Exploiting Hierarchical Context on a Large Database of Object Categories - Supplemental Materials -

Myung Jin Choi, Joseph J. Lim, Antonio Torralba, Alan S. Willsky
Massachusetts Institute of Technology
myungjin@mit.edu, lim@csail.mit.edu, torralba@csail.mit.edu, willsky@mit.edu

1. Collection and annotation procedure of the dataset

The annotation of the SUN 09 dataset have three phases:

- Images were collected from multiple sources (Google, Flickr, Altavista, LabelMe) and only images corresponding to scenes were kept in the collection. Any close up of an object or images on white backgrounds were removed.
- The images were labeled by one annotator. This procedure was carried out over one year at about 40 images annotated each day.
- The labeled images were carefully verified for consistency and synonymous labels were consolidated. The resulting annotations have a higher quality than what can be achieved by LabelMe or Amazon Mechanical Turk. Therefore, this dataset can be used both for training and testing.

Figure 1 shows a set of typical images in the SUN 09 dataset.

In order to create enough training data for the local detectors, we collected additional examples for over 100 object classes. For this complementary collection the process was:

- Download images from Google and Flickr by querying for those 100 objects.
- The images were then posted on Amazon Mechanical Turk and the task of the workers was to segment only the valid instances of the queried object. This removed most of the noise from the images returned by Google and Flickr. Then, finally, the authors verified the quality of the annotations and removed any additional noise. This dataset was used to train the object detectors only and not to train the context model.

2. Image annotation results on SUN 09

Figure 2 shows more image annotation results comparing the baseline detector and our context model.

3. Average precision-recall for SUN 09

Table 1 shows the average precision-recall (APR) for the object localization task on the SUN 09 dataset.

4. ROC curves for localization and presence prediction tasks

Figures 3 and 4 show the ROC curves for the object localization task and the presence prediction task for the object categories in the PASCAL 07 dataset. Red curves correspond to the performance of the context model, and blue curves show performance of the baseline detector. We used convex hull to make the ROC curves strictly concave.

Figures 5 and 6 show the ROC curves for the object localization and presence prediction tasks for the SUN 09 dataset, respectively.

Figure 1. A set of typical images in SUN 09 dataset.

