Segment Anything Model for Pedestrian Infrastructure Inventory: Assess Zero-Shot Segmentation on Multi-Mode Geospatial Data

Jiahao Xia[1], Gavin, Gong[2], Jiawei Liu[3], Zhigang Zhu[4], Hao Tang[5]

[1] Department of Civil and Environmental Engineering, Rutgers, the State University of New Jersey; [2] East Brunswick High School; [3] Data Science and Engineering Program, The CUNY City College; [4] Department of Computer Science, The CUNY City College; 5 Department of Computer Information Systems, The CUNY BMCC.

INTRODUCTION

In this study, a Segment Anything Model (SAM)-based pedestrian infrastructure segmentation workflow is designed and optimized, which is capable of efficiently processing multi-sourced geospatial data including LiDAR data and satellite imagery data. We used an expanded definition of pedestrian infrastructure inventory which goes beyond the traditional transportation elements to include street furniture objects often omitted from the traditional definition. Our contributions lie in producing the necessary knowledge to answer the following two questions. First, which data representation can facilitate zero-shot segmentation of infrastructure objects with SAM? Second, how well does the SAM-based method perform on segmenting pedestrian infrastructure objects? Our findings indicate that street view images generated from mobile LiDAR point cloud data, when paired along with satellite imagery data, can work efficiently with SAM to create a scalable pedestrian infrastructure inventory approach with immediate benefits to GIS professionals, city managers, transportation owners, and walkers, especially those with travel-limiting disabilities.

DATASET
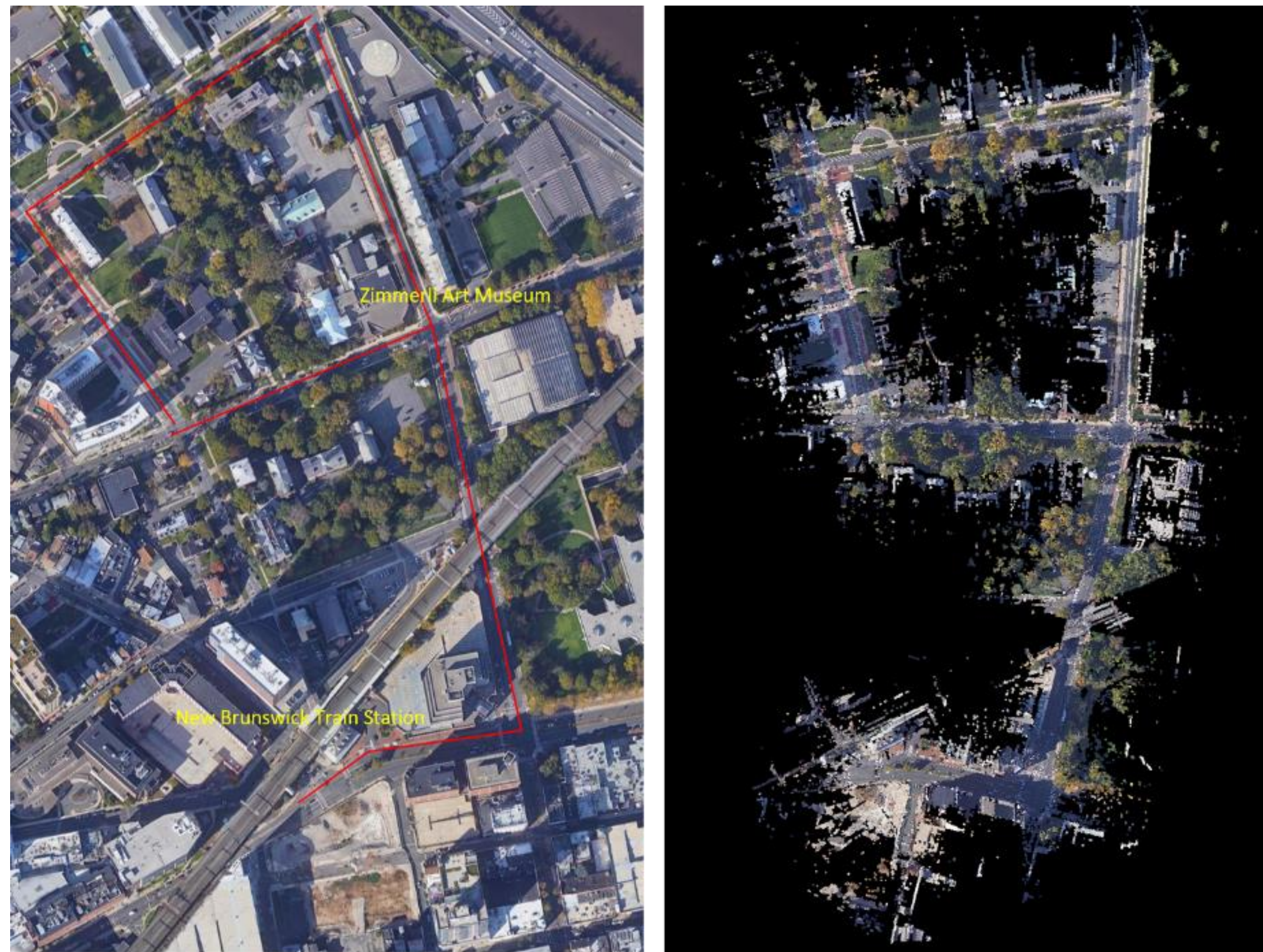
Evaluation Data Set (left: Google satellite imagery and right: the mobile LiDAR data) along a route in downtown New Brunswick, NJ.

