Learning to Co-Generate Object Proposals with a Deep Structured Network Supplementary Material

Zeshan Hayder[1,2], Xuming He[2,1]
[1]Australian National University & [2]NICTA *
{zeeshan.hayder, xuming.he}@anu.edu.au

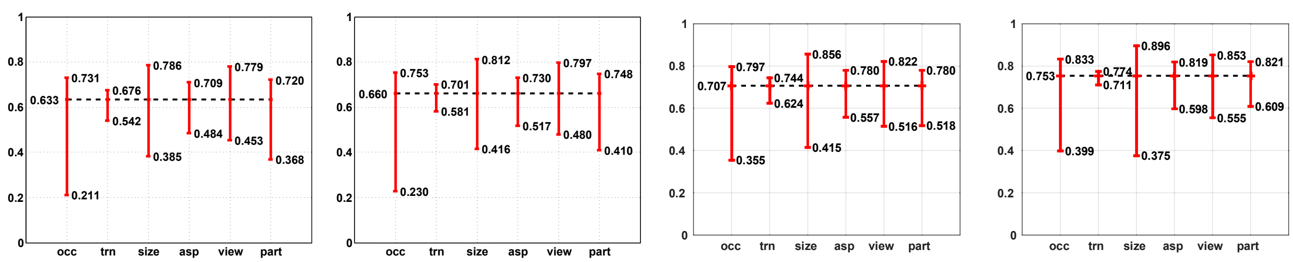
Mathieu Salzmann[1,3]
[3]CVLab, EPFL, Switzerland
mathieu.salzmann@epfl.ch

1. Deep Co-Objectness and Detection

To demonstrate the impact of having better object proposals on object detection, we employed our proposals within the Fast RCNN algorithm of [1]. In Table 1, we compare the resulting Average Precisions with the of state-of-the-art object detectors on Pascal VOC 2007. We used selective search proposals by Hosang et al. [4] as initial pool of bounding-boxes to co-generate the object proposals. Following [3], we also provide a detailed comparison of different performance criteria in Fig. 1. Note that we outperform the original Fast RCNN by 4.2% on average, with up to 12% for some classes, such as *bottle*. Note also that we outperform the state-of-the-art object co-detection approach of [2] and the region proposal network of [5]. The detection trends for all the object categories are detailed in Fig. 2.

VOC 2007 (test)	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv	mAP
Fast RCNN (SS) [1]	74.5	78.3	69.2	53.2	36.6	77.3	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	30.1	65.4	**70.2**	75.8	65.8	66.9
CoDet-G-DL (SS) [2]	**75.8**	78.1	69.3	53.8	36.9	77.5	79.0	82.5	40.1	73.5	67.7	81.4	**82.2**	75.4	70.0	33.4	65.4	70.0	74.3	67.2	67.7
RPN [5]	70.0	**80.6**	70.1	57.3	**49.9**	78.2	80.4	82.0	**52.2**	75.3	67.2	80.3	79.8	75.0	**76.3**	**39.1**	68.3	67.3	**81.1**	67.6	69.9
Co-Obj (SS) + Fast RCNN	73.4	79.4	**71.9**	**61.6**	48.4	**81.2**	**81.9**	**84.8**	47.5	**78.2**	**69.6**	**81.5**	81.9	**77.8**	75.5	38.1	**73.4**	68.9	78.0	**68.0**	**71.1**

Table 1: **Detection Average Precision(%) on the PASCAL VOC 2007 test set:** Rows 1-3 show the multi-class object detection state-of-the-art results. The results of our approach are provided in Row 4.



(a) Fast RCNN (SS) (b) CoDeT-G-DL (c) Co-Obj (B) + Fast RCNN (d) Co-Obj (SS) + Fast RCNN

Figure 1: **Sensitivity and impact analysis:** Overall detailed performance comparison using different metrics (i.e., occlusion, ll average normalized precision AP_N.

provided by Infoscience - Ecole polytechnique fédérale de Lausanne
View metadata, citation and similar papers at core.ac.uk
brought to you by CORE

References

[1] R. B. Girshick. Fast R-CNN. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
[2] Z. Hayder, X. He, and M. Salzmann. Structural kernel learning for large scale multiclass object co-detection. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
[3] D. Hoiem, Y. Chodpathumwan, and Q. Dai. Diagnosing error in object detectors. In *The European Conference on Computer Vision (ECCV)*, pages 340–353. Springer, 2012.

*NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the ARC through the ICT Centre of Excellence program.