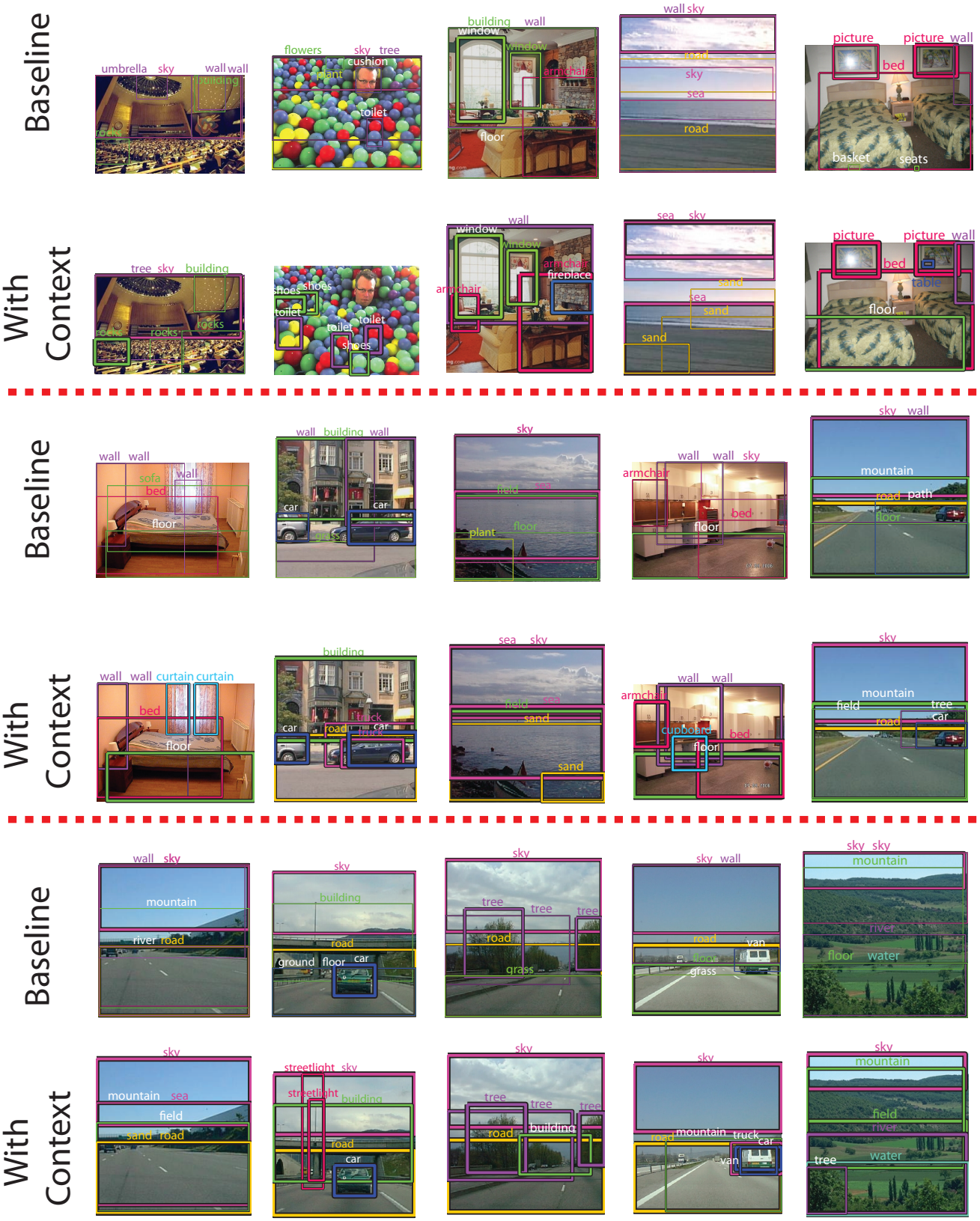
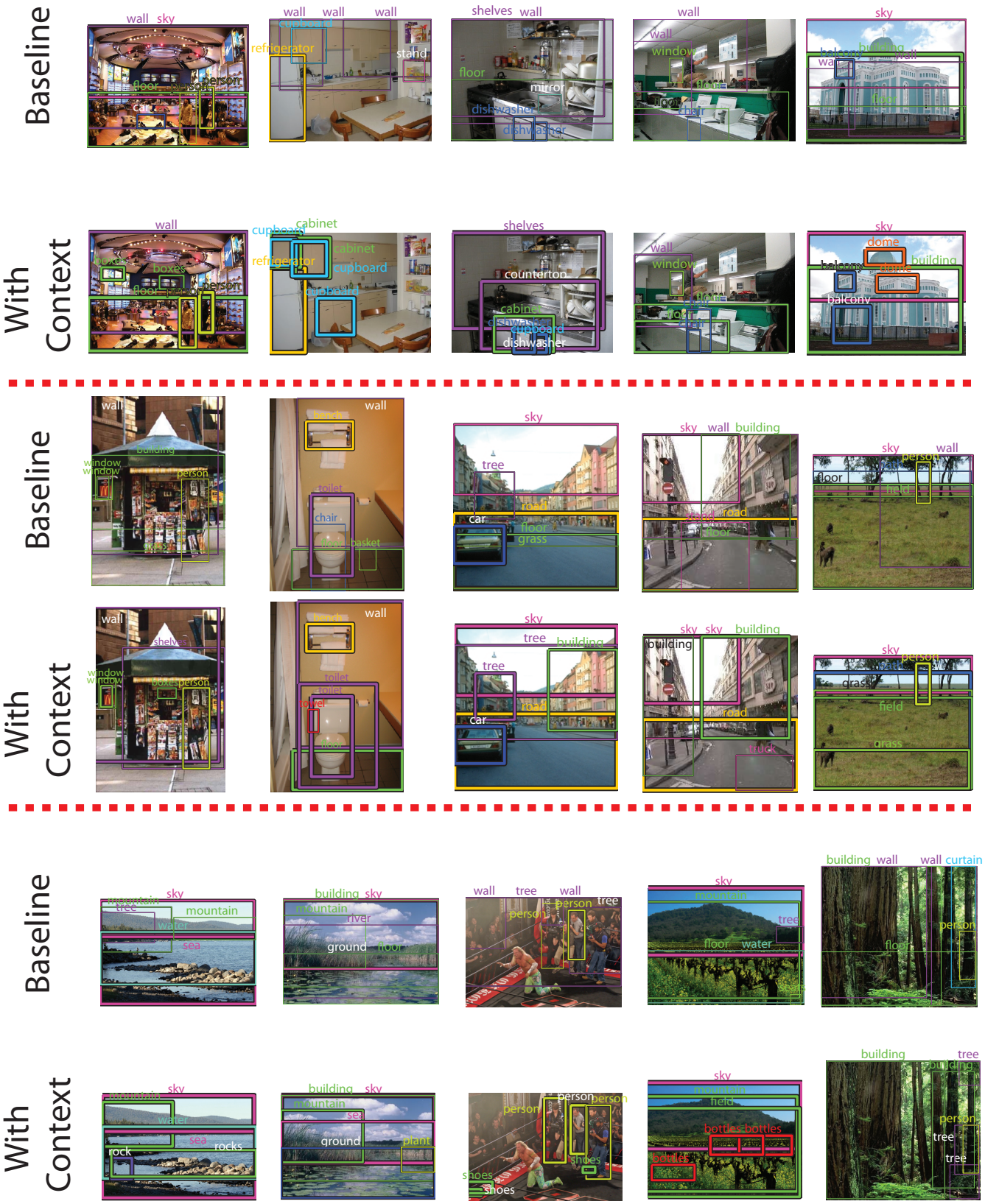


