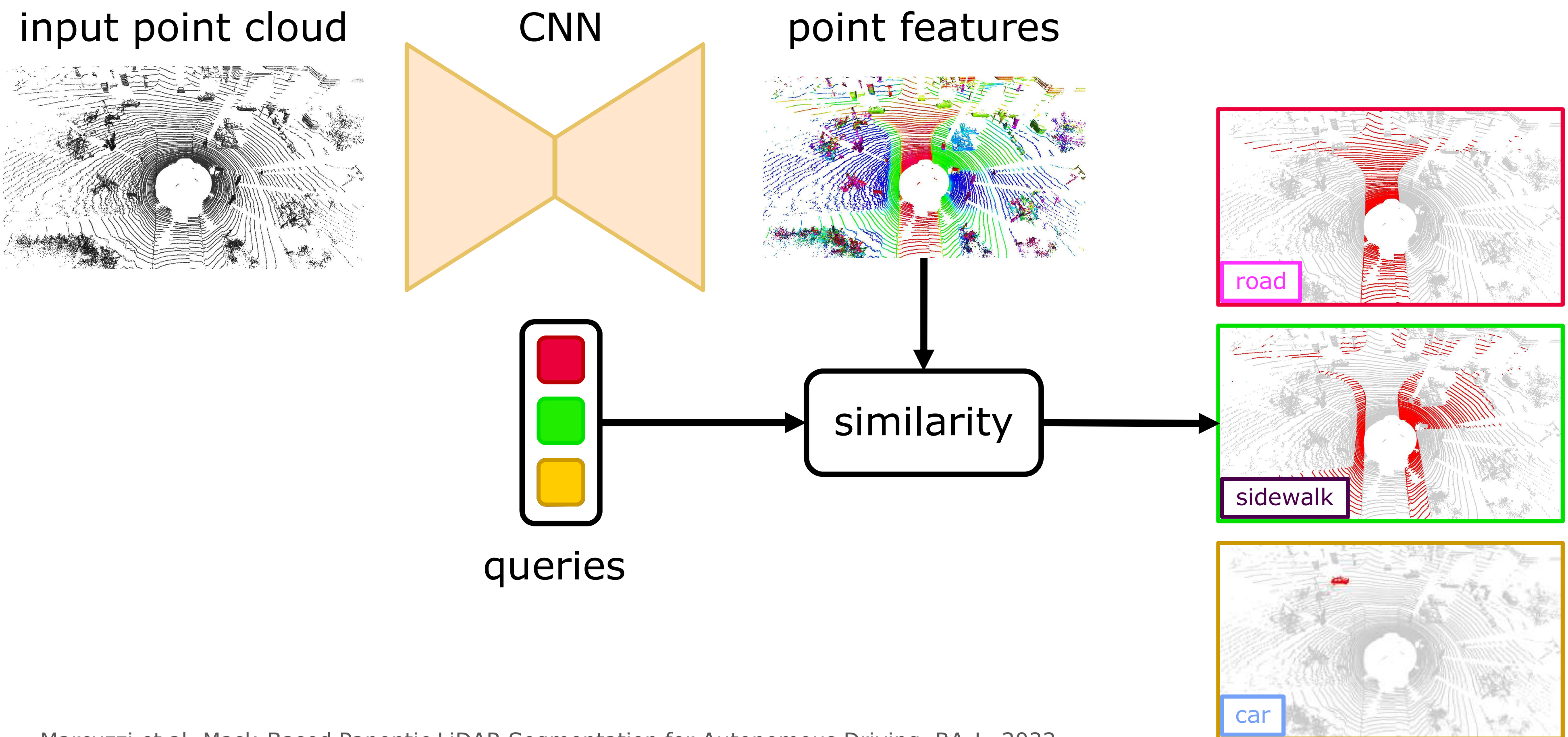
Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Adding Positional Information



Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Mask-Based 3D Panoptic Segmentation



