The Role of Context for Object Detection and Semantic Segmentation in the Wild

Roozbeh Mottaghi[1] Xianjie Chen[2] Xiaobai Liu[2] Nam-Gyu Cho[3] Seong-Whan Lee[3]
Sanja Fidler[4] Raquel Urtasun[4] Alan Yuille[2]
Stanford University[1] UCLA[2] Korea University[3] University of Toronto[4]

Due to a bug in the evaluation code the results that we published in Tables 1 and 2 are not accurate. This does not change our conclusions and does not affect the results in the rest of the paper. The correct accuracies can be found below:

	Recall	IOU		Recall	IOU
bag	2.1	1.2	**food**	16.4	10.7
bed	2.8	0.7	**mouse**	1.0	0.9
bedcloth	0.0	0.0	**plate**	10.2	5.6
bench	0.2	0.1	**platform**	9.9	7.5
book	13.5	5.0	**rock**	8.0	6.7
cabinet	6.7	4.4	**shelves**	15.1	3.7
clothes	3.3	1.8	**sidewalk**	0.6	0.5
computer	0.0	0.0	**sign**	11.2	7.0
cup	1.9	1.4	**snow**	20.8	16.4
curtain	22.1	11.6	**truck**	0.6	0.2
door	3.6	2.3	**window**	31.7	14.6
fence	10.9	6.6	**wood**	1.2	0.8
flower	14.6	6.8	**light**	14.3	8.5
			Avg.	8.6	4.8

Table 1. The subset of 59 most frequent classes that have low segmentation accuracy according to O_2P [1] results.

References

[1] J. Carreira, R. Caseiroa, J. Batista, and C. Sminchisescu. Semantic segmentation with second-order pooling. In *ECCV*, 2012. 1

[2] J. Tighe and S. Lazebnik. Superparsing: Scalable nonparametric image parsing with superpixels. In *ECCV*, 2010. 1

	Recall		IOU	
	SuperParsing [2]	O_2P [1]	SuperParsing [2]	O_2P [1]
sky	88.8	93.9	65.6	75.6
grass	68.0	77.7	45.3	56.0
water	44.8	72.0	34.5	54.8
person	72.8	57.6	30.1	44.5
tree	66.2	66.7	37.8	44.3
bus	22.8	70.1	14.0	43.2
wall	66.6	68.1	30.8	40.5
cat	36.5	66.4	20.1	36.7
aeroplane	29.3	67.2	19.5	36.4
car	31.2	55.5	15.0	33.5
motorbike	25.7	66.1	14.3	32.8
road	22.8	50.0	15.8	31.2
track	22.9	44.3	17.5	29.5
ground	48.9	41.8	24.0	27.6
dog	18.6	46.3	11.5	26.9
train	16.6	47.9	10.4	26.7
horse	2.2	44.8	2.0	26.4
floor	25.6	46.1	14.4	25.7
bird	4.9	42.7	4.1	24.6
building	45.7	31.4	19.8	24.3
tvmonitor	10.5	48.9	9.0	24.3
sheep	5.0	38.0	4.2	23.7
bicycle	16.6	52.5	11.3	23.5
boat	0.1	37.8	0.0	22.3
mountain	10.3	30.4	8.8	19.2
keyboard	0.1	34.6	0.1	18.2
cow	0.1	24.6	0.1	16.2
sofa	4.4	29.2	3.6	16.1
pottedplant	1.2	40.7	1.1	15.9
bottle	1.3	35.8	1.2	15.0
ceiling	9.6	20.1	6.4	12.7
table	9.6	11.3	6.4	7.0
chair	3.5	10.1	2.9	6.8
Avg.	25.3	47.6	15.2	29.1

Table 2. **Segmentation:** Nearest-neighbor methods such as [2] do not work well on PASCAL due to the high variability of images. In contrast the O_2P classifier [1] on superpixels performs well.