Subcategory-aware Convolutional Neural Networks for Object Proposals and Detection

Yu Xiang[1], Wongun Choi[2], Yuanqing Lin[3] and Silvio Savarese[4]

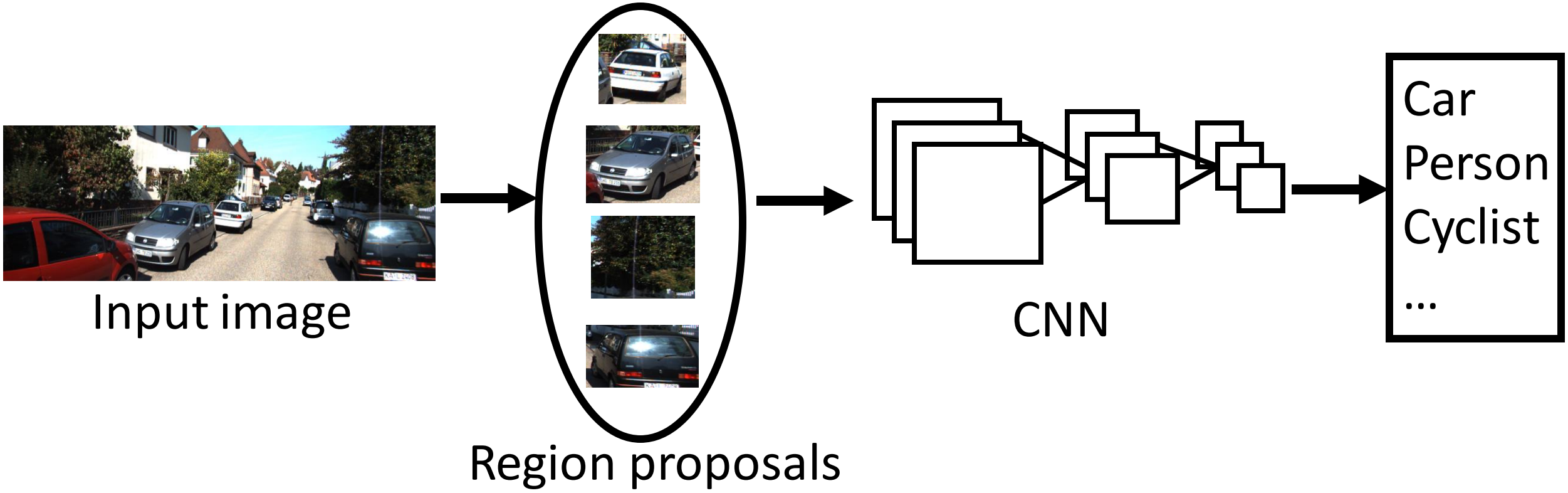
[1]University of Washington, [2]NEC Laboratories America, Inc., [3]Baidu, Inc., [4]Stanford University