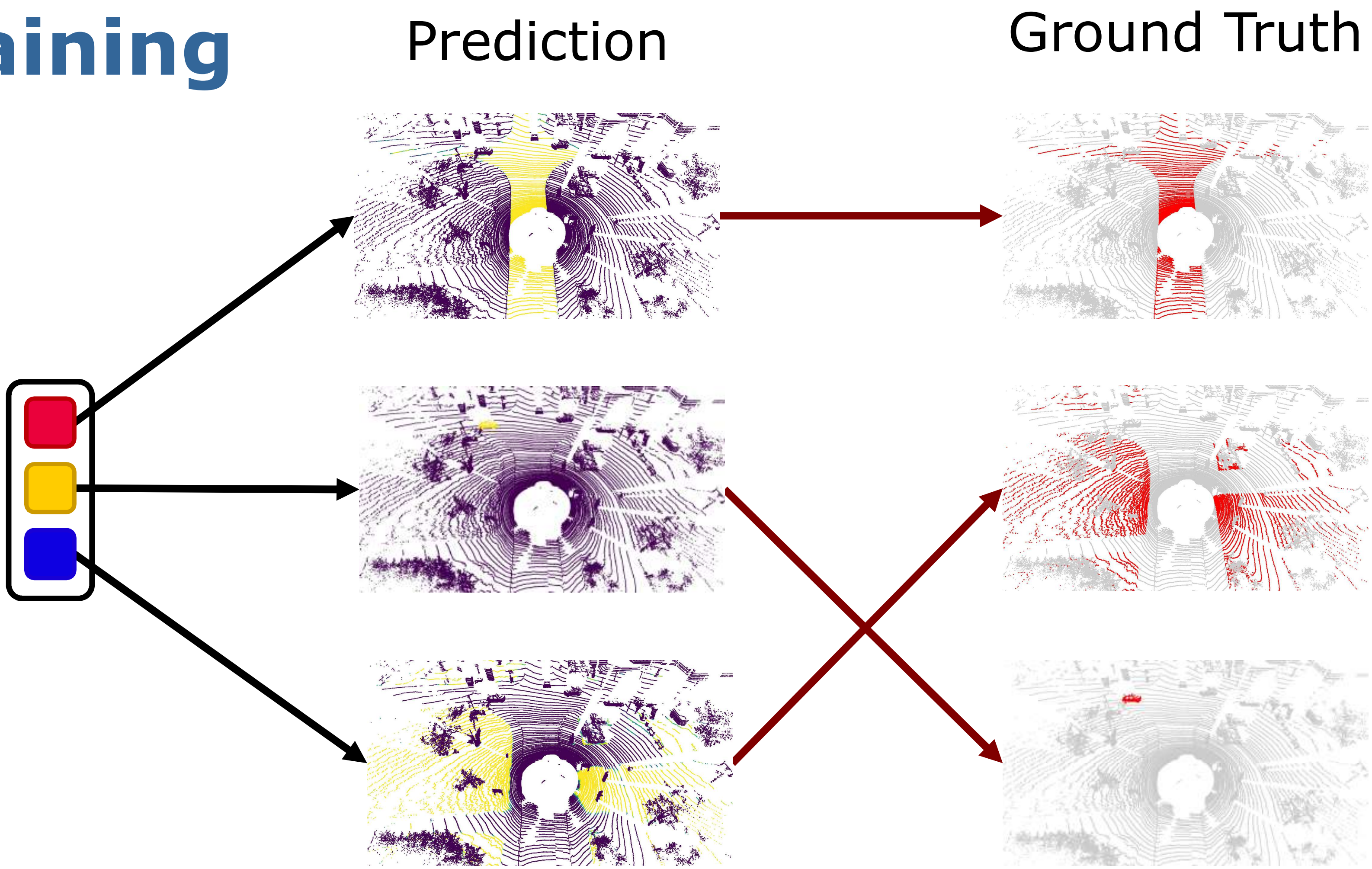

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Mask-Based 3D Panoptic Segmentation