Figure 2. Examples of scenes showing the six most confident detections with and without context.



(Figure 2 continued)

Category	Baseline	Gist	Context	Bound	Category	Baseline	Gist	Context	Bound
airplane	**35.11**	34.01	33.79	46.55	monitor	4.07	10.45	**10.64**	29.63
armchair	9.62	9.97	**11.15**	34.2	mountain	17.23	18.36	**18.38**	38.59
awning	**1.52**	0.48	0.35	4.8	oven	8.07	8.18	**10.27**	36.36
bag	0.28	2.02	**3.64**	8.45	path	0.59	0.84	**1.1**	18.12
balcony	3.06	**9.15**	4.6	11.96	person	17.66	**18.06**	17.98	31.02
ball	9.09	9.09	9.09	8.74	picture	1.65	**2.49**	2.29	13.8
bars	0.03	**0.13**	0.09	16.67	pillow	9.13	9.33	**9.51**	15.81
basket	9.15	9.16	**9.21**	10.74	plant	1.1	4.75	**4.84**	14.78
bed	26.27	32.73	**34.72**	68.42	plate	1.55	2.1	**4.64**	10.42
bench	0.04	0.02	**0.06**	8.7	platform	0.01	0.02	**0.12**	21.88
bookcase	2.32	**4.84**	4.71	25.83	poster	0.32	0.27	**0.49**	3.77
books	1.46	2.05	**2.15**	8.76	pot	**9.15**	9.13	9.13	11.68
bottle	4.6	4.6	**4.68**	10.63	railing	**9.09**	1.52	6.06	6.72
bottles	0.01	0.01	**0.04**	10.53	refrigerator	6.14	13.51	**17.72**	20.63
bowl	0.9	**2.7**	2.45	13.19	river	2.9	4.77	**4.92**	38.83
box	0.19	**0.57**	0.41	6.56	road	33.2	39.59	**39.75**	58.81
boxes	3.03	**9.09**	**9.09**	8.29	rock	0.2	1.03	**4.58**	13.27
bread	0.2	0.4	**0.91**	6.56	rocks	0.45	0.17	**0.7**	7.32
building	14.38	15.5	**15.67**	25.45	rug	0.15	1.1	**4.83**	22.08
bus	9.13	9.26	**9.29**	11.29	sand	3.95	**5.1**	4.79	32.74
cabinet	0.43	**1.1**	1	25.19	screen	9.57	**10.97**	10.91	12.56
candle	0.25	**0.63**	0.1	7.69	sea	28.66	**33.59**	33.43	62.88
car	15.92	16.38	**16.68**	21.25	seats	1.85	9.14	**9.19**	10.5
chair	11.64	12.04	**12.34**	26.48	shelves	2.58	4.75	**5.04**	29.27
chandelier	18.3	18.4	**20.51**	40.58	shoes	**0.05**	0.03	0.02	3.16
clock	9.09	9.09	9.09	7.69	showcase	0.03	0.52	**0.77**	28.99
closet	1.13	**3.44**	3.21	10.2	sink	2.11	**2.84**	2.68	23.08
clothes	0.24	0.1	**0.28**	9.23	sky	55.34	60.13	**60.48**	78.77
counter	0.1	0.1	**0.24**	17.33	sofa	11.47	13.24	**15.3**	46.32
countertop	0.18	0.56	**0.66**	15.91	staircase	**4.57**	2.82	1.39	11.93
cupboard	5.3	**10.66**	8.05	31.5	stand	0.27	0.08	**0.47**	12.2
curtain	6.62	8.33	**12.51**	40.31	steps	**0.65**	0.24	0.18	2.65
cushion	6.11	6.24	**9.52**	15.2	stone	0	0.07	**0.11**	7.14
desk	0.76	1.48	**1.55**	19.48	stones	**1.52**	0.35	0.25	7.25
dish	9.15	3.73	**10.56**	15.79	stool	10.78	11.52	**12.78**	34.43
dishwasher	20.45	20.59	**21.73**	30	stove	9.88	10	**10.05**	20.45
dome	**10.61**	7.81	6.53	21.74	streetlight	3.17	**9.44**	9.38	19.78
door	8.01	6.69	**8.11**	24.44	table	9.84	10.36	**10.56**	26.81
drawer	1.7	2.22	**2.58**	27.78	television	**11.94**	7.34	8.08	28.81
easel	0	0.01	**0.02**	10.87	text	9.09	9.09	9.09	2.6
fence	0.19	0.49	**0.79**	14.57	toilet	**22.01**	21.68	21.18	42.55
field	19.77	25.64	**25.71**	46.78	towel	4.61	4.74	**6.19**	17.53
fireplace	7.66	5.7	**10.72**	26.32	tower	**1.29**	0.97	1.15	20
floor	31.34	42.92	**43.22**	69.1	tray	0.48	0.24	**0.52**	7.14
flowers	0.2	**0.64**	0.6	10.91	tree	10.88	12.43	**12.69**	26.94
gate	0.91	0.83	**2.27**	4.95	truck	**9.65**	5.33	5.08	19.84
glass	0.17	0.2	**0.24**	5.56	umbrella	**0.22**	0.02	0.04	7.46
grass	11.04	12.01	**12.09**	31.08	van	**10.26**	4.19	6.2	15.04
ground	1.49	1.8	**1.82**	32.86	vase	0.2	0.37	**0.49**	11.49
handrail	**0.88**	0.23	0.59	16.83	videos	2.27	2.27	2.27	5.88
headstone	0.14	**0.25**	0.07	12.9	wall	14.7	14.14	**15.43**	32.43
machine	**9.1**	9.09	**9.1**	29.41	water	1.51	**3.92**	2.74	17.93
microwave	27.71	25.44	**28.36**	50	window	**9.95**	9.87	9.94	15.67
mirror	0.83	0.82	**1.61**	34.05	AVERAGE	7.06	7.84	**8.37**	21.55