CVGL Computational Vision & Geometry Lab

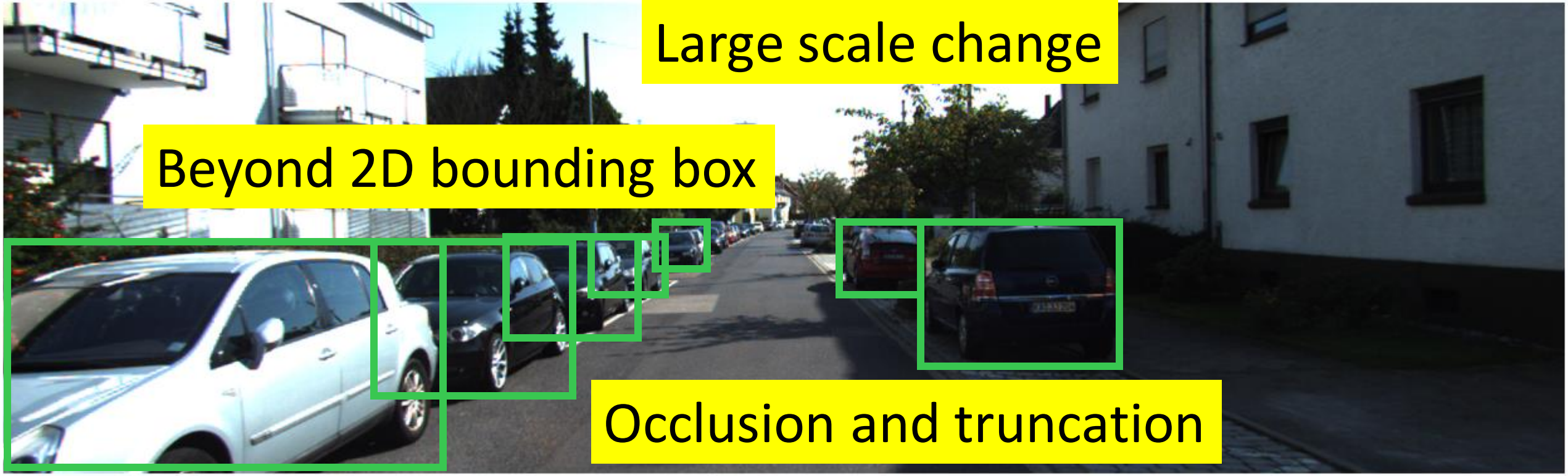
NEC Laboratories America

Relentless passion for innovation

Convolutional Neural Networks for Object Detection

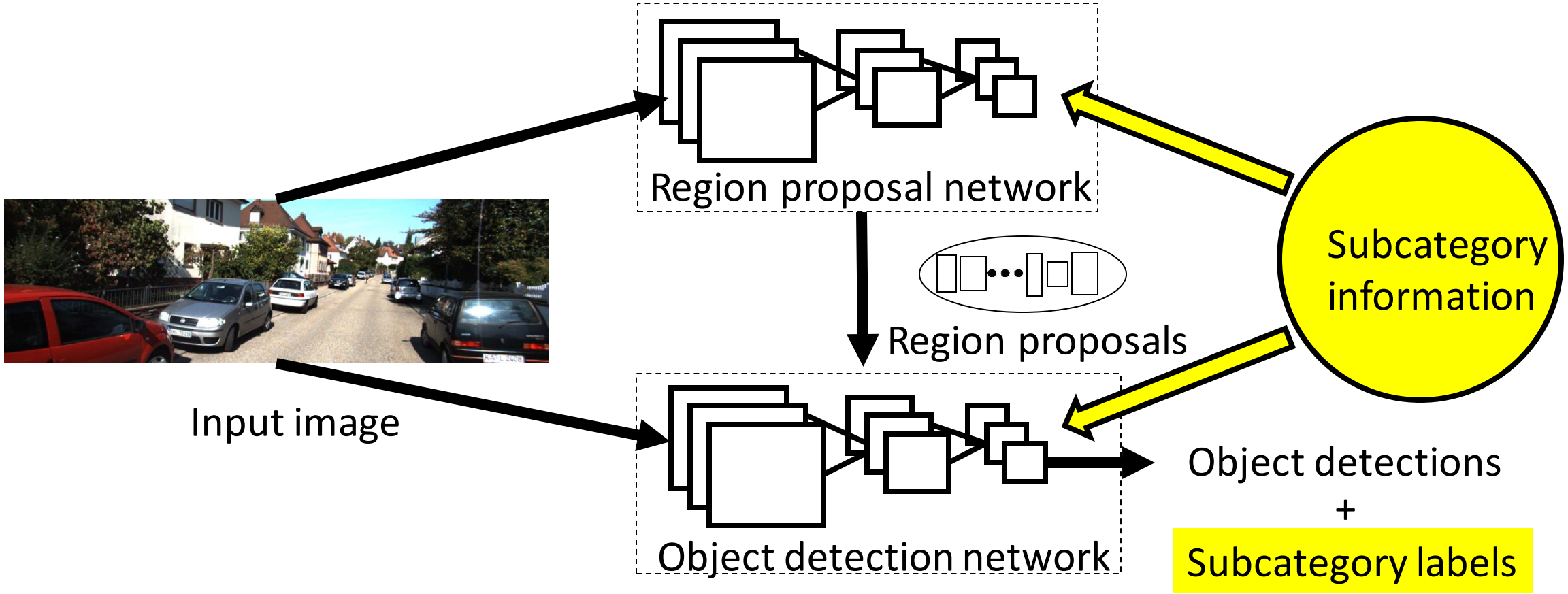