Satellite imagery and 3D point cloud data are among common types of data that have been studied for extracting pedestrian infrastructure features at the network level. In our project, we use mobile LiDAR point cloud data as the primary data source and bring in Google satellite imagery as the secondary data source.

METHODLOGY

☐ *Elements of the proposed SAM-based workflow.*

Data pre-processing → Projection and Alignment → Zero-shot segmentation → 2D-3D re-projection
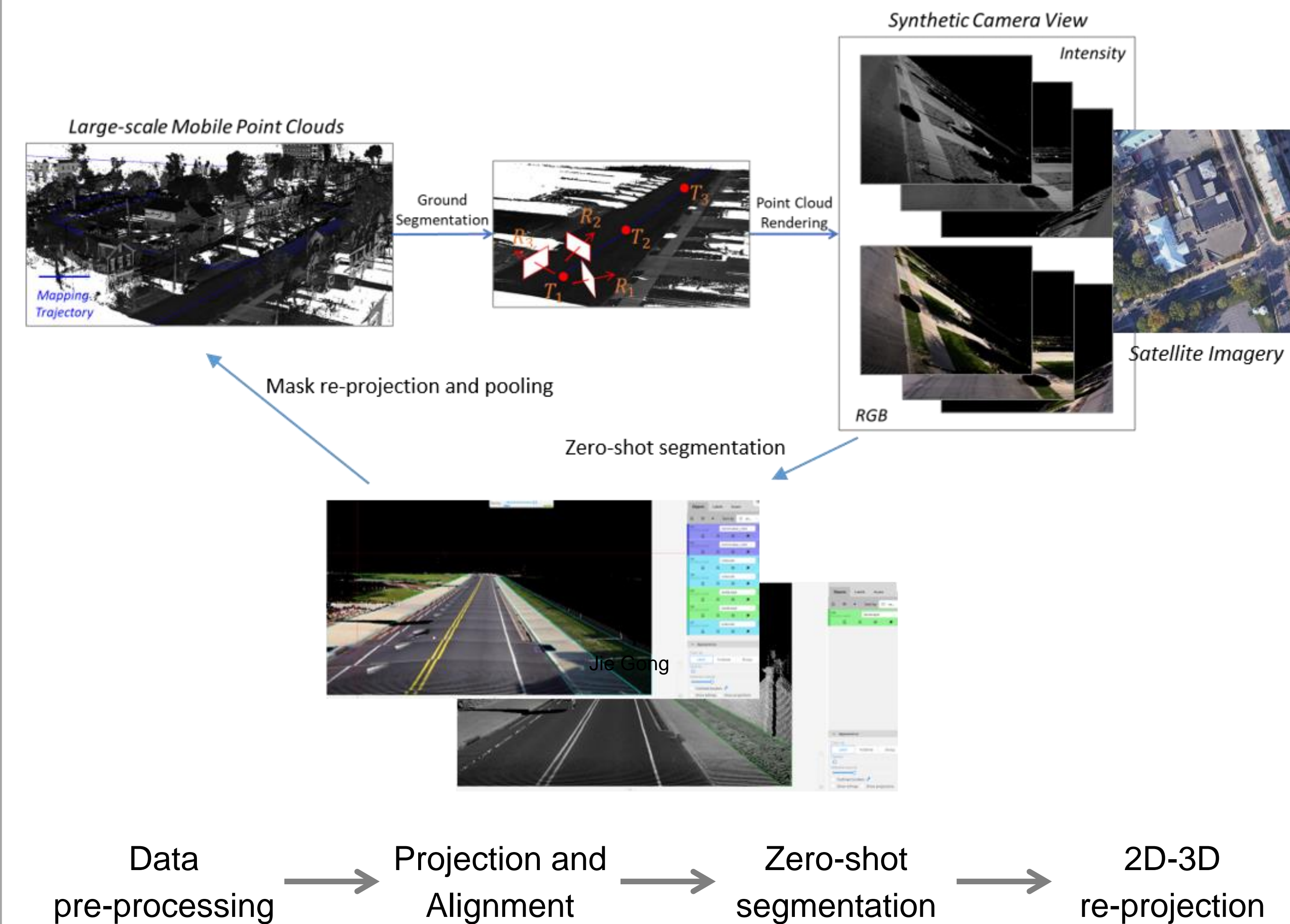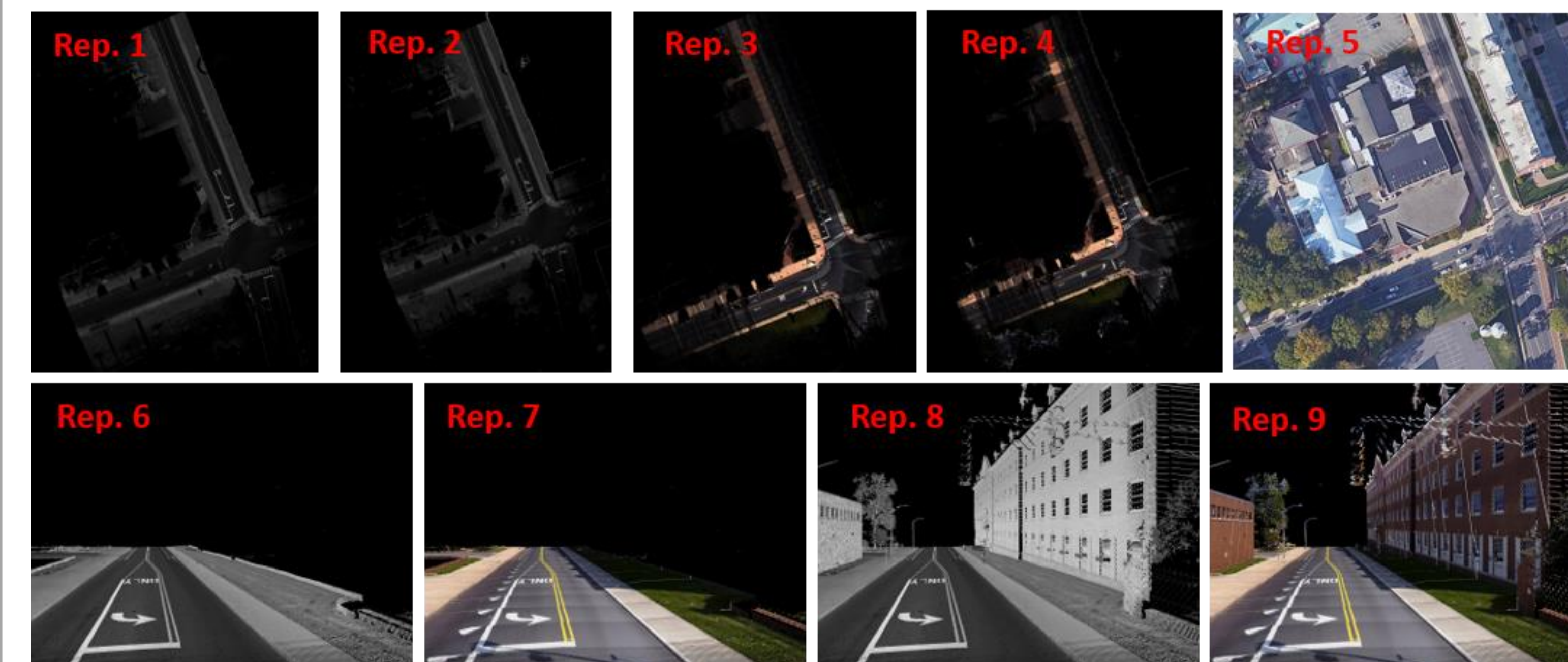
☐ *Image Representation*

Rep. 1 – BEV with only ground class and intensity values, Rep. 2 – BEV with all classes and intensity values, Rep. 3 – BEV with only ground class and color values, Rep. 4 – BEV with all classes and color values, Rep. 5 – Satellite imagery, Rep. 6 – Street view with only ground class and intensity values, Rep. 7 – Street view with only ground class and color values, Rep. 8 – Street view with all classes and intensity values, and Rep. 9 – Street view with all classes and color values.