Table 1. Average precision-recall for object categories in the SUN 09 dataset.

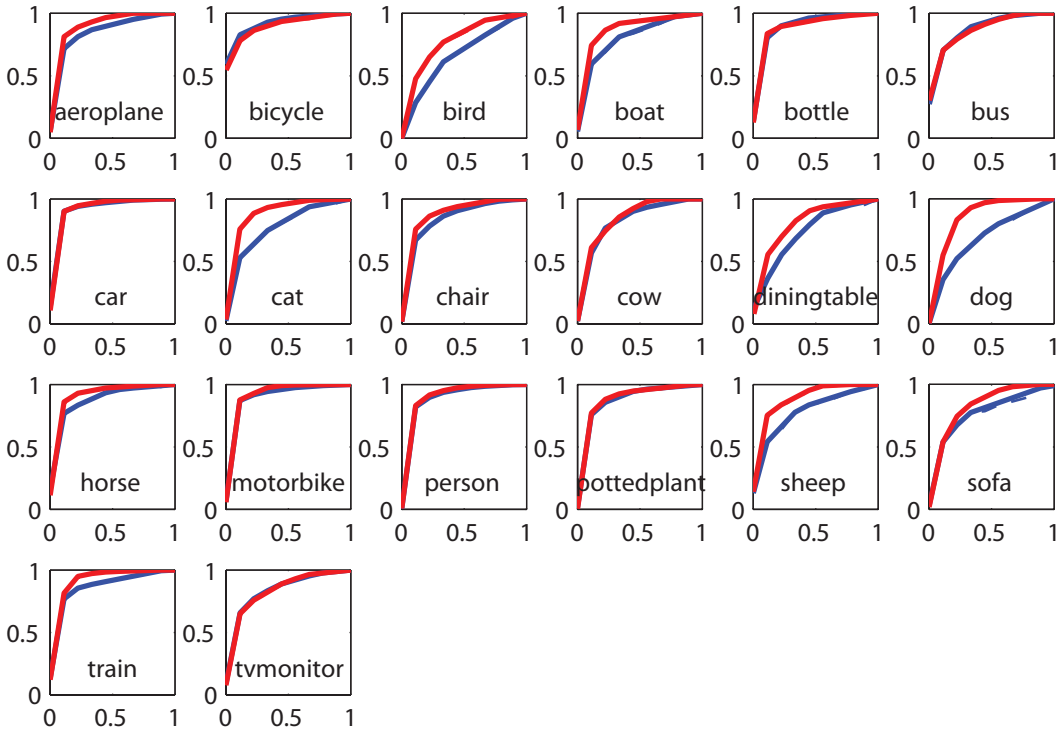


Figure 3. The ROC curves for the localization task on PASCAL 07. Red curve: context model. Blue curve: baseline detector.