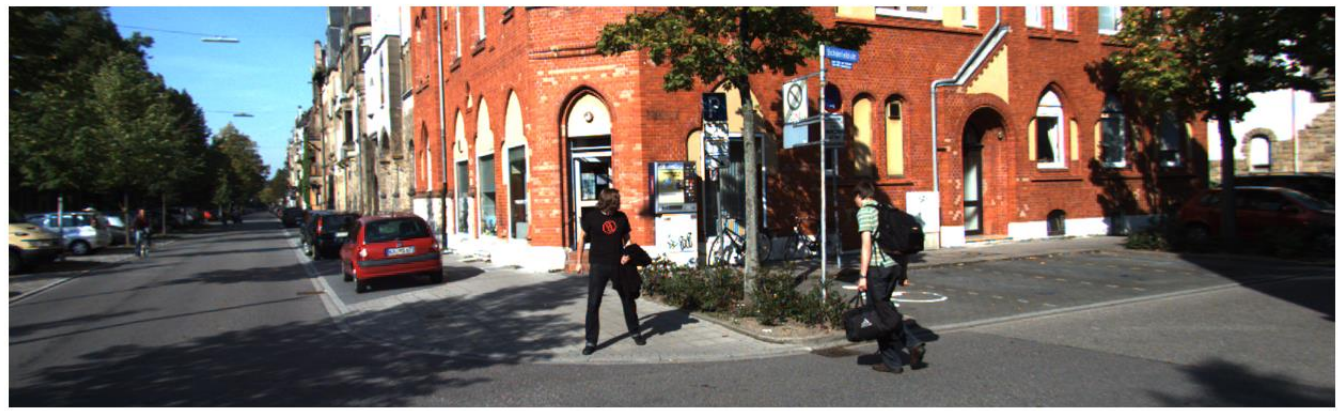
Challenges

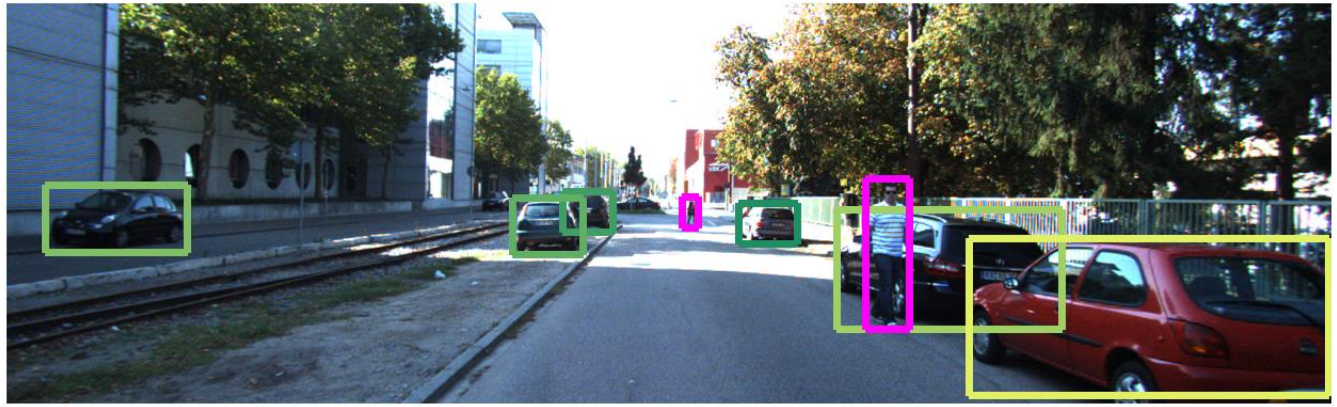


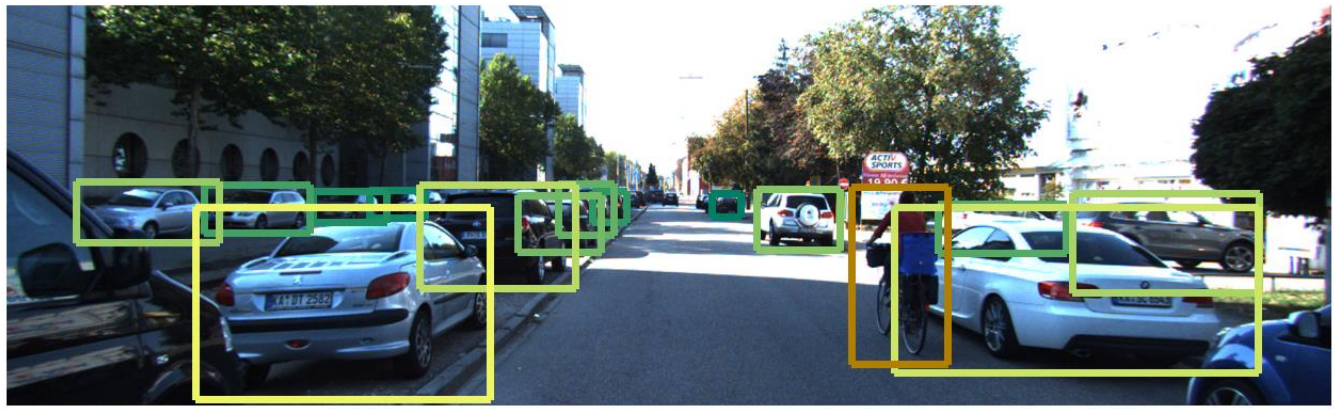
The image is from the KITTI detection benchmark (Geiger et al. CVPR'12)

