

A Coarse-to-Fine Model for 3D Pose Estimation and Sub-category Recognition –Supplementary Material–

Roozbeh Mottaghi¹, Yu Xiang^{2,3}, and Silvio Savarese³

¹Allen Institute for AI, ²University of Michigan-Ann Arbor, ³Stanford University

We provide per-category results in Tables 1 and 2. These tables correspond to Tables 1 and 2 in the paper, respectively.

	Bounding Box	All	Sub-category & Viewpoint	Sub-category	Viewpoint (8 views)
RCNN [2]					
aeroplane	69.1	✗	✗	✗	✗
boat	29.3	✗	✗	✗	✗
car	55.8	✗	✗	✗	✗
DPM-VOC+VP [3]					
aeroplane	40.5	✗	✗	✗	28.6
boat	0.5	✗	✗	✗	0.2
car	47.6	✗	✗	✗	36.6
V-DPM [1]					
aeroplane	39.8	✗	✗	✗	23.8
boat	5.8	✗	✗	✗	1.0
car	37.3	✗	✗	✗	23.8
SV-DPM [1]					
aeroplane	41.8	✗	14.0	24.1	26.9
boat	6.7	✗	4.4	6.1	4.7
car	34.9	✗	6.7	11.2	23.0
FSV-DPM [1]					
aeroplane	40.0	0.12	16.2	25.8	25.7
boat	5.4	0.51	3.1	4.7	3.2
car	32.0	0.43	4.5	7.6	19.5

Table 1. Results of variation of DPM [1], DPM-VOC+VP [3] and RCNN [2] on PASCAL3D+ [4] for all three or a subset of tasks. The result of DPM-VOC+VP [3] was adopted from [4]. The first column (‘Bounding Box’) is equivalent to the standard detection AP of PASCAL VOC. The meaning of ✗ is that the method is not capable of doing that task.

	Bounding Box	All	Sub-category & Viewpoint	Sub-category	Viewpoint (8 views)
1-layer hierarchy (ours)					
aeroplane	69.1	✗	✗	✗	40.5
boat	26.3	✗	✗	✗	9.8
car	53.0	✗	✗	✗	36.4
2-layer hierarchy (ours)					
aeroplane	69.2	✗	28.6	45.0	41.3
boat	29.5	✗	8.4	21.3	9.6
car	54.4	✗	11.1	16.1	37.7
3-layer hierarchy (ours)					
aeroplane	69.2	4.0	28.7	46.7	40.8
boat	29.5	4.0	8.8	21.9	10.6
car	56.1	1.6	15.4	23.2	37.2
Flat model					
aeroplane	69.2 [†]	3.0	25.6	42.0	38.3
boat	29.5	3.1	7.4	22.8	10.6
car	56.1	1.2	11.3	18.2	29.6
Separate					
aeroplane	69.2 [†]	1.4	27.0	47.6	39.2
boat	29.5	2.9	8.6	24.8	10.1
car	56.1	1.5	12.8	20.6	36.7

Table 2. Results of variations our hierarchical model, a flat model that uses the same set of features as those of the 3-layer hierarchy, and also separate classifiers on PASCAL3D+[4]. [†] We consider the same confidence values as the 3-layer model. So the bounding box detection results are identical.

Sub-category information:

We considered the following sub-categories in our experiments:

- **Aeroplane:** Airline, Fighter, Propeller, Shuttle
- **Boat:** Cabin, Cruise, Rowing, Sailing
- **Car:** Hatchback, Mini, Minivan, Race, Sedan, SUV, Truck, Wagon

Confusion matrices for sub-category recognition:

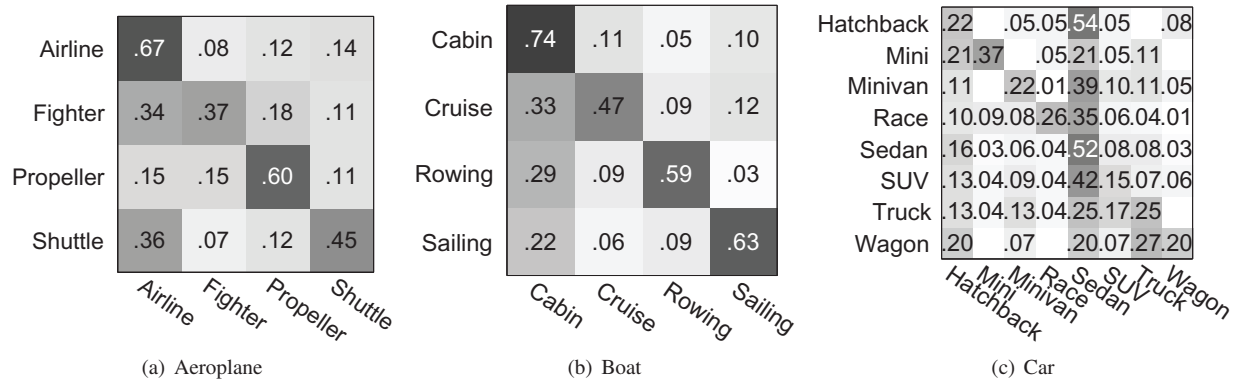


Figure 1. Confusion matrices for sub-category recognition. The recall rate is 81.7, 66.0, and 71.1% for *aeroplane*, *boat*, and *car*, respectively.

References

- [1] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. *PAMI*, 2010.
- [2] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *CVPR*, 2014.
- [3] B. Pepik, M. Stark, P. Gehler, and B. Schiele. Teaching 3d geometry to deformable part models. In *CVPR*, 2012.
- [4] Y. Xiang, R. Mottaghi, and S. Savarese. Beyond pascal: A benchmark for 3d object detection in the wild. In *WACV*, 2014.