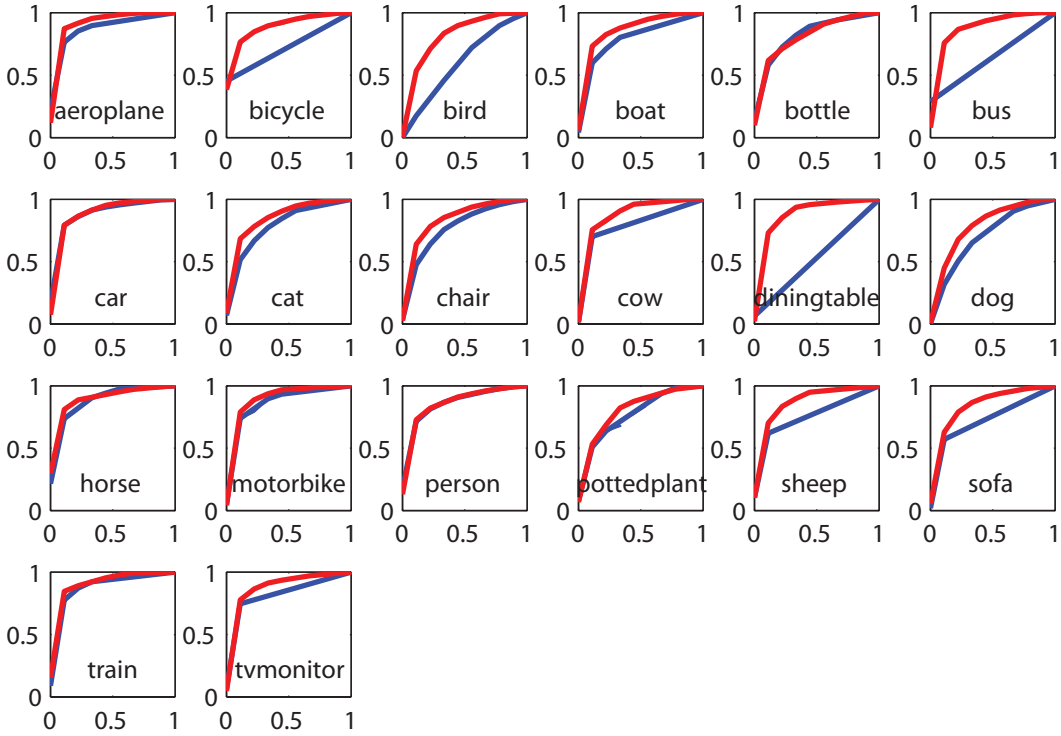


Figure 4. The ROC curves for the presence prediction task on PASCAL 07. Red curve: context model. Blue curve: baseline detector.