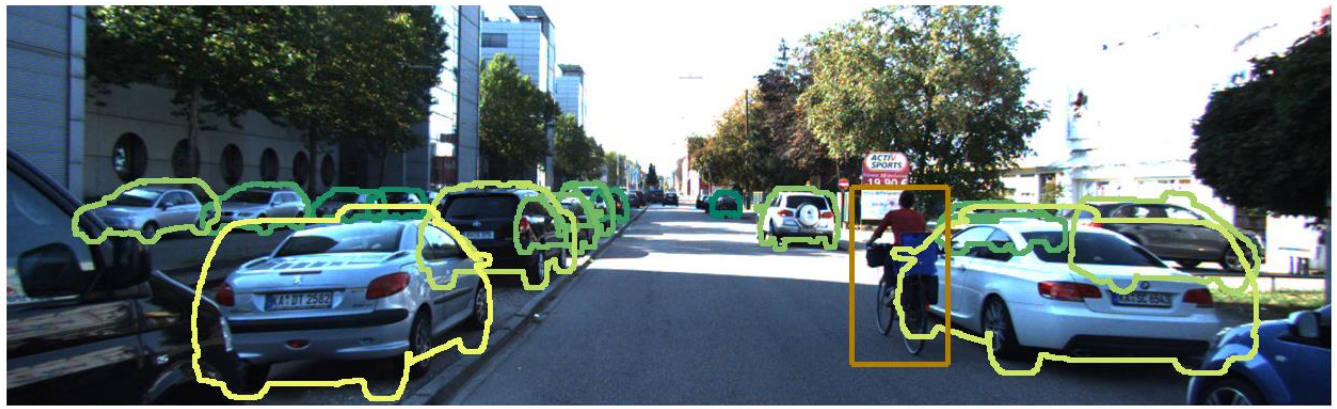
Our Work: Subcategory-aware CNNs

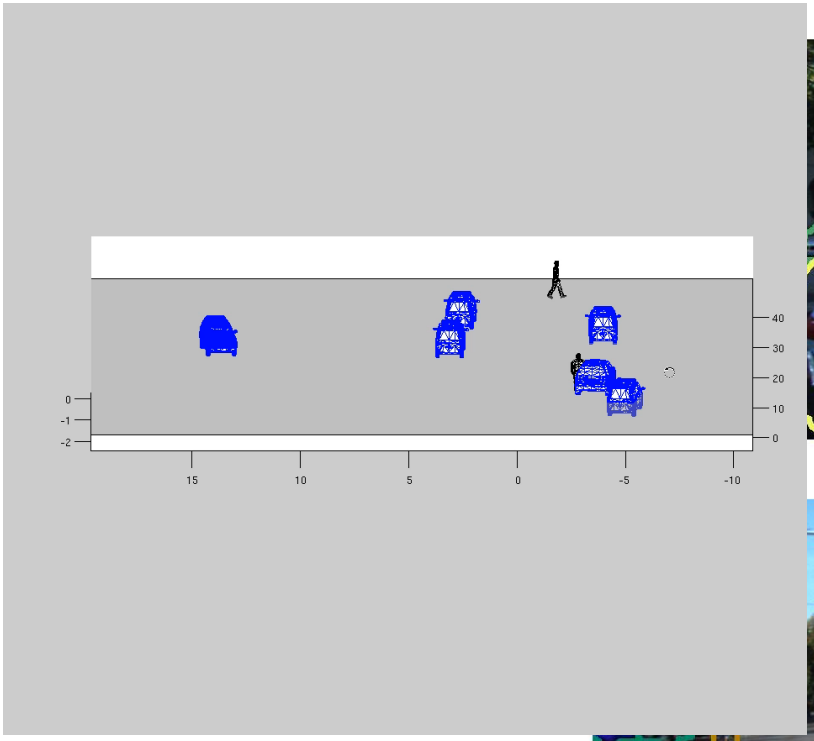


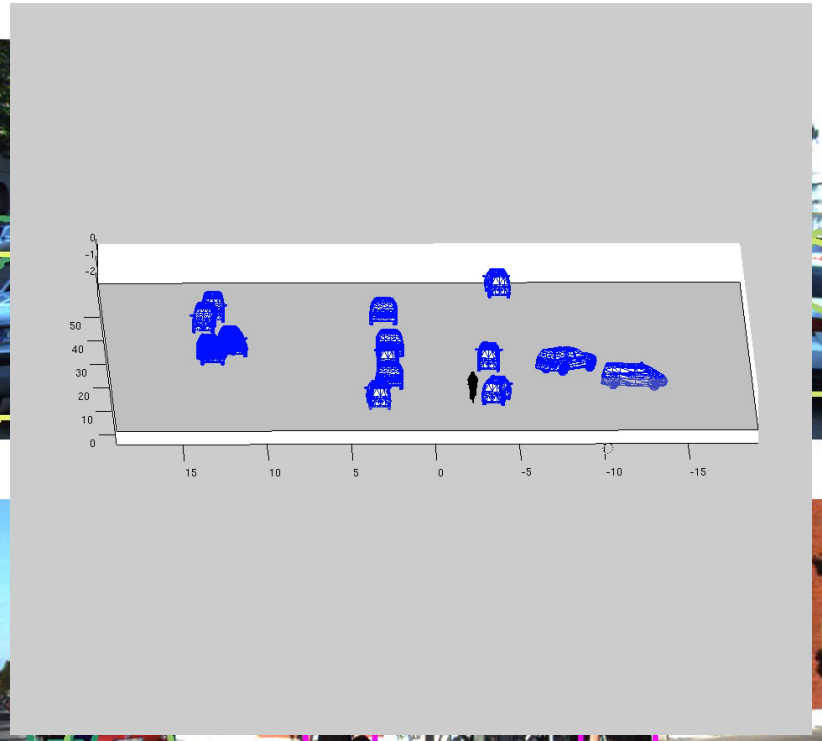


