NeuroNCAP: Photorealistic Closed-loop Safety Testing for Autonomous Driving
Supplementary Material

William Ljungbergh[*,1,2], Adam Tonderski[*,1,3], Joakim Johnander[1], Holger Caesar[4], Kalle Åström[3], Michael Felsberg[2], and Christoffer Petersson[1]

[1] Zenseact, Sweden
[2] Linköping Univeristy, Sweden
[3] Lund University, Sweden
[4] Delft University of Technology, Netherlands

A Further Scenario Details

We base our stationary scenarios on the following 10 NuScenes validation set sequences: `0099, 0101, 0103, 0106, 0108, 0278, 0331, 0783, 0796, 0966`. For each sequence we pick 1-4 actors that have been observed from a sufficiently nearby distance and determine a collision location somewhere along the ego-trajectory. We jitter this target position by 5-10m in the logitundinal direction; 0.5-1.5m in the lateral direction; and the yaw by 0.5-1.0 radians. Finally we determine a starting time of the scenario, so that a collision happens in slightly more than 4 seconds. See examples of each scenario in Fig. 1 and Fig. 2.

Frontal and side collisions follow a similar procedure, and use `0103, 0106, 0110, 0346, 0923` and `0103, 0108, 0110, 0278, 0921` respectively. We determine a trajectory that is either anti-parallel or perpendicular to the ego-trajectory, and time it so that a collision occurs unless an avoidance maneuver is performed. See examples of each frontal scenario in Fig. 3 and of side scenarios in Fig. 4. In cases where the target is not visible in the front camera we show the side camera instead.

B Qualitative Video Examples

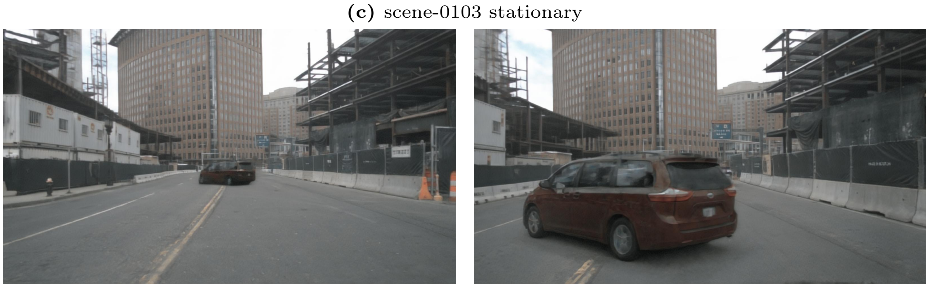
We provide visualizations of our photorealistic closed-loop safety testing with UniAD [1] in `qualitative_examples.mp4`. The visualization contains a visualization of the AD model output (left) and the rendered views for the front camera (right, top) as well as the front-left and front-right cameras (right, bottom). The AD model (here UniAD) output visualization shows the detected target actor, its possible future trajectories, and the plan in the form of a trajectory. For reference, ground truth objects are drawn as shaded gray boxes.

In the example videos, we see how the target actor is robustly and accurately detected. The views produced by the neural-based renderer sometimes contain clearly visible artifacts, especially on objects that are close. However, these artifacts do not seem to adversely affect the quality of the AD model detections.