RESULTS AND DISCUSSIONS

☐ *Evaluation of different pedestrian infrastructure features using the SAM.* (no extraction-N, partial extraction-P, complete extraction-C)

	Pedestrian Infrastructure Features	Rep. 1	Rep. 2	Rep. 3	Rep. 4	Rep. 5	Rep. 6	Rep. 7	Rep. 8	Rep. 9
Polygon features	Sidewalk	P	P	P	P	P	C	C	C	C
	Crosswalk	C	C	C	C	C	C	C	C	C
	Curb ramp	P	P	P	P	N	C	C	C	C
	Landscape	P	P	P	P	P	C	C	P	P
	Stair	N	N	P	P	N	N	N	C	C
	Detectable warning surface	N	C	N	C	N	C	C	C	C
	Traffic barrier	N	N	N	N	N	N	N	C	C
	Retaining wall	N	N	N	N	N	C	C	C	C
	Tree Canopy	N	N	P	P	P	N	N	P	P
Point features	Bench	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Bollard	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Fire hydrant	N	N	N	N	N	N	N	C	C
	Mailbox	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Manhole cover	N	N	N	N	C	C	C	C	C
	Memorial	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Phone booth	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Parking meter	N	N	N	N	C	C	C	C	C
	Post	N	N	N	N	N	N	N	C	C
	Public Sculpture	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Public vending machine	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Storm water inlet	N	N	N	N	C	C	C	C	C
	Tree trunk	N	N	N	N	N	N	N	C	C
	Waste container	N	N	N	N	N	N	N	C	C

☐ *visualization of the segmentation masks generated by SAM*
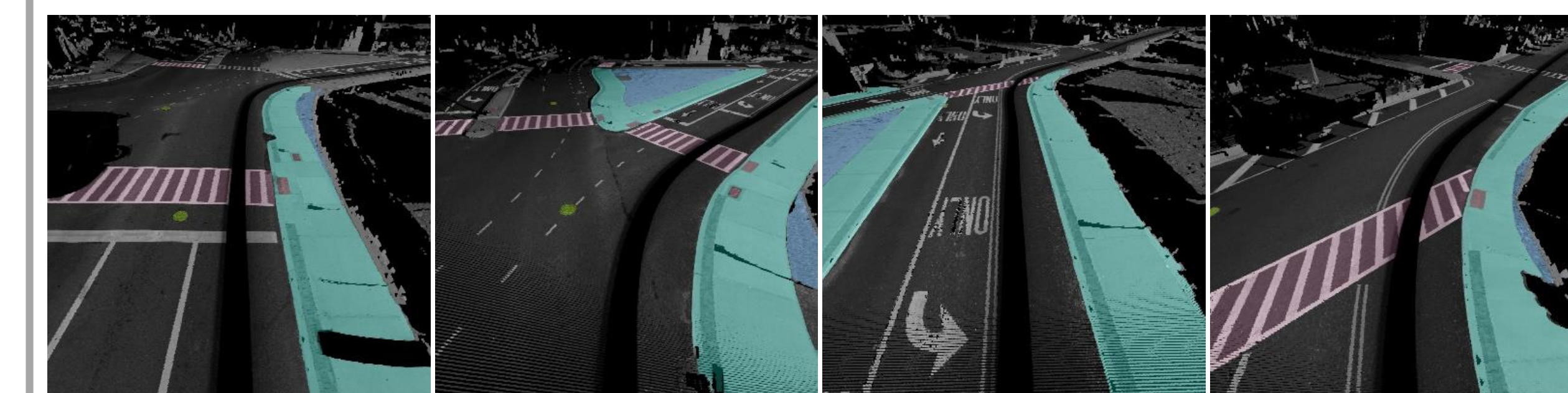
☐ *Geometry characteristics of the sidewalks*