3D panoptic segmentation

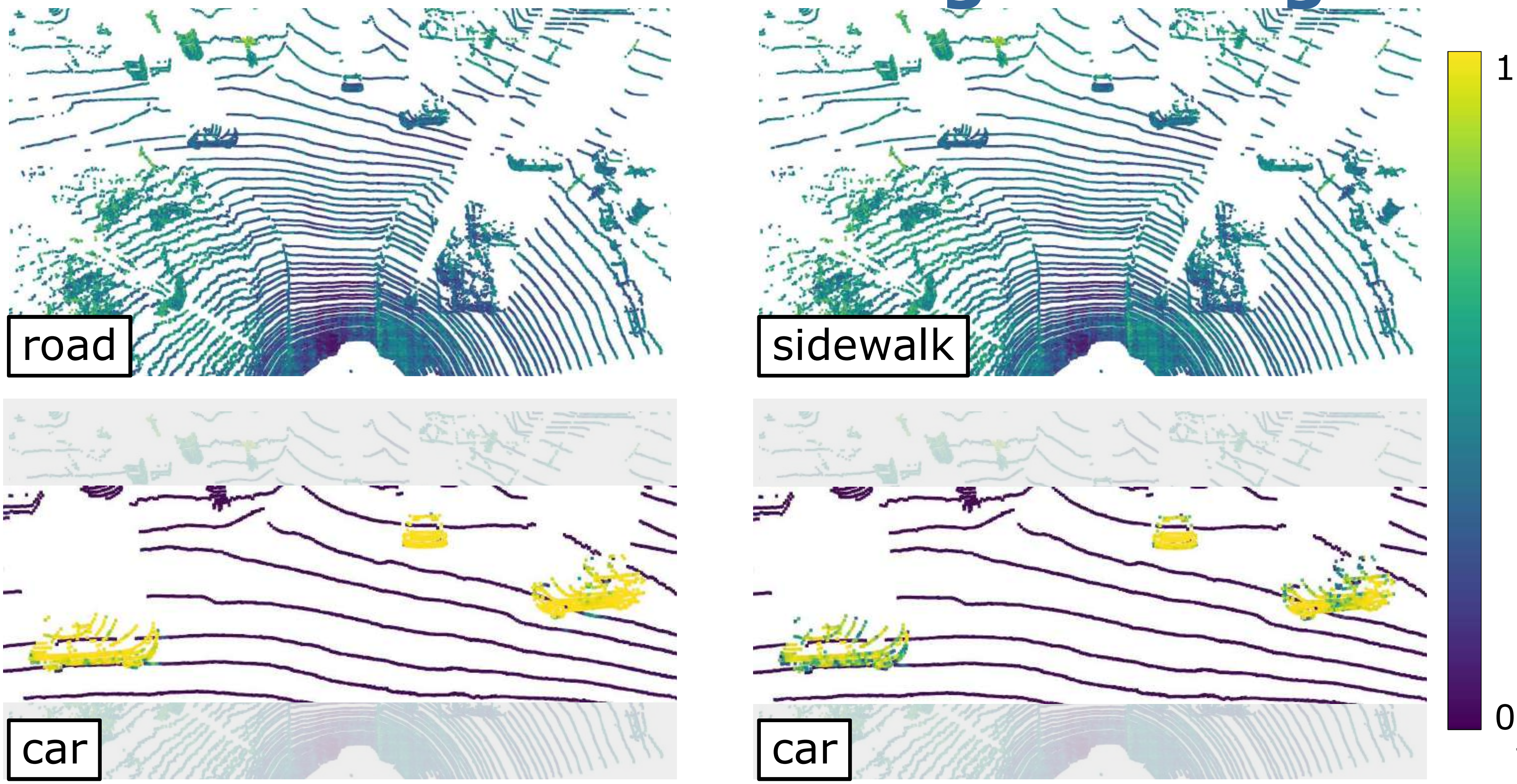

Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Training