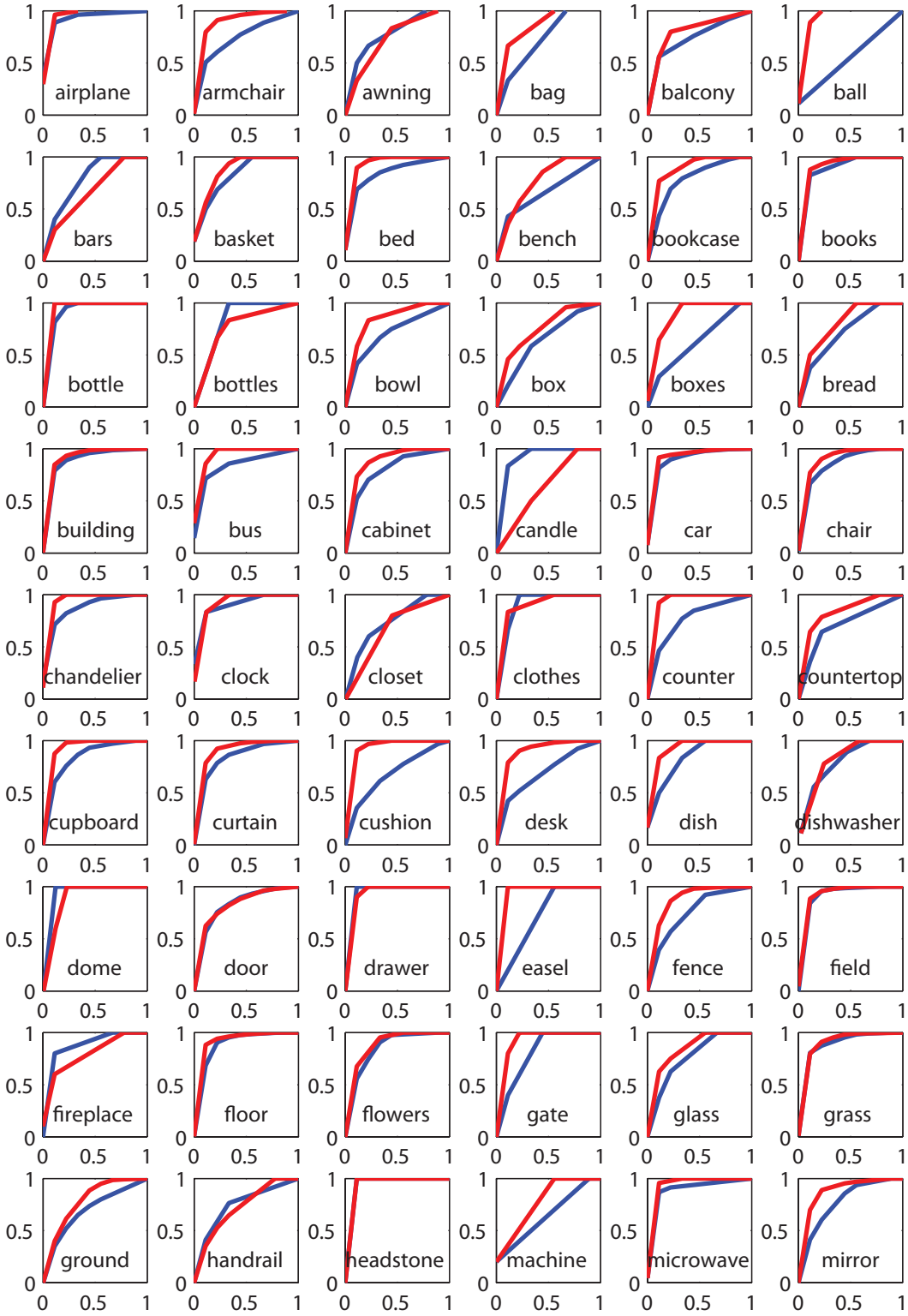
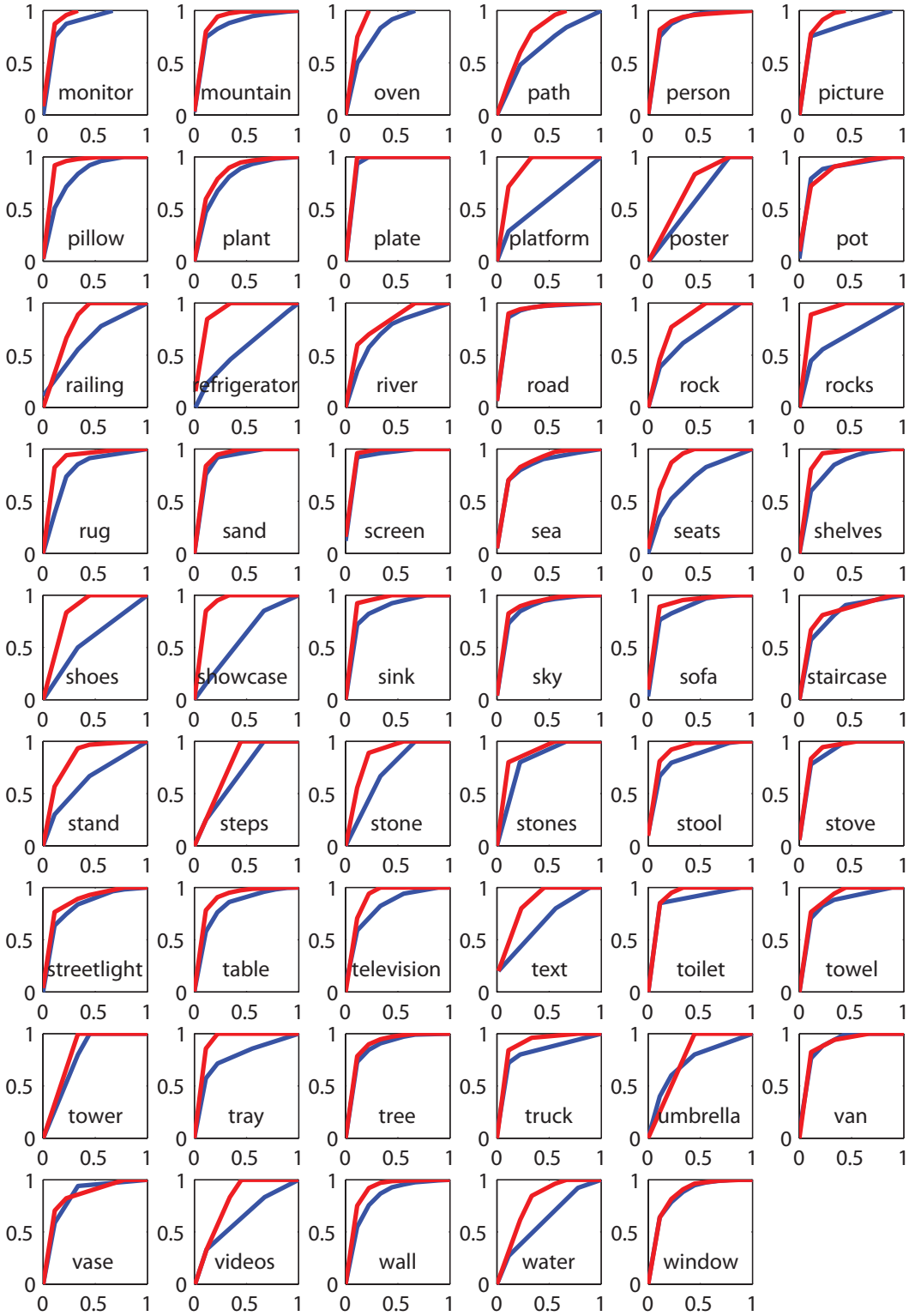