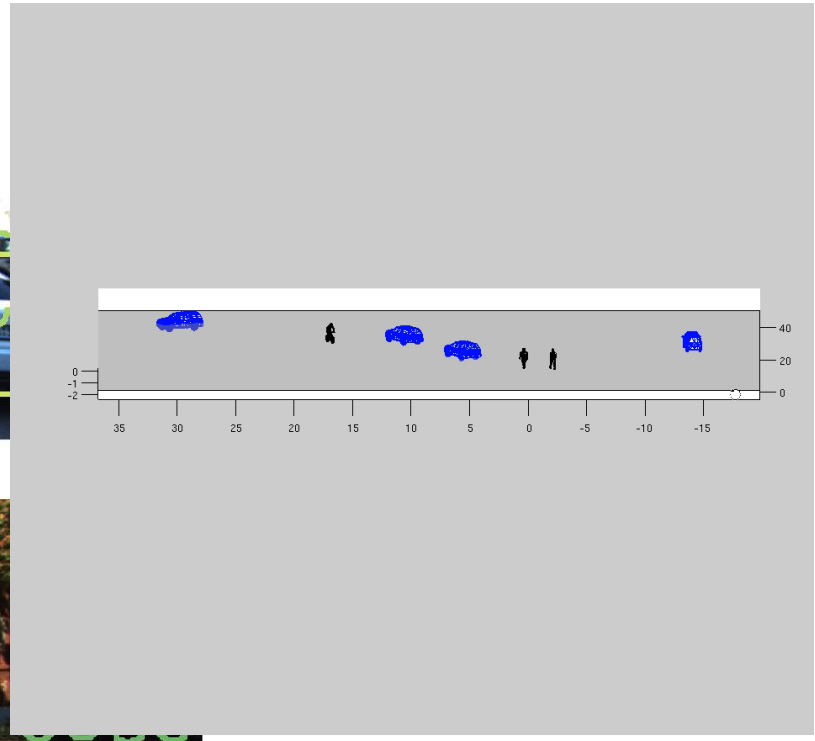







0
-1
-2
15
10
5
0
-5
-10
40
30
20
10
0

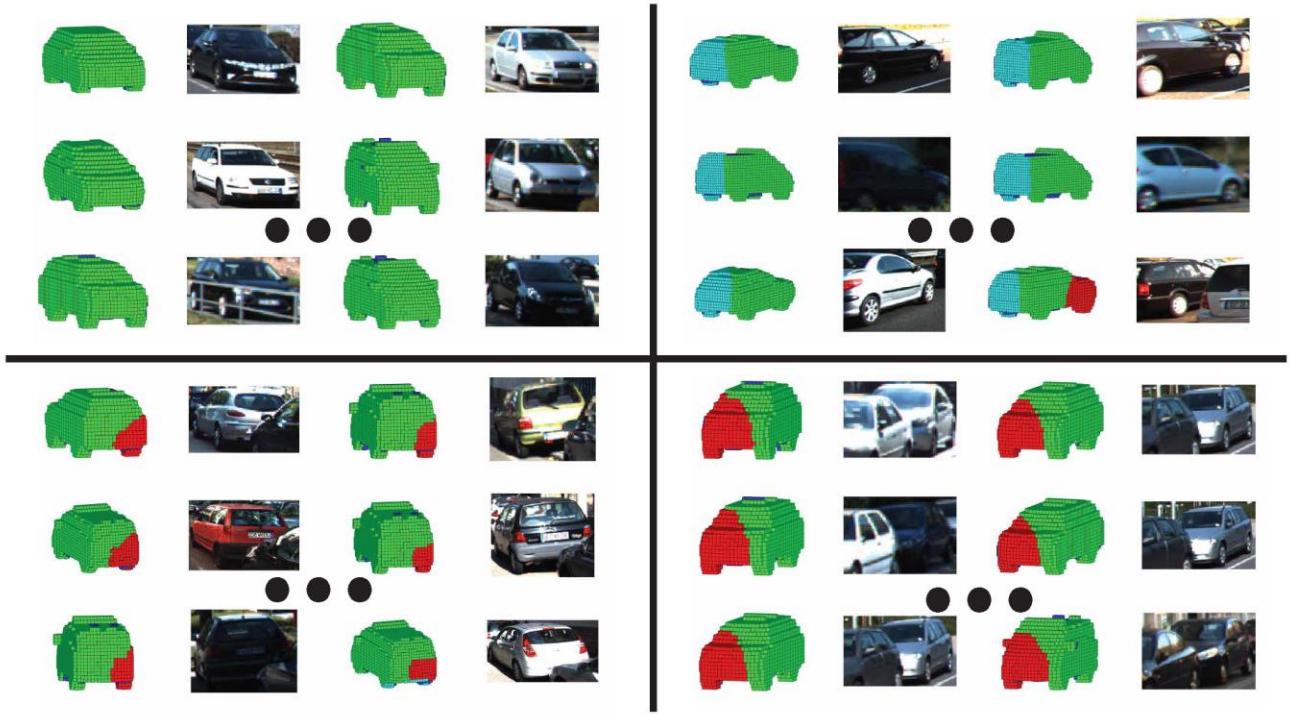

0
-1
-2
50
40
30
20
10
0
15
10
5
0
-5
-10
-15


40
20
0
0
-1
-2