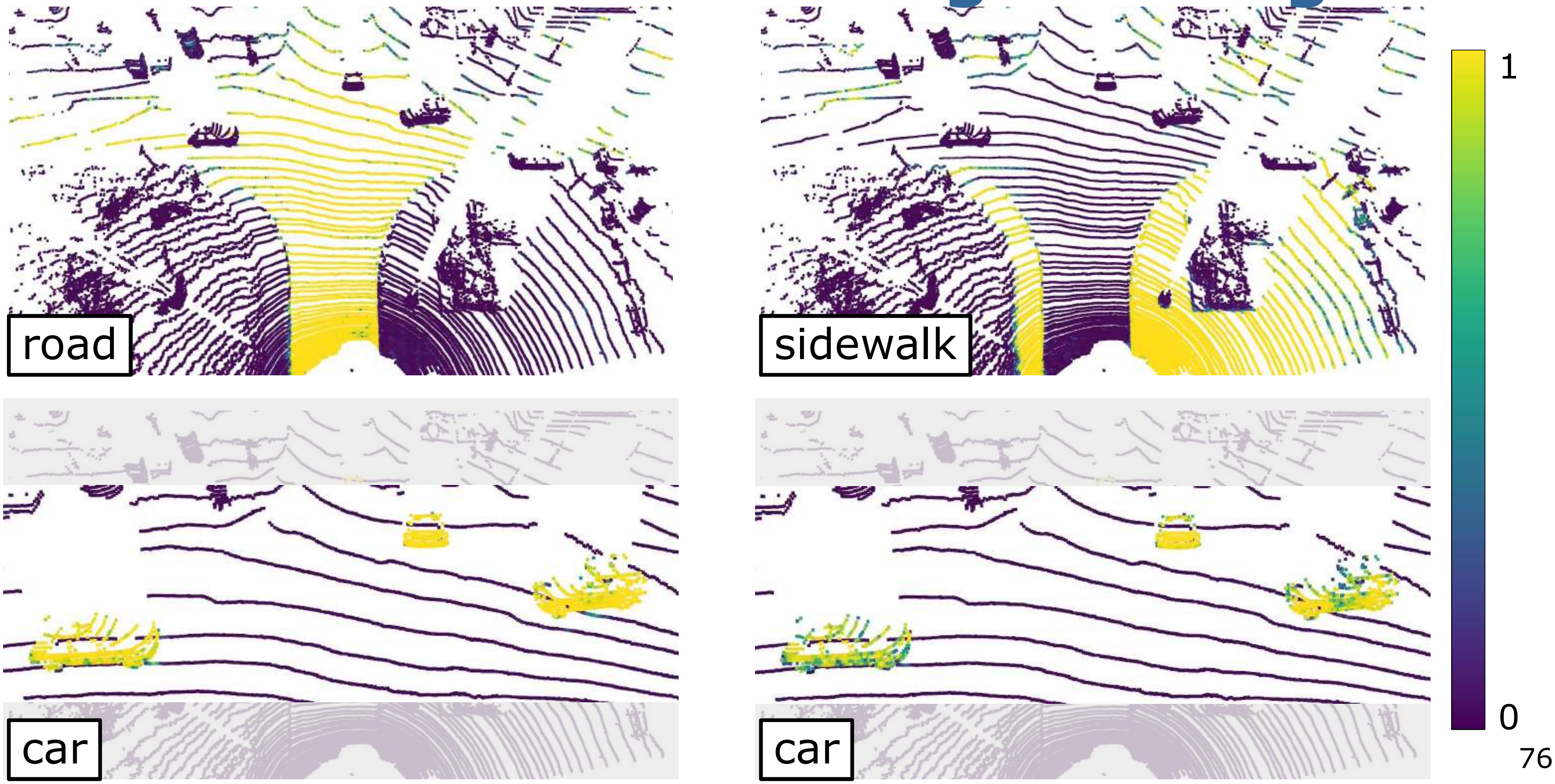



Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

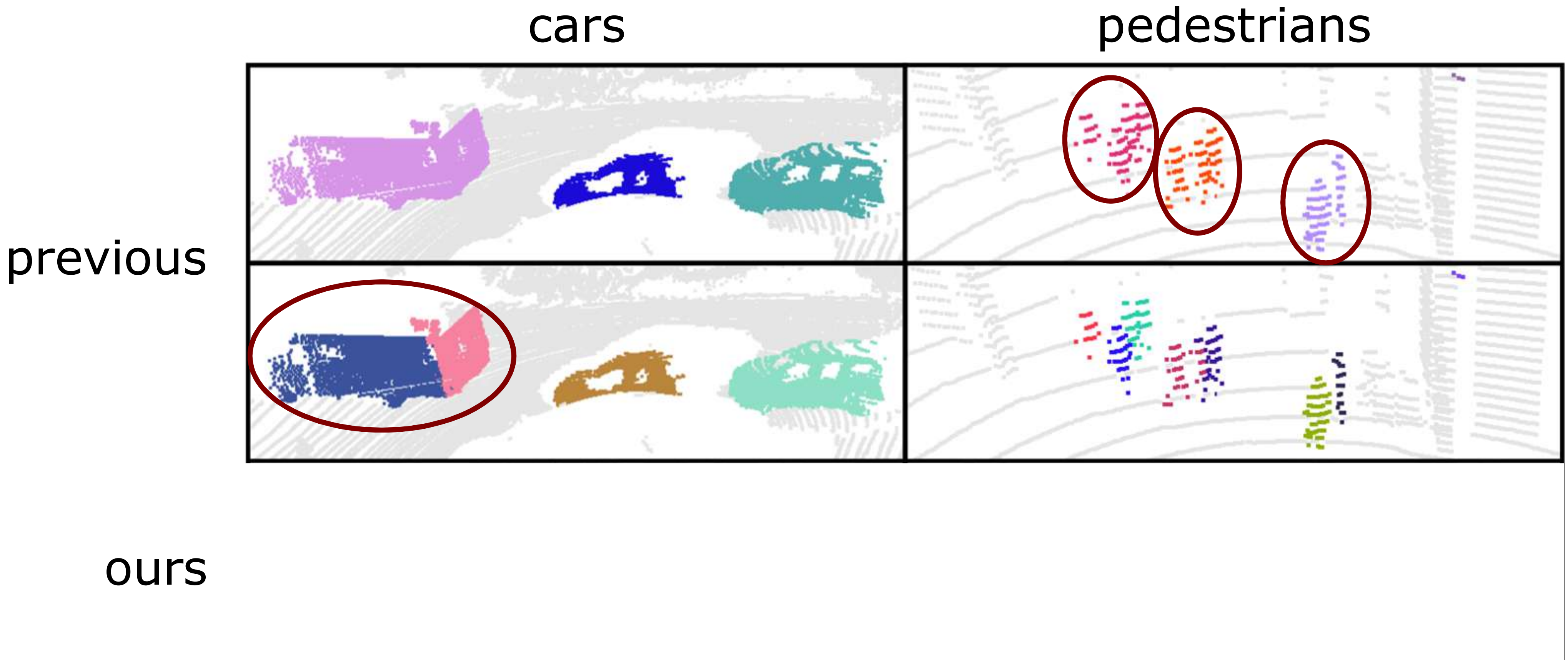
