Supplementary Material: Monocular 3D Object Detection Leveraging Accurate Proposals and Shape Reconstruction

Jason Ku* Alex D. Pon* Steven L. Waslander
University of Toronto
{kujason.ku, alex.pon}@mail.utoronto.ca, stevenw@utias.utoronto.ca

1. Additional Qualitative Results

In Fig. 1, we show additional detection results on several scenes in the KITTI [2] validation split, *val1* [1]. Fig. 2 shows common failure cases for the network which are noted when there is heavy occlusion or truncation of the object. Additional results on several KITTI driving sequences can be found at https://youtu.be/_iJpEpXB7j4

Figure 1. **Additional Qualitative Results:** 2D detections (top) are shown in **orange**. 3D detections in **green** are shown projected into the image (top) and in the 3D scene (bottom). Ground truth 3D boxes (bottom) are shown in **red**. Points within the detection boxes are the estimated point clouds from the network, while the background points are taken from the colorized interpolated LiDAR scan.

*Equal contribution.

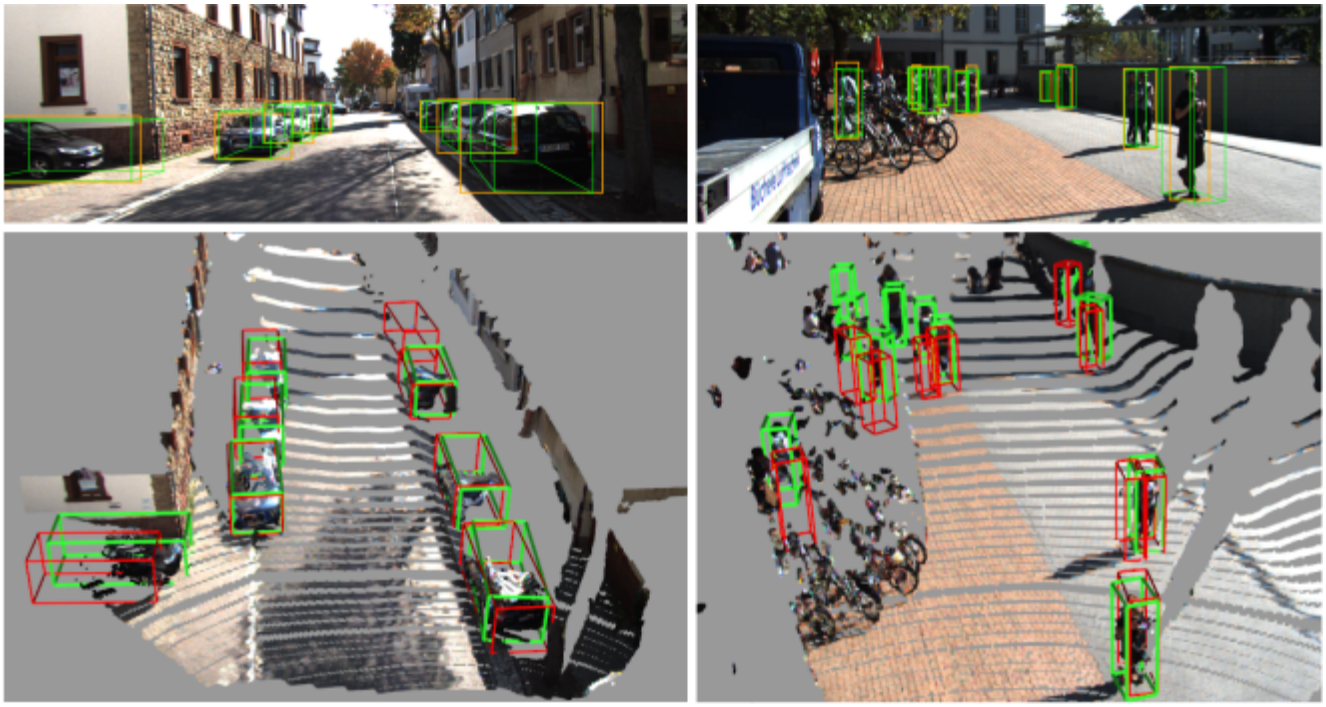

Figure 2. **Common Failure Cases:** Truncation and occlusion are two common causes of localization errors. Truncation is evident for the far left car in the first image, and occlusion is common in large groups of pedestrians walking together as shown in the second image. 2D detections (top) are shown in **orange**. 3D detections in **green** are shown projected into the image (top) and in the 3D scene (bottom). Ground truth 3D boxes (bottom) are shown in **red**. Points within the detection boxes are the estimated point clouds from the network, while the background points are taken from the colorized interpolated LiDAR scan.

References

[1] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia. Multi-view 3d object detection network for autonomous driving. In *CVPR*, 2017. 1

[2] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *CVPR*, 2012. 1