

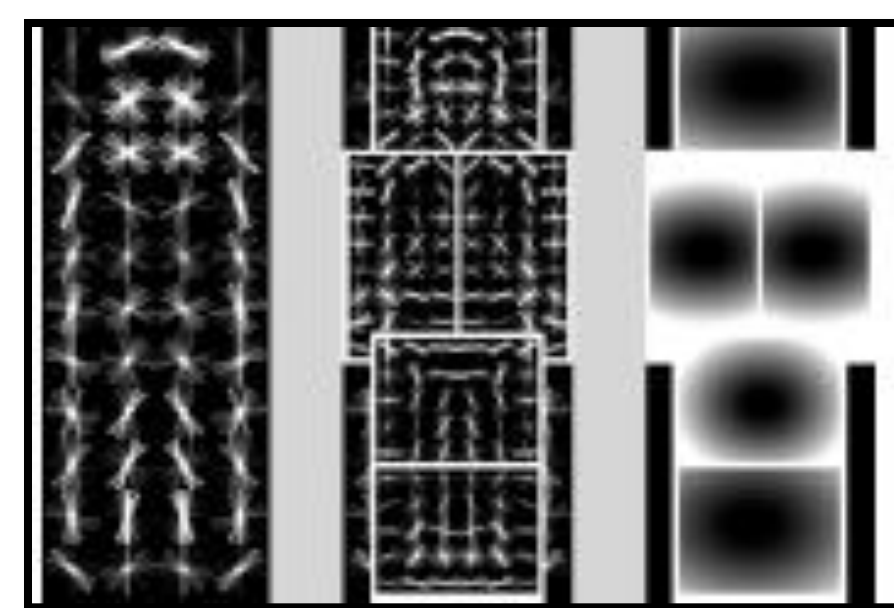
Goal

Evaluation of *context* on the challenging PASCAL VOC 2008 dataset using state-of-the-art object detectors

We analyze

- 6 sources of context (4 old, 2 new)
- 4 uses of context (3 old, 1 new)
- Effects on localization and recognition across 20 object categories
- Most/least effective scenarios

Object Detector




- Sliding-window based
- Deformable part-based HOG (Felzenszwalb *et al.*)
- Top performer in PASCAL challenge ('UoCTTI' entry)

Sources of Context


Any information that aids general scene perception

Scene Gist



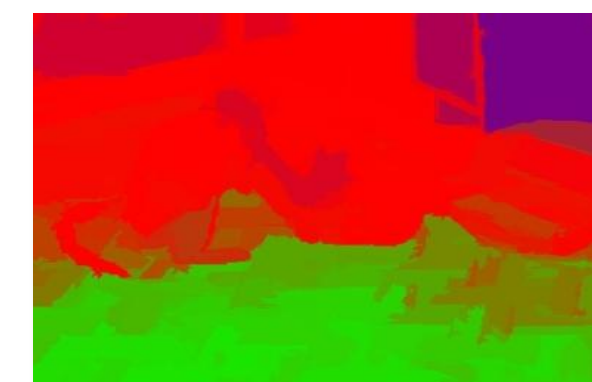
Global Image Statistics (Oliva & Torralba)

Local Pixel



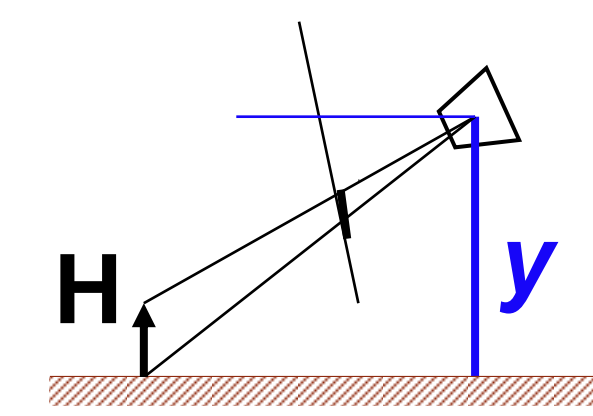
Boundary cues, Shape mask

Geometric



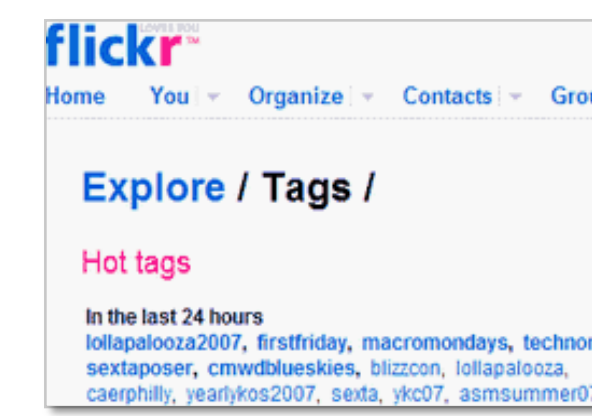
Scene layout, Occlusion (Hoiem *et al.*)

Photogrammetric




Camera/object relationships

Semantic



Keywords, Scene category

Geographic



Geographic properties (Hays & Efros)

Uses of Context

Application of context sources to detection (motivated from Biederman's relational constraints)

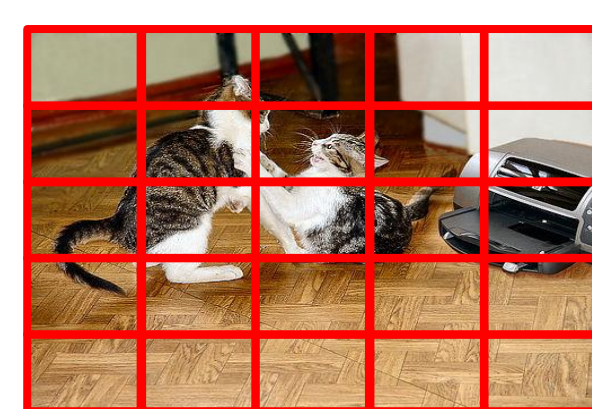
- Presence, Location, Size ('Scene' context)
- Spatial Support/Interposition context

Presence



Gist, Geometric, Semantic, and Geographic

Location



Gist, and Geometric

Size