# Evolution of Masks during Training



# Evolution of Masks during Training

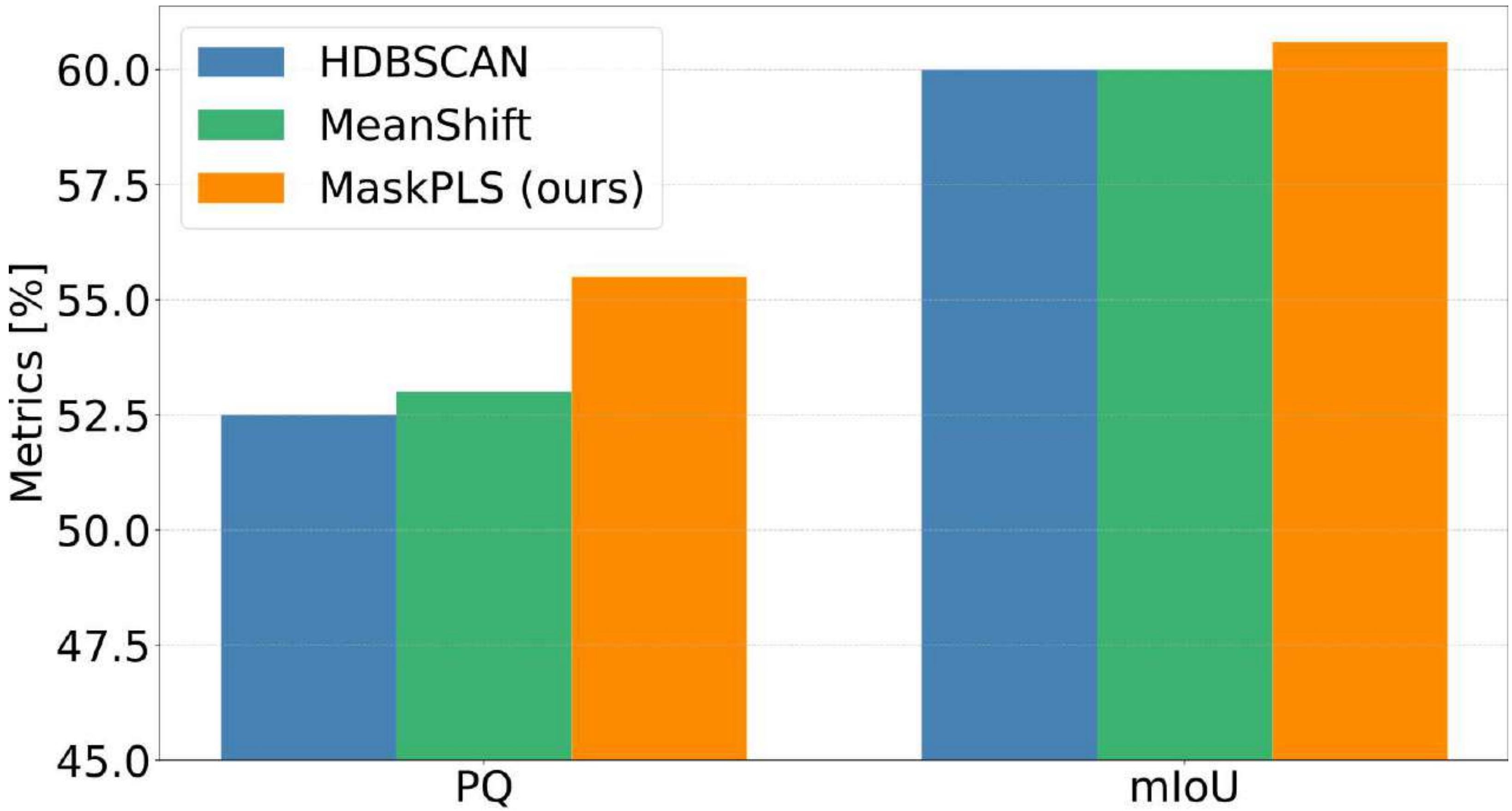


# Results for Different Classes

