3D Object Proposals using Stereo Imagery for Accurate Object Class Detection

Xiaozhi Chen*, Kaustav Kundu*, Yukun Zhu, Huimin Ma, Sanja Fidler and Raquel Urtasun

APPENDIX

PROPOSAL RECALL

In this appendix we show more results of the proposals by evaluating 2D and 3D bounding box recall across all categories and difficulty regimes.

Recall vs Distance: Fig. 1 shows 2D bounding box recall as a function of the object distance with 2000 proposals. In all three difficulty regimes, we observe only very small decrease of our method in recall with the distance of objects from the camera increasing. On the contrary, the recall of all other methods drop significantly for objects at large distance.

3D bounding box Recall: We plot 3D box recall of our proposals in Fig. 2. When evaluated on easy data, the class-dependent 3DOP achieves more than 98%, 90% and 80% 3D recall for Car, Pedestrian and Cyclist, respectively. Recall on moderate and hard data are very similar.

Stereo vs LIDAR: We also show 2D and 3D box recall of our proposals when applied to stereo and LIDAR data respectively. For 3D box recall shown in Fig. 3, LIDAR point cloud clearly outperforms stereo on Pedestrian and Cyclist, while being similar with stereo on Car. This is reasonable as LIDAR point cloud has much more precise depth and thus doing better on detecting small objects. For 2D box recall shown in Fig. 4, stereo and LDIAR achieve very close recall on Pedestrian and Cyclist. The stereo approach has slightly higher recall on Car. This suggests that stereo can work very well on large objects like cars, while LIDAR is superior for small objects.

REFERENCES

• * Denotes equal contribution.
Fig. 1: 2D bounding box Recall vs Distance with 2000 proposals. We use overlap threshold of 0.7 for Car, and 0.5 for Pedestrian, Cyclist.
Fig. 2: 3D bounding box Recall vs #Candidates. IoU threshold is set to 0.25.
Fig. 3: 3D bounding box Recall vs #Candidates at IoU threshold of 0.25 for different approaches.
Fig. 4: 2D bounding box Recall vs #Candidates for different approaches. We use an overlap threshold of 0.7 for Car, and 0.5 for Pedestrian and Cyclist.