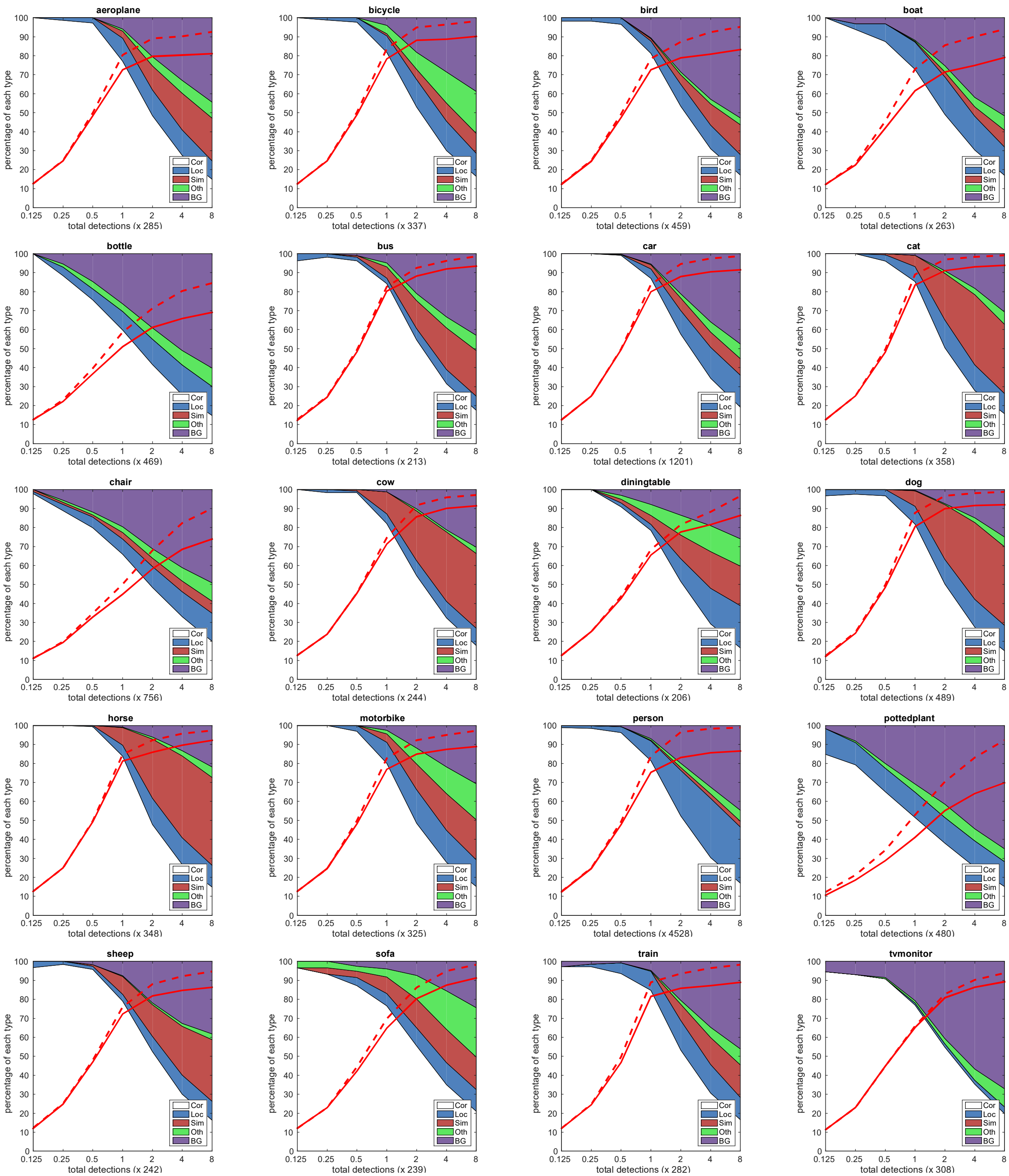


Figure 2: **Deep co-objectness and detection trends.** Following [3], we show the evolution of the type of detection as the number of detections increases; The white areas correspond to the correct detections; The blue areas represent the detections with localization error; The red areas correspond to confusion with a similar category; The green areas represent the confusion with a dissimilar category. Finally, the curves show the recall as a function of the number of objects (dashed=weak localization, solid=strong localization).

[4] J. H. Hosang, R. Benenson, P. Dollár, and B. Schiele. What makes for effective detection proposals? *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2015.

[5] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In *The Conference on Neural Information Processing Systems (NIPS)*, 2015.