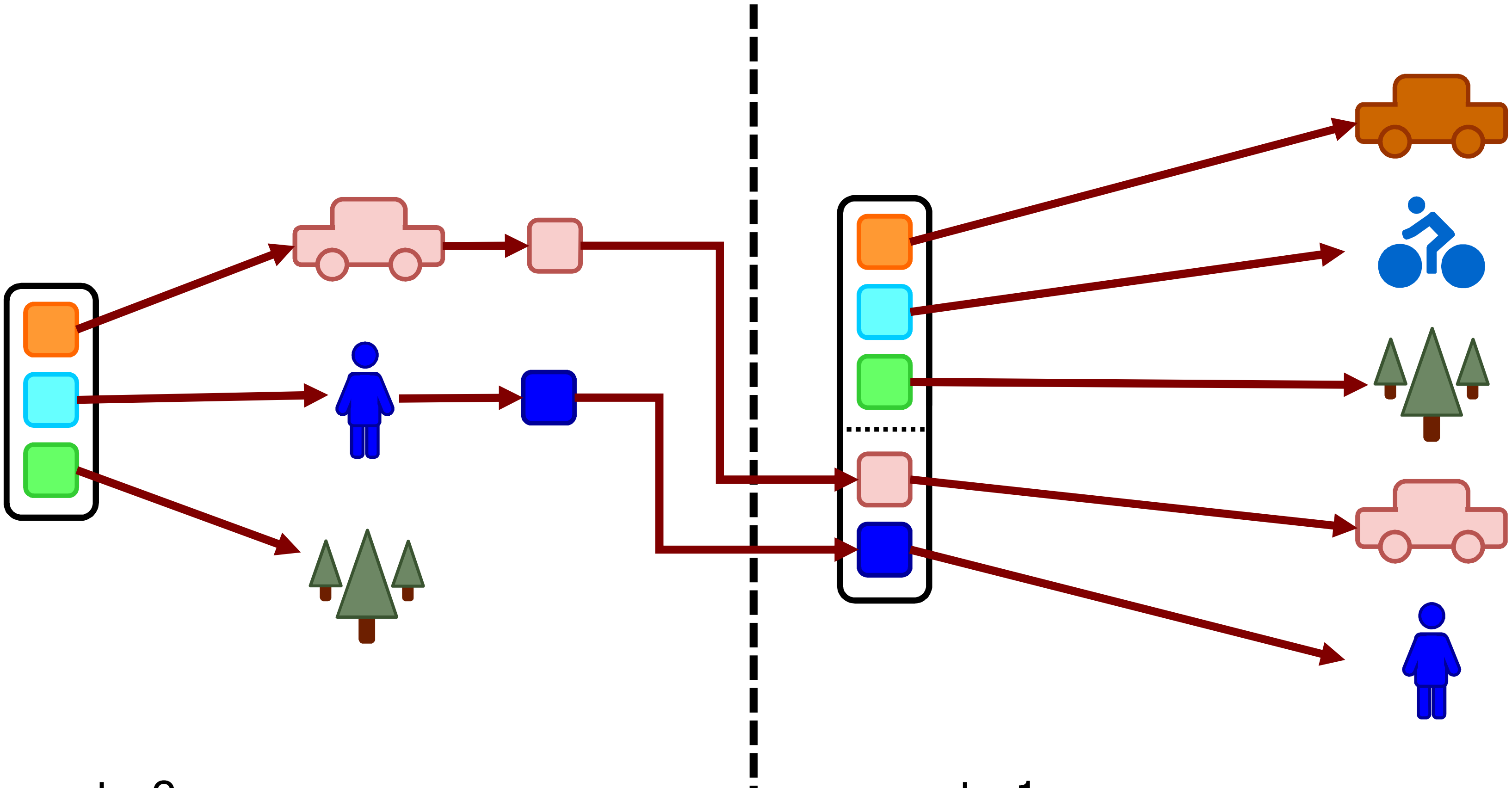



Marcuzzi et al. Mask-Based Panoptic LiDAR Segmentation for Autonomous Driving, RA-L, 2022.