35
30
25
20
15
10
5
0
-5
-10
-15

Subcategories

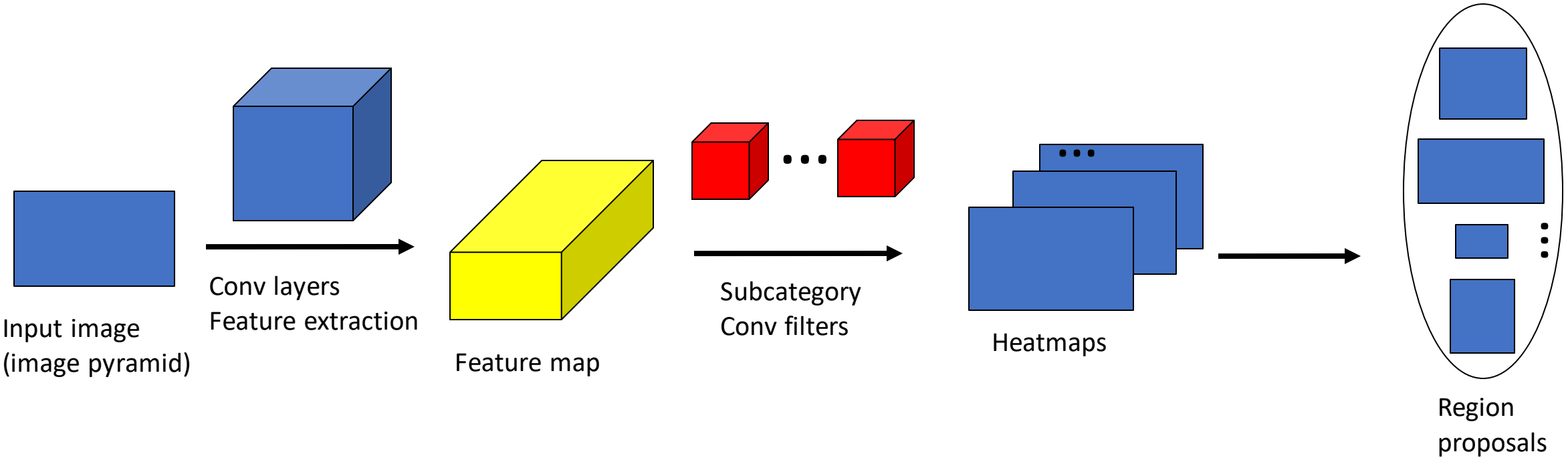
- Subcategory is a general concept.

- 3D Voxel Pattern (3DVP, Xiang et al., CVPR'15)

