

The Role of Context Selection in Object Detection - Supplementary Material

Ruichi (Rich) Yu¹
richyu@cs.umd.edu

Xi (Stephen) Chen²
chnxi@microsoft.com

Vlad I. Morariu¹
morariu@umiacs.umd.edu

Larry S. Davis¹
lsd@umiacs.umd.edu

¹ University of Maryland
College Park, MD. USA.

² Microsoft Corporation
One Microsofy Way,
Redmond, WA. USA.

1 Predicting Object Class using Pure Contextual Relationships

We show the accuracy for each object class when predicting with pure context in Table 1. Some examples of histograms that represent the learned object-to-object contextual relationships are shown in Figure 1.

	bath- tub	bed	book- shelf	box	chair	coun- ter	desk	door	dress- er	gar- bage bin	lamp	moni- tor	night stand	pillow	sink	sofa	table	tv	toilet
Accu- racy	0.87	0.91	0.70	0.54	0.73	0.76	0.50	0.65	0.73	0.68	0.59	0.71	0.95	0.83	0.71	0.67	0.48	0.47	0.95

Table 1: Prediction Accuracy: Pure Context

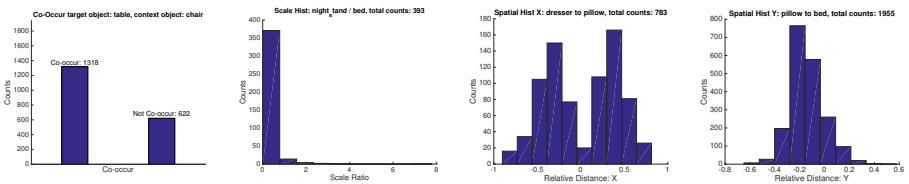


Figure 1: **Likelihood Histograms** The first histogram measures the likelihood of co-occurrence of table and chair. The second one measures the relative scale ratio between night stand and bed. The third and forth histograms measure the relative spatial distance between pillow and dresser along X-axis and the one between bed and pillow along Y-axis.

2 Experiments

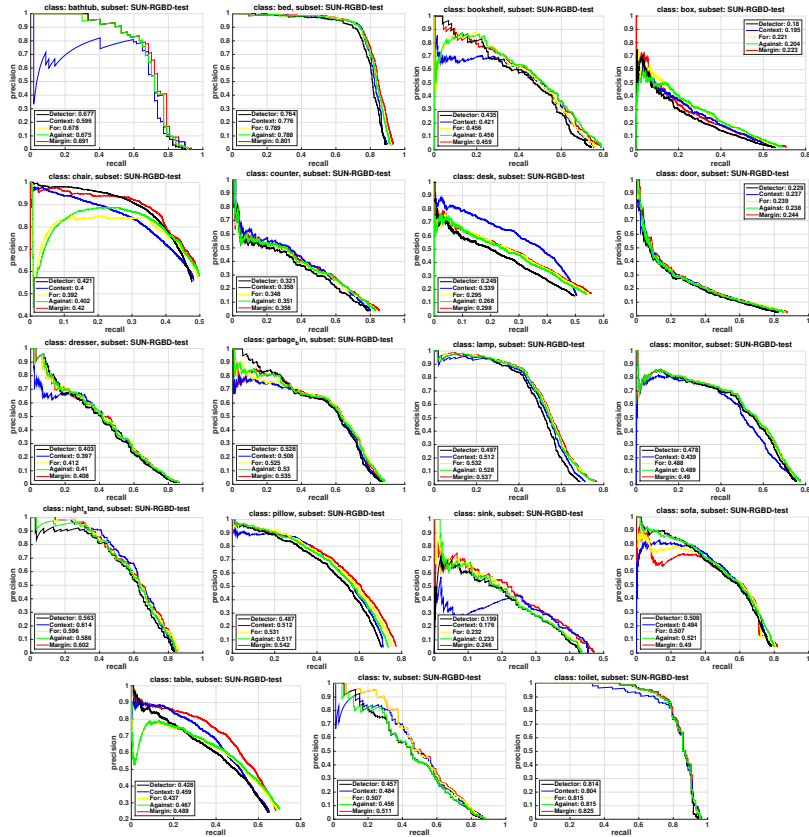


Figure 2: **Precision-Recall (PR) Curves for 19 classes** We show the PR curves for 19 classes with different methods: *Detector* shows the original appearance-based detection results; *Context* shows the results of region-based re-scoring model without context selection; *For* shows the results of FUB method; *Against* shows the results of AUB method; *Margin* shows the results of using the margin between FUB and AUB as the final score.

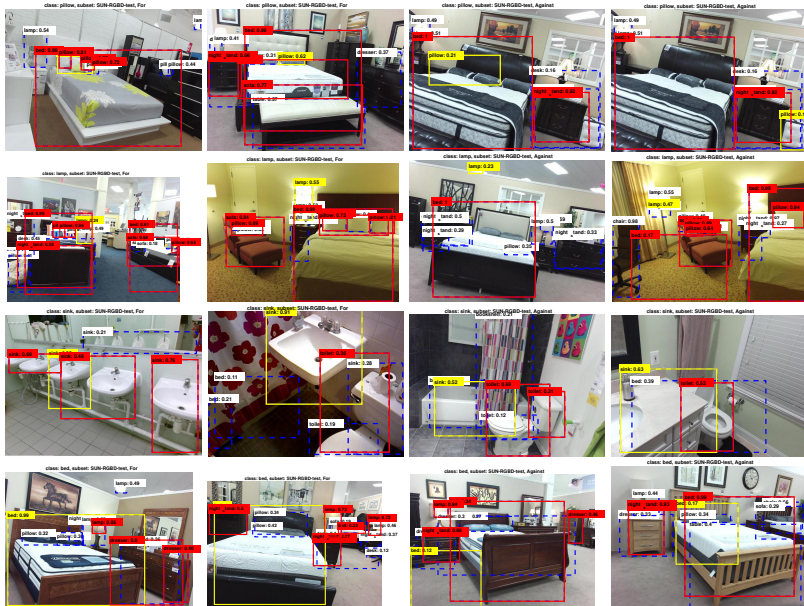


Figure 3: **Visualization of For-Against Context Selection** We visualize the selected contextual regions with the labels that have the highest appearance-based confidence scores among all possible labels for four classes: *pillow*, *lamp*, *sink* and *bed*. All figures are drawn when the precision threshold for potential contextual regions is set as 0.4. The first two columns show the selection results based on the FUB model and the last two columns are the results from the AUB model. The yellow boxes are the target objects, the red boxes are the selected contextual regions, and the blue dashed boxes are the ones that are not selected.