# Comparison with Clustering



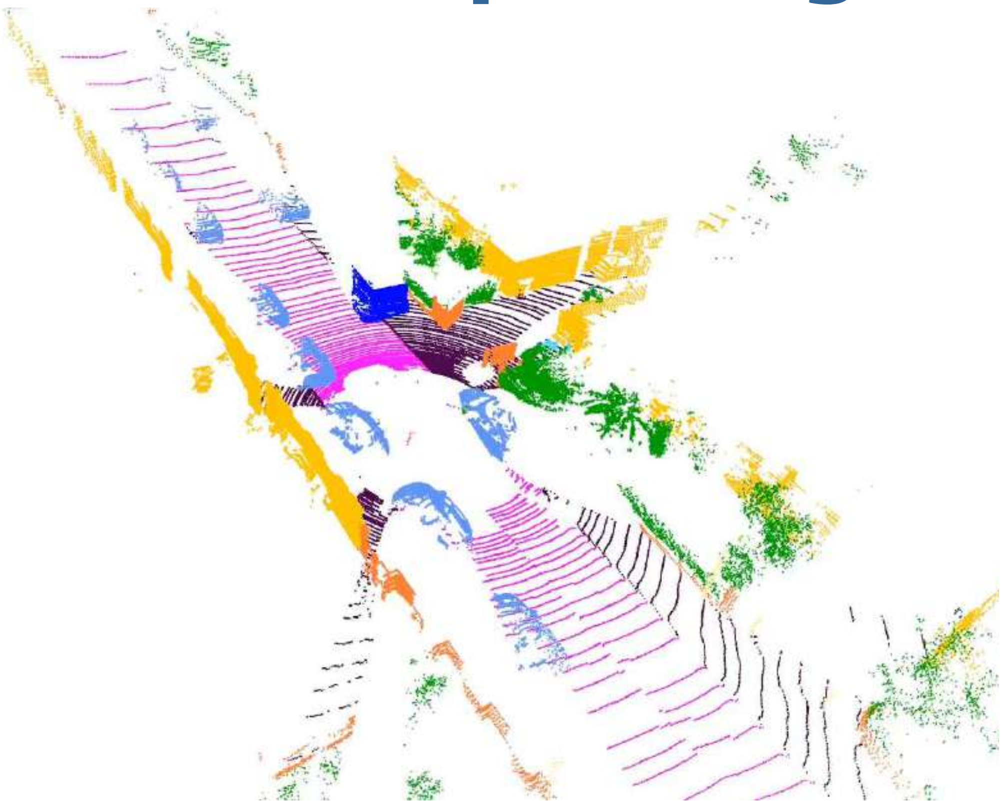
# 4D Panoptic Segmentation



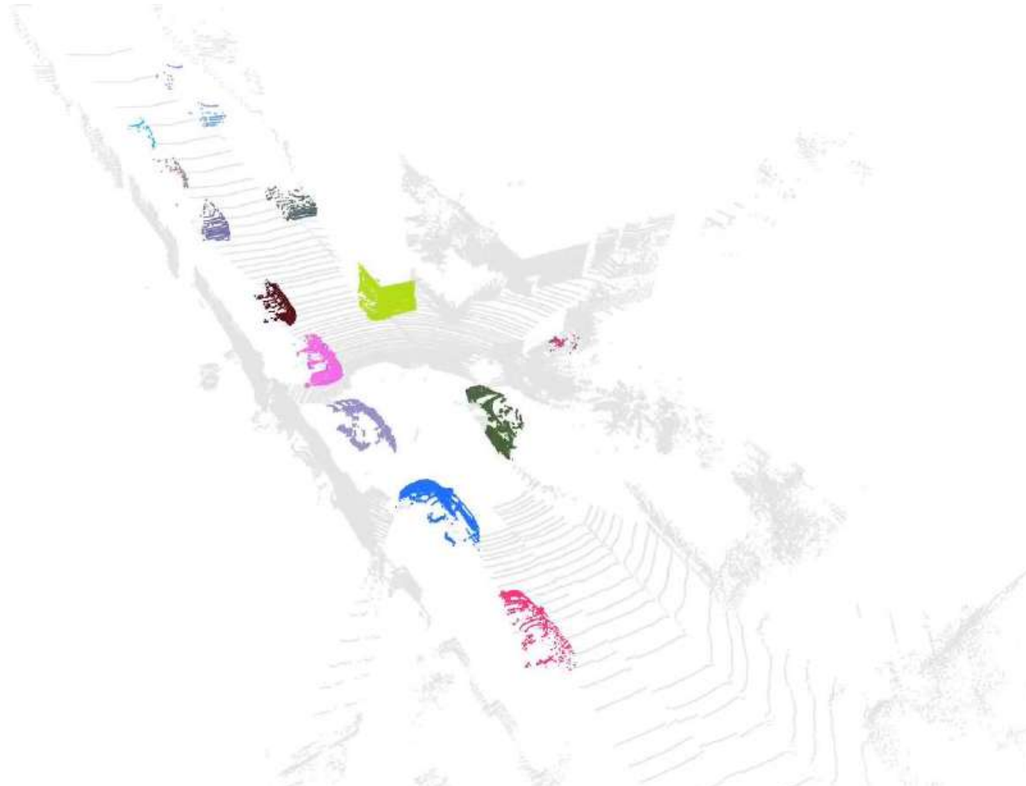
scan t=0

scan t=1

# 4D Panoptic Segmentation

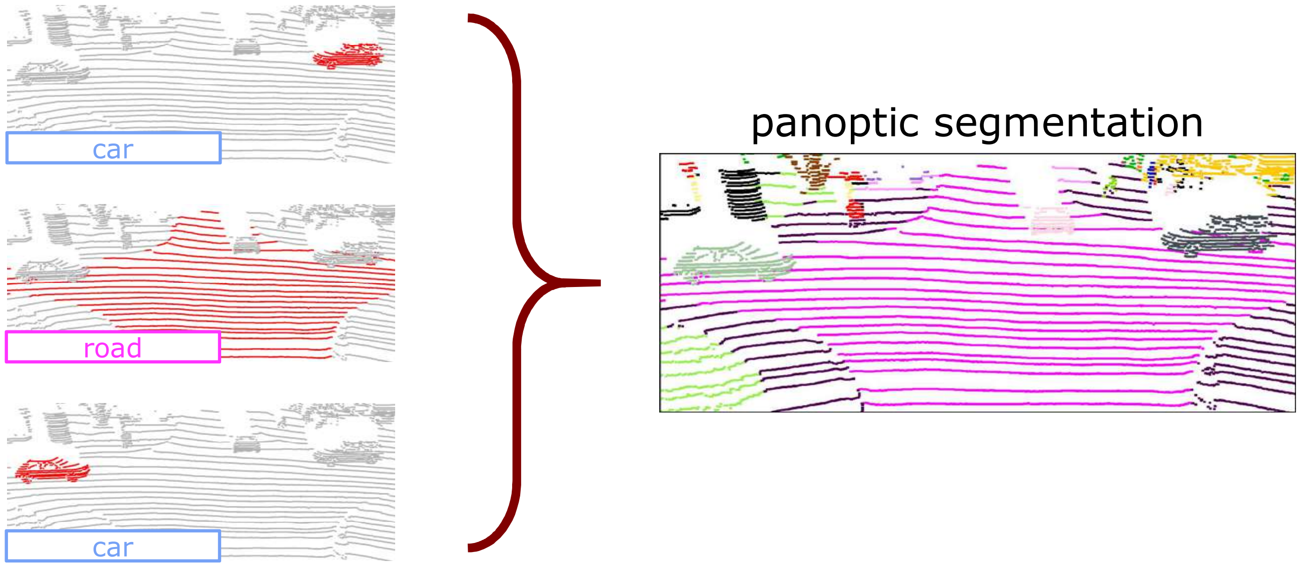


semantic segmentation



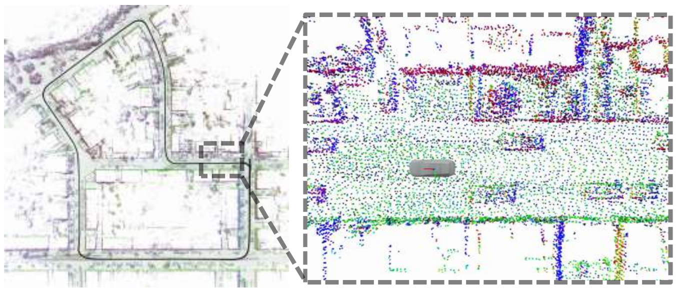
instance segmentation

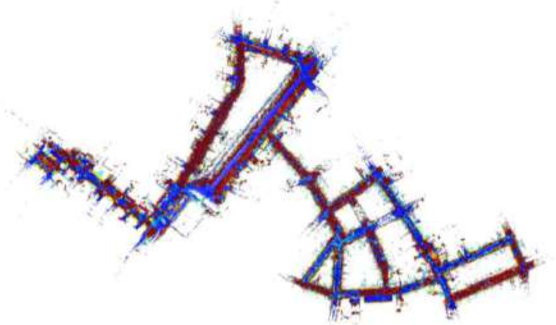
# Summary: Mask-based Segmentation

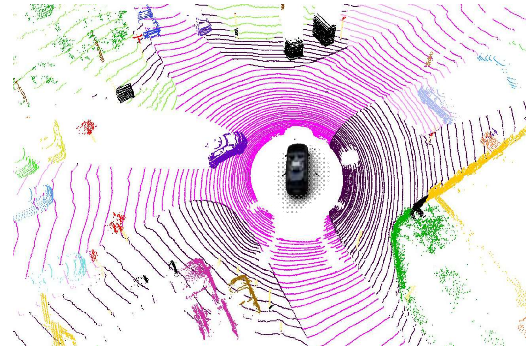



- Rethinking 3D panoptic segmentation as mask prediction
- **Unified handling** of “stuff” and “thing” classes
- No hyperparameter tuning for clustering needed!

# Conclusion







- 3D LiDAR-based mapping using a flexible neural point-based representation
- Refinement with continuous-trajectory bundle adjustment
- Mask-based 3D panoptic segmentation

**Thank you for your attention!**