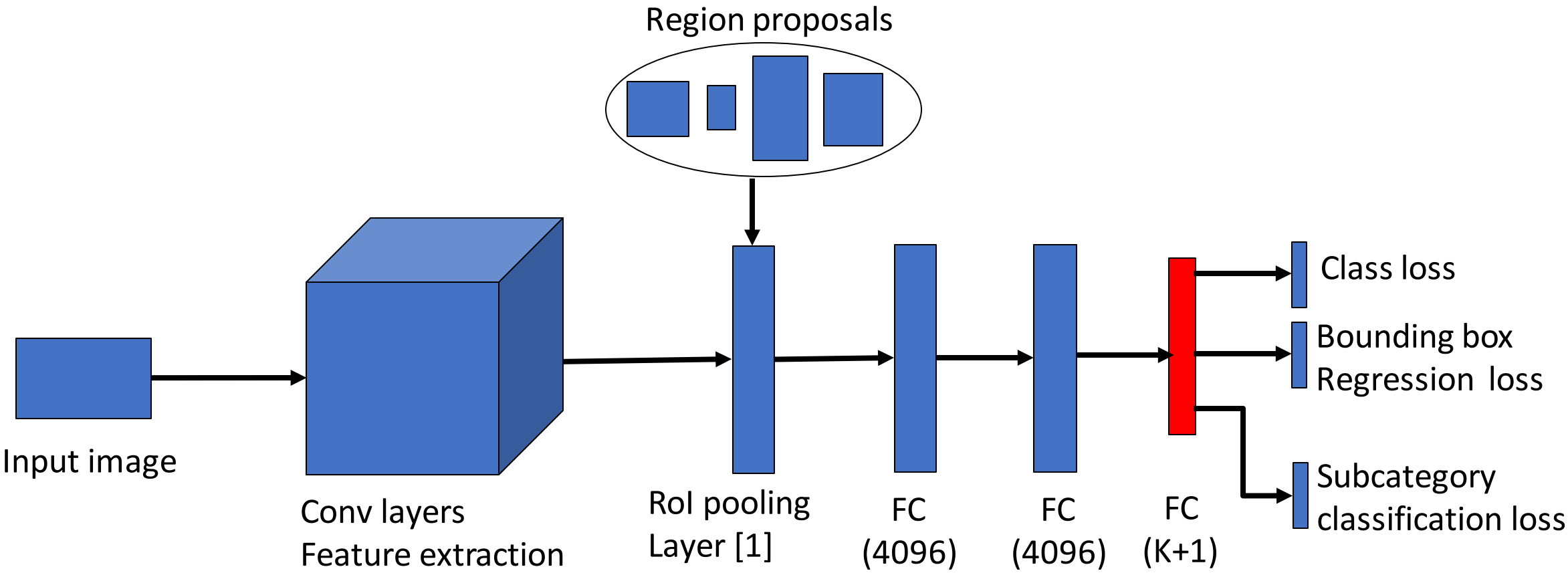


Cluster objects with similar 3D pose, occlusion and truncation.

Subcategory-aware Region Proposal Network



Subcategory-aware Detection Network



[1] R. Girshick. Fast R-CNN. ICCV, 2015.

Car Detection and Orientation Estimation on KITTI

	Object Detection (AP)			Object Detection and Orientation estimation (AOS)		
Method	**Easy**	**Moderate**	**Hard**	**Easy**	**Moderate**	**Hard**
ACF [1]	55.89	54.77	42.98	N/A	N/A	N/A
DPM-VOC+VP [2]	74.95	64.71	48.76	72.28	61.84	46.54
OC-DPM [3]	74.94	65.95	53.86	73.50	64.42	52.40
SubCat [4]	84.14	75.46	59.71	83.41	74.42	58.83
Regionlets [5]	84.75	76.45	59.70	N/A	N/A	N/A
3DVP [6]	84.81	73.02	63.22	84.31	71.99	62.11
3DOP [7]	**93.04**	88.64	79.10	**91.44**	86.10	76.52
Mono3D [8]	92.33	88.66	78.96	91.01	86.62	76.84
SDP+RPN [9]	90.14	88.85	78.38	N/A	N/A	N/A
MS-CNN [10]	90.03	89.02	76.11	N/A	N/A	N/A
Ours SubCNN	90.81	**89.04**	**79.27**	90.67	**88.62**	**78.68**

[1] P. Dolla´r, R. Appel, S. Belongie, and P. Perona. Fast feature pyramids for object detection. TPAMI, 2014.
[2] B. Pepik, M. Stark, P. Gehler, and B. Schiele. Multi-view and 3d deformable part models. TPAMI, 2015.
[3] B. Pepikj, M. Stark, P. Gehler, and B. Schiele. Occlusion patterns for object class detection. In CVPR, 2013.
[4] E. Ohn-Bar and M. M. Trivedi. Learning to detect vehicles by clustering appearance patterns. T-ITS, 2015.
[5] X. Wang, M. Yang, S. Zhu, and Y. Lin. Regionlets for generic object detection. In ICCV, 2013.
[6] Y. Xiang, W. Choi, Y. Lin, and S. Savarese. Data-driven 3d voxel patterns for object category recognition. In CVPR, 2015.
[7] X. Chen, K. Kundu, Y. Zhu, A. G. Berneshawi, H. Ma, S. Fidler, and R. Urtasun. 3d object proposals for accurate object class detection. In NIPS, 2015.
[8] X. Chen, K. Kundu, Z. Zhang, H. Ma, S. Fidler, R. Urtasun. Monocular 3D Object Detection for Autonomous Driving, in CVPR, 2016.
[9] F. Yang, W. Choi, and Y. Lin. Exploit all the layers: Fast and accurate cnn object detector with scale dependent pooling and cascaded rejection classifiers. In CVPR, 2016.
[10] Z. Cai, Q. Fan, R. Feris, and N. Vasconcelos. A unified multi-scale deep convolutional neural network for fast object detection. In ECCV, 2016.

