Supplementary Material – LiDARsim: Realistic LiDAR Simulation by Leveraging the Real World

Sivabalan Manivasagam[1,2] Shenlong Wang[1,2] Kelvin Wong[1,2] Wenyuan Zeng[1,2]
Mikita Sazanovich[1] Shuhan Tan[1] Bin Yang[1,2] Wei-Chiu Ma[1,3] Raquel Urtasun[1,2]
[1]Uber Advanced Technologies Group [2]University of Toronto
[3]Massachusetts Institute of Techonology
{manivasagam, slwang, kelvin.wong, wenyuan, sazanovich, shuhan, byang10, weichiu, urtasun}@uber.com

Abstract

*In this supplementary, we cover additional details of LiDARsim. In Sec. 1 we provide dynamic object visualizations to highlight the diversity of our vehicles collected from real data. We also show some of the dynamic objects generated from KITTI (Sec. 1). Then in Sec. 2.1 we perform a quantitative analysis of the raydrop network to better understand why it helps bridge the real2sim gap. Finally, in Sec 2.2 we provide additional details and results on the safety-critical scenario evaluation using LiDARsim and the modifications we did on the Neural Motion Planner [2]. Additionally, we include a video (***submission612.mp4***) showcasing (1) our full LiDARSIM pipeline, (2) the application of LiDARsim to unknown object detection and end-to-end safety case evaluation.*

1. Dynamic Objects

To highlight the expressiveness of our dynamic object collection compared to using a set of CAD models, we showcase approximately 1000 dynamic objects in Figures 1 - 3. This has been captured by driving around several north-american cities. Figure 4 shows a selection of 36 vehicles generated from the KITTI public dataset [1]. We use a similar process as described in the main paper to generate KITTI dynamic objects. However, a difficulty involved with KITTI objects is that the ground-truth bounding box labels are only provided for vehicles in the camera view. This means that using only GT bounding box labels, we can only aggregate LiDAR on the objects when they are in camera view, which means the meshes generated are partial. To overcome this and obtain complete dynamic object shapes that have all their sides, we estimate the relative velocity of the object with respect to the ego-vehicle based on consecutive bounding boxes and extrapolate the last ground-truth bounding box's location into the future for 5 seconds. We then leverage hierarchical color-based ICP to aggregate and align the LiDAR and do a manual check to ensure the generated objects look reasonable.

2. Additional Results

2.1. Measuring Realism of LiDAR Simulation

We now demonstrate that LiDARsim is comparable to real LiDAR. We cover LiDAR statistics on simulated vehicles compared to real vehicles, an indicator for realism in downstream tasks. Fig. 5 shows cumulative counts of objects based on number of LiDAR points that hit the dynamic object and a qualitative example of LiDARsim vs. real LiDAR (Note that LiDARsim may have different dynamic objects in the scene due to selection from our object bank). Learning raydrop results in simulated LiDAR that better matches with the real distribution compared to standard raycasting. Without raydrop, raycasting overestimates the number of points that hit a vehicle because it does not account for higher order physics (e.g., non-lambertian reflectance, far range). Learning raydrop helps better match the real data.

2.2. Neural Motion Planner Modifications and Results

We now evaluate the effect of perception on the end-metric performance of the autonomy system: safety. We evaluate an enhanced neural motion planner's (NMP) [2] ability to maneuver safety-critical scenarios. Our version of NMP is modified