Figure 5. The ROC curves for the localization task on SUN 09. Red curve: context model. Blue curve: baseline detector.



(Figure 5 continued)

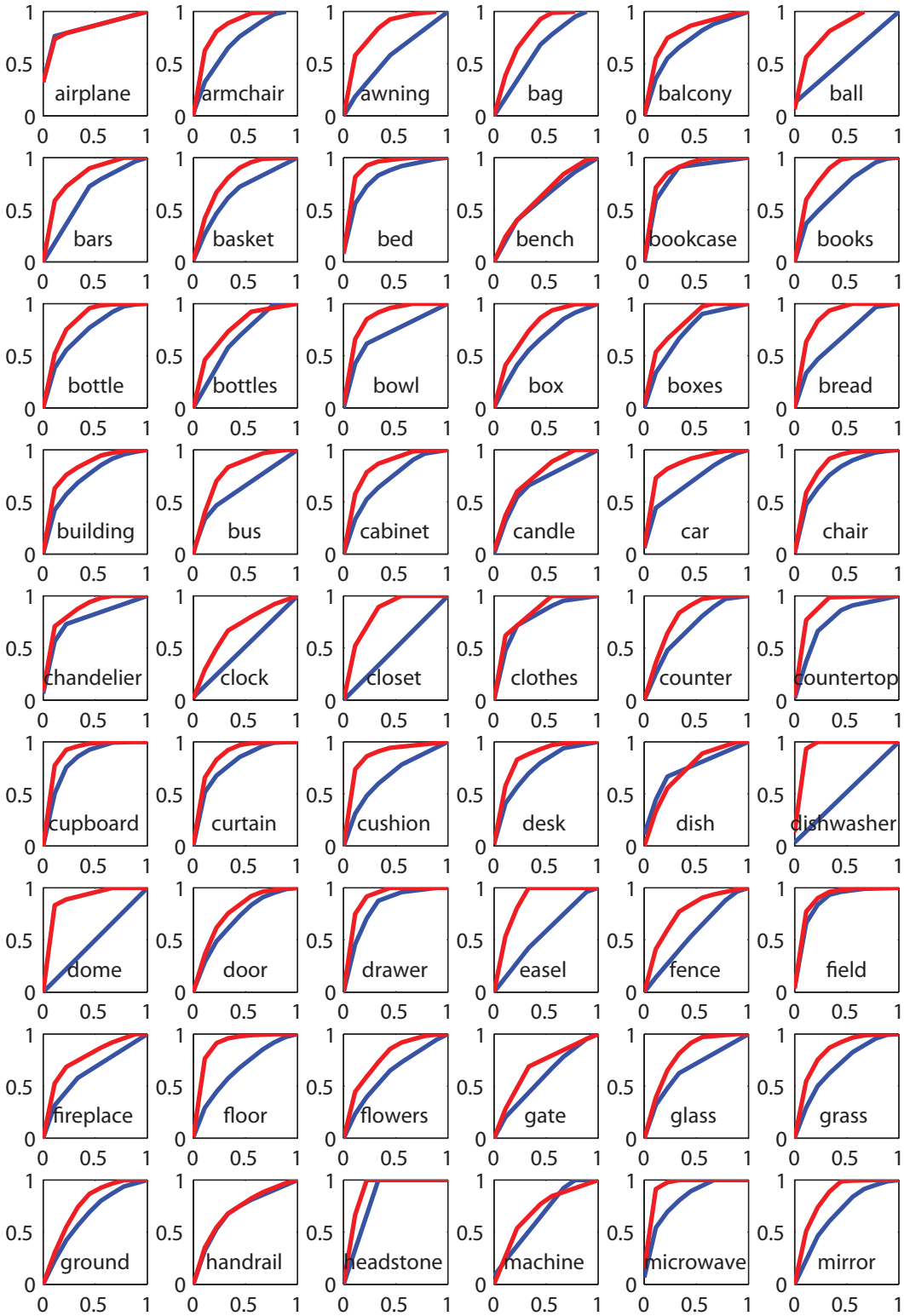