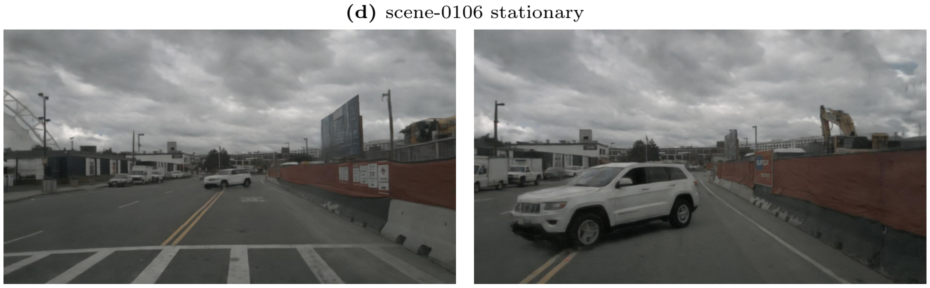
(a) scene-0099 stationary

(b) scene-0101 stationary

(c) scene-0103 stationary



(d) scene-0106 stationary

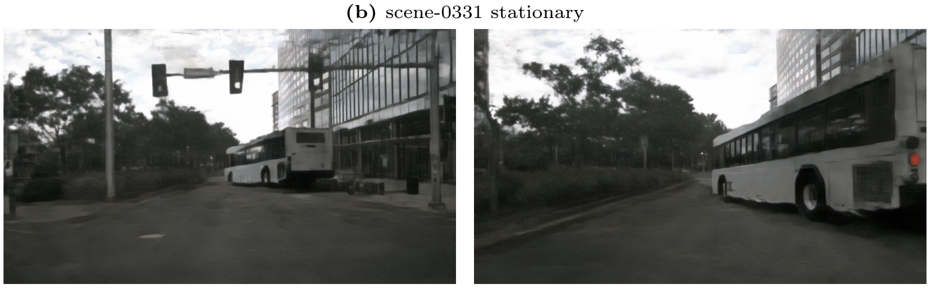


(e) scene-0108 stationary

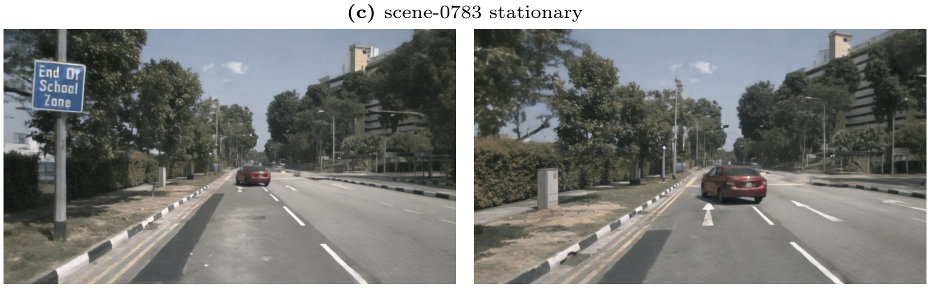
Fig. 1: One example from each stationary scenario (part 1).