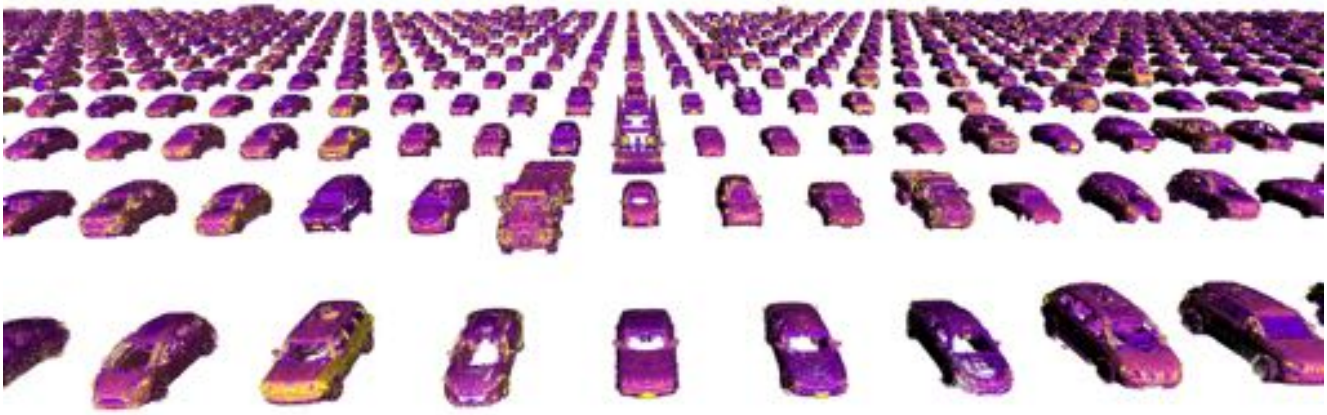

Figure 1: Selection of dynamic objects colored by intensity

to constrain trajectory samples to lie on the road and to not collide with other vehicles' predicted trajectories. We take the safety-critical scenario described in Fig. 6 and generate 110 scenarios of the test case in geographic areas in different cities and traffic configurations. To understand the safety buffers of NMP, we vary the initial velocity and trigger time of the occluded vehicle entering the SDV lane. See our video **submission612.mp4** for qualitative results. Fig. 7 shows average success rate at different conditions. On average, the NMP successfully avoids colliding with the vehicle turning into the lane 90% of the time across different initial velocity and trigger time variations across the different scenarios. We are able to identify at what point do higher speeds and earlier trigger times causes the planner to less successfully avoid the incoming vehicle.

References

[1] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361. IEEE, 2012. 1

[2] W. Zeng, W. Luo, S. Suo, A. Sadat, B. Yang, S. Casas, and R. Urtasun. End-to-end interpretable neural motion planner. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8660–8669, 2019. 1

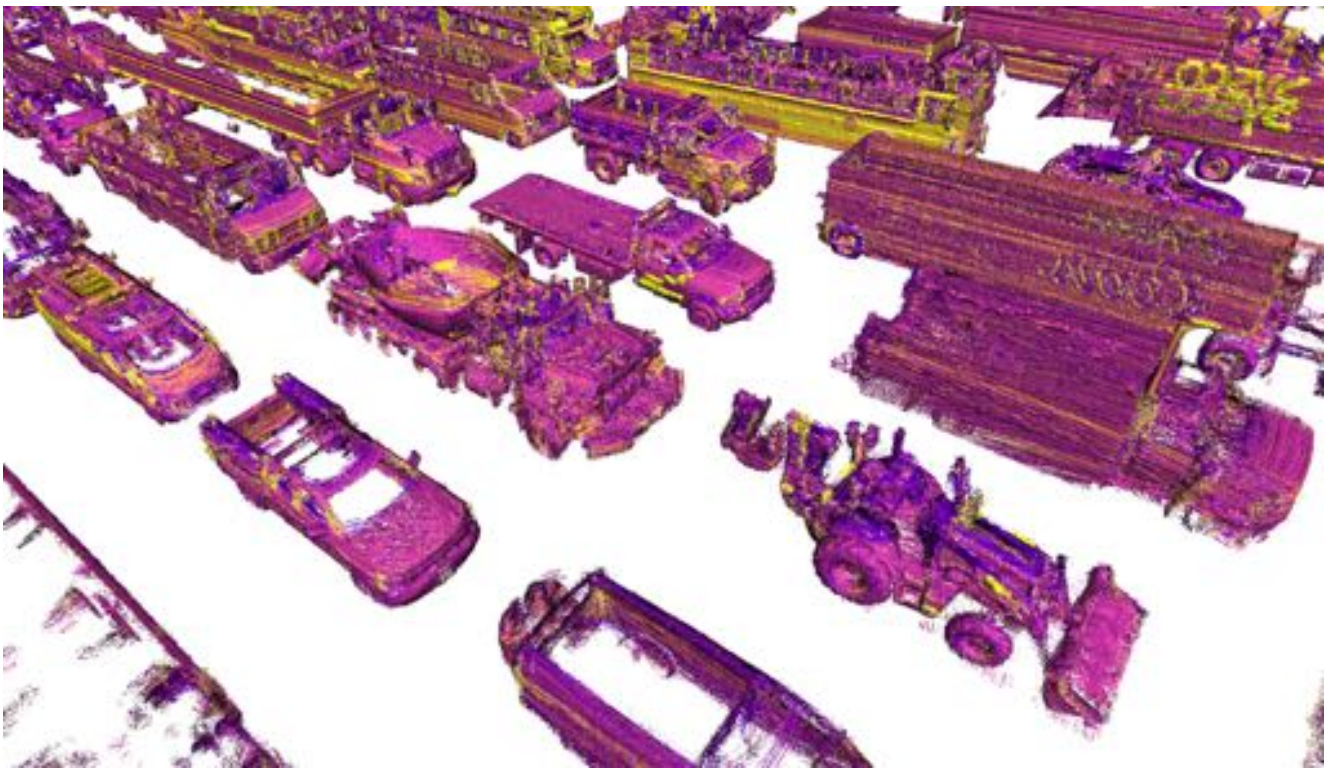

Figure 2: Interesting dynamic objects colored by intensity, showcasing diversity (construction vehicles, 16-wheeler)

Figure 3: Interesting dynamic objects colored by intensity, showcasing text

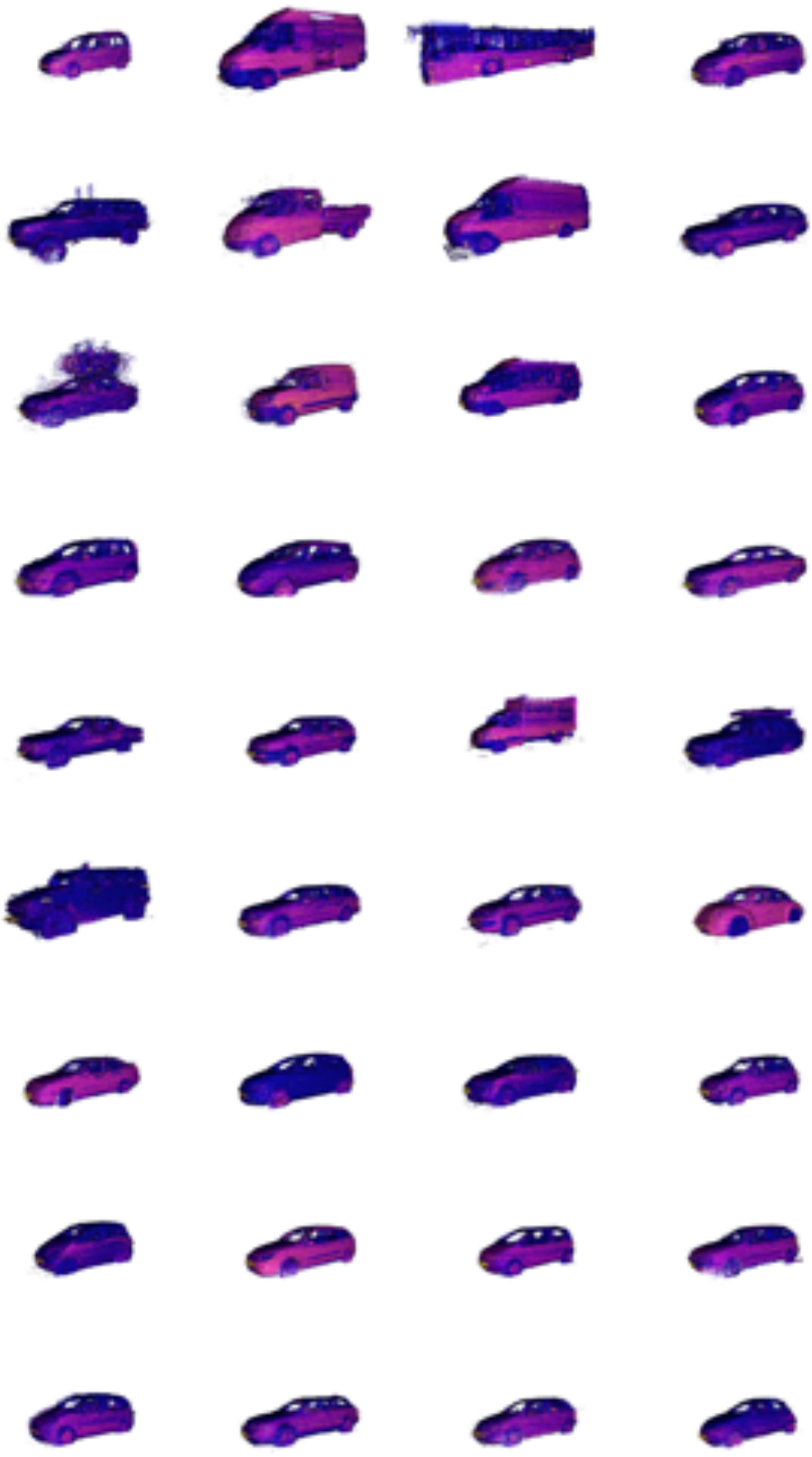

Figure 4: KITTI dynamic objects colored by intensity

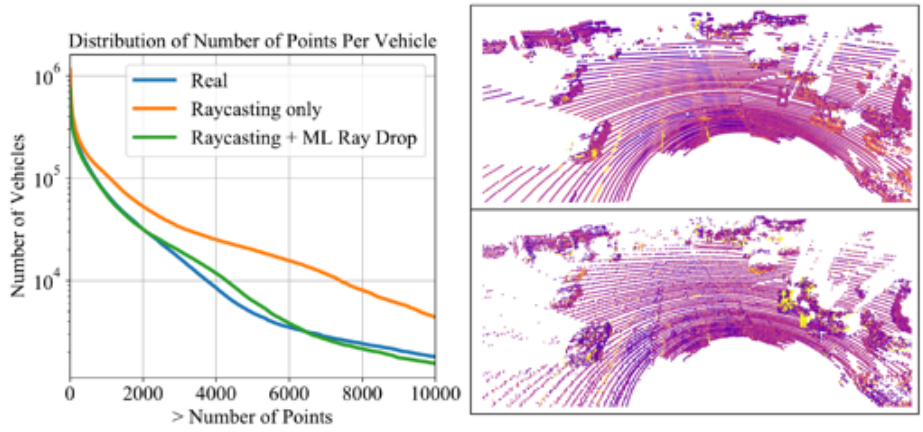

Figure 5: Left: Points per vehicle for different LiDAR.
Top Right: Real LiDAR. Bot. Right: LiDARsim LiDAR

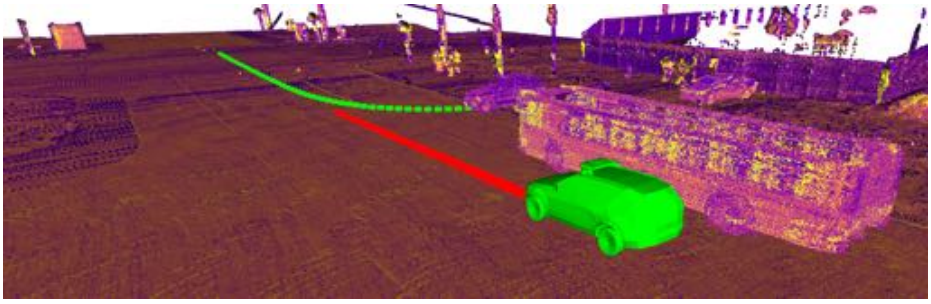

Figure 6: Safety-critical scenario: Occluded vehicle entering the SDV lane

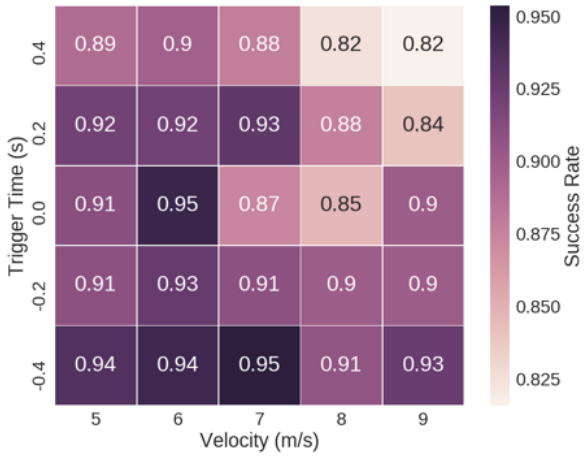

Figure 7: NMP Results (Success Rate) on safety-critical scenario