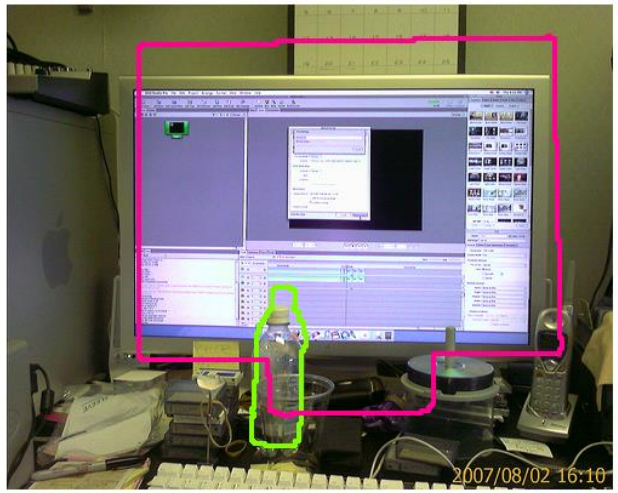
Detection: Rank 2
Pose : Rank 4

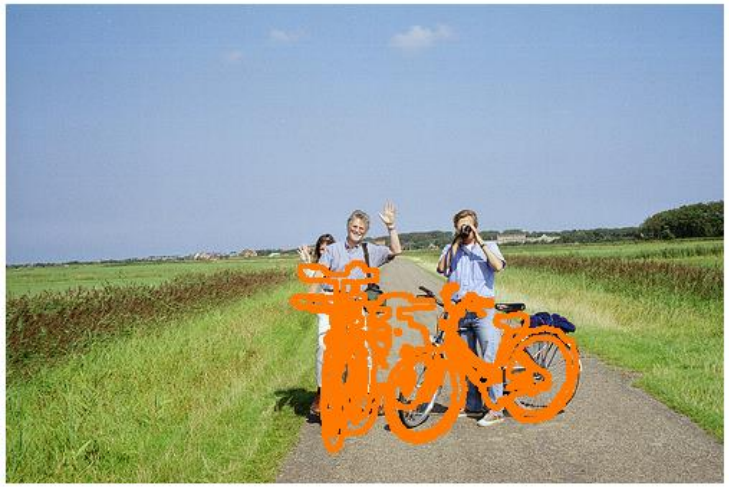
Detection and Pose Estimation on PASCAL3D+

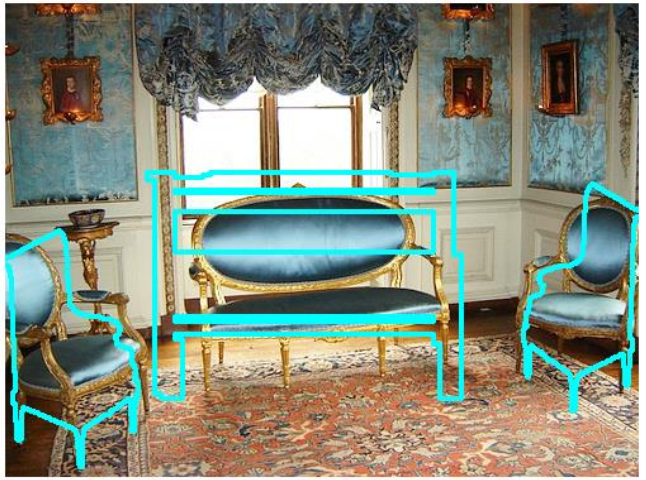
Method	Detection (AP)
DPM [1]	29.6
R-CNN [2]	56.9
Ours SubCNN	**60.7**

Method	4 Views (AVP)	8 Views (AVP)	16 Views (AVP)	24 Views (AVP)
VDPM [3]	19.5	18.7	15.6	12.1
DPM-VOC+VP [4]	24.5	22.2	17.9	14.4
Ours SubCNN	**47.5**	**31.9**	**24.5**	**19.3**

[1] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part-based models. TPAMI, 2010.
[2] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. arXiv preprint arXiv:1311.2524, 2013.
[3] Y. Xiang, R. Mottaghi, and S. Savarese. Beyond pascal: A benchmark for 3d object detection in the wild. In WACV, 2014.
[4] B. Pepik, M. Stark, P. Gehler, and B. Schiele. Multi-view and 3d deformable part models. TPAMI, 2015.


2007/08/02 16:10






48305

Conclusion

- A new network architecture for object proposal generation using subcategory information
- A new network for joint object detection and subcategory classification
- Our method improves over the state-of-the-art methods on both KITTI and PASCAL3D+.

Acknowledgements

Thank you!