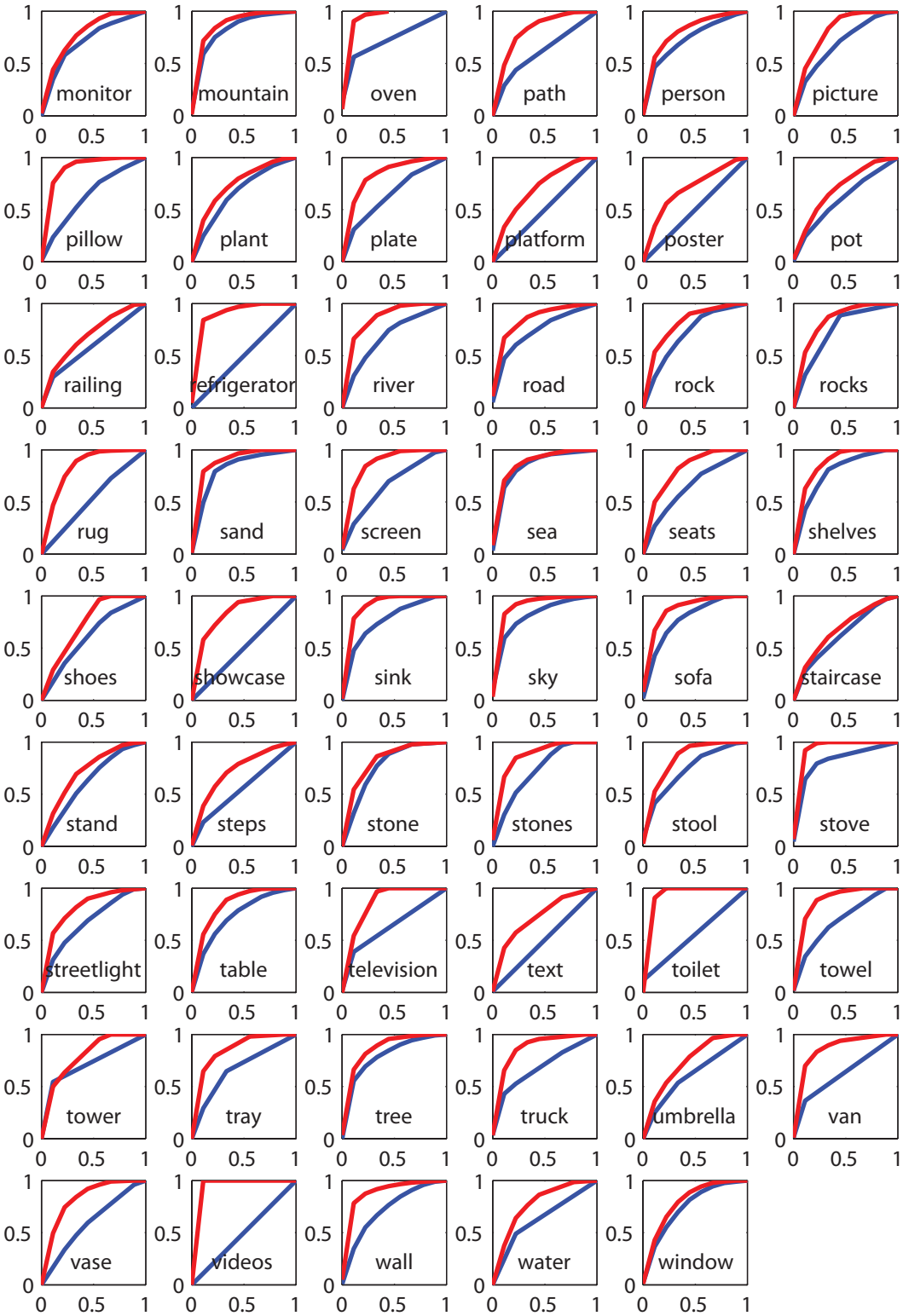


Figure 6. The ROC curves for the presence prediction task on SUN 09. Red curve: context model. Blue curve: baseline detector.



(Figure 6 continued)