(a) scene-0278 stationary

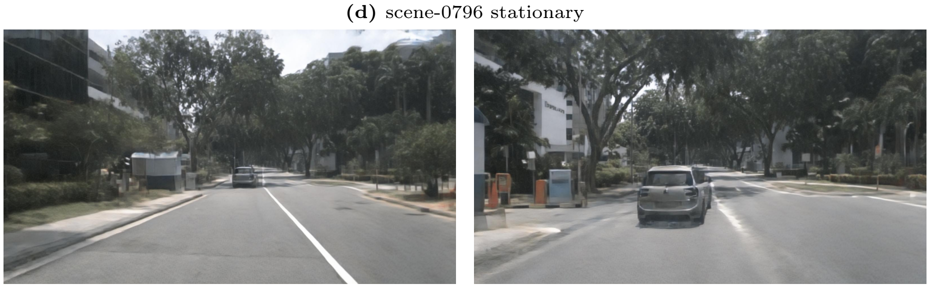
(b) scene-0331 stationary



(c) scene-0783 stationary



(d) scene-0796 stationary

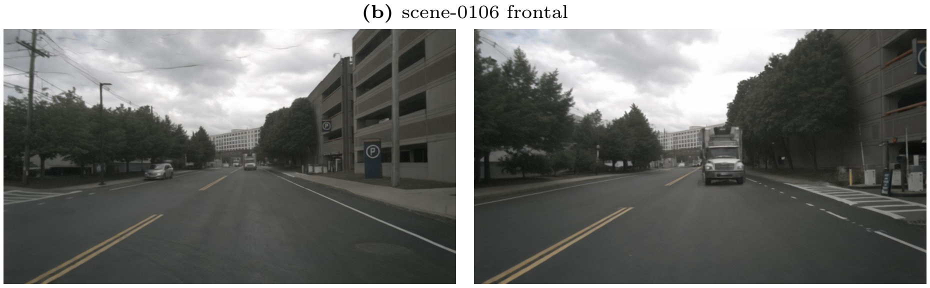


(e) scene-0966 stationary

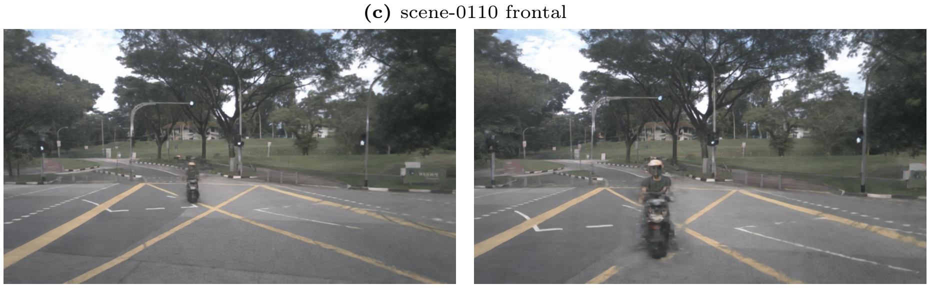
Fig. 2: One example from each stationary scenario (part 2).

(a) scene-0103 frontal

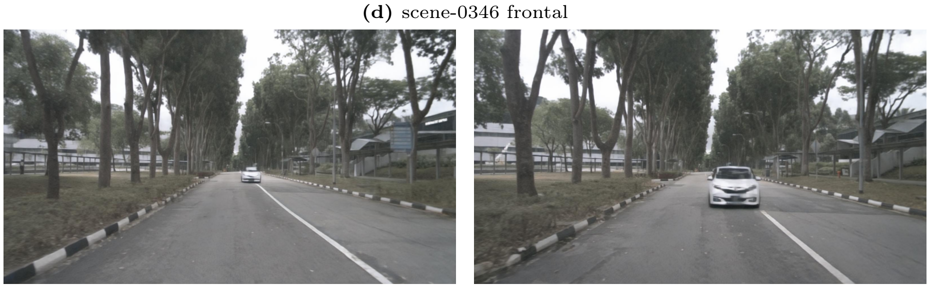
(b) scene-0106 frontal



(c) scene-0110 frontal



(d) scene-0346 frontal



(e) scene-0923 frontal

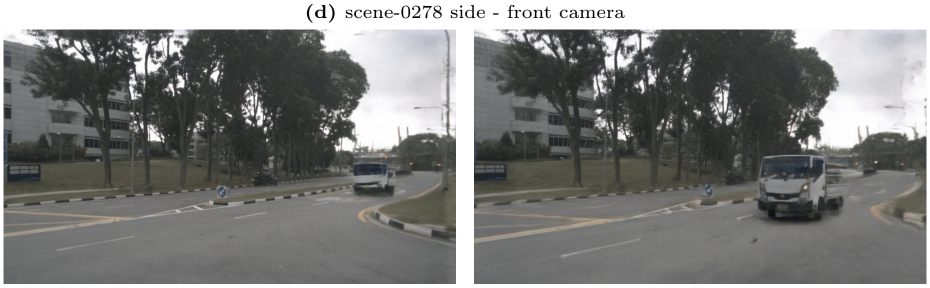
Fig. 3: One example from each frontal scenario.

(a) scene-0103 side - front camera

(b) scene-0108 side - side camera

(c) scene-0110 side - front camera

(d) scene-0278 side - front camera



(e) scene-0921 side - right camera

Fig. 4: One example from each side scenario.

The planned trajectory on the other hand, tends to pass straight through the target agent. This behavior is consistent over time – the AD model detects the target actor but decides to drive through it – and does not change as we get closer to the target actor. This results in severe collisions.

References

1. Hu, Y., Yang, J., Chen, L., Li, K., Sima, C., Zhu, X., Chai, S., Du, S., Lin, T., Wang, W., et al.: Planning-oriented autonomous driving. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17853–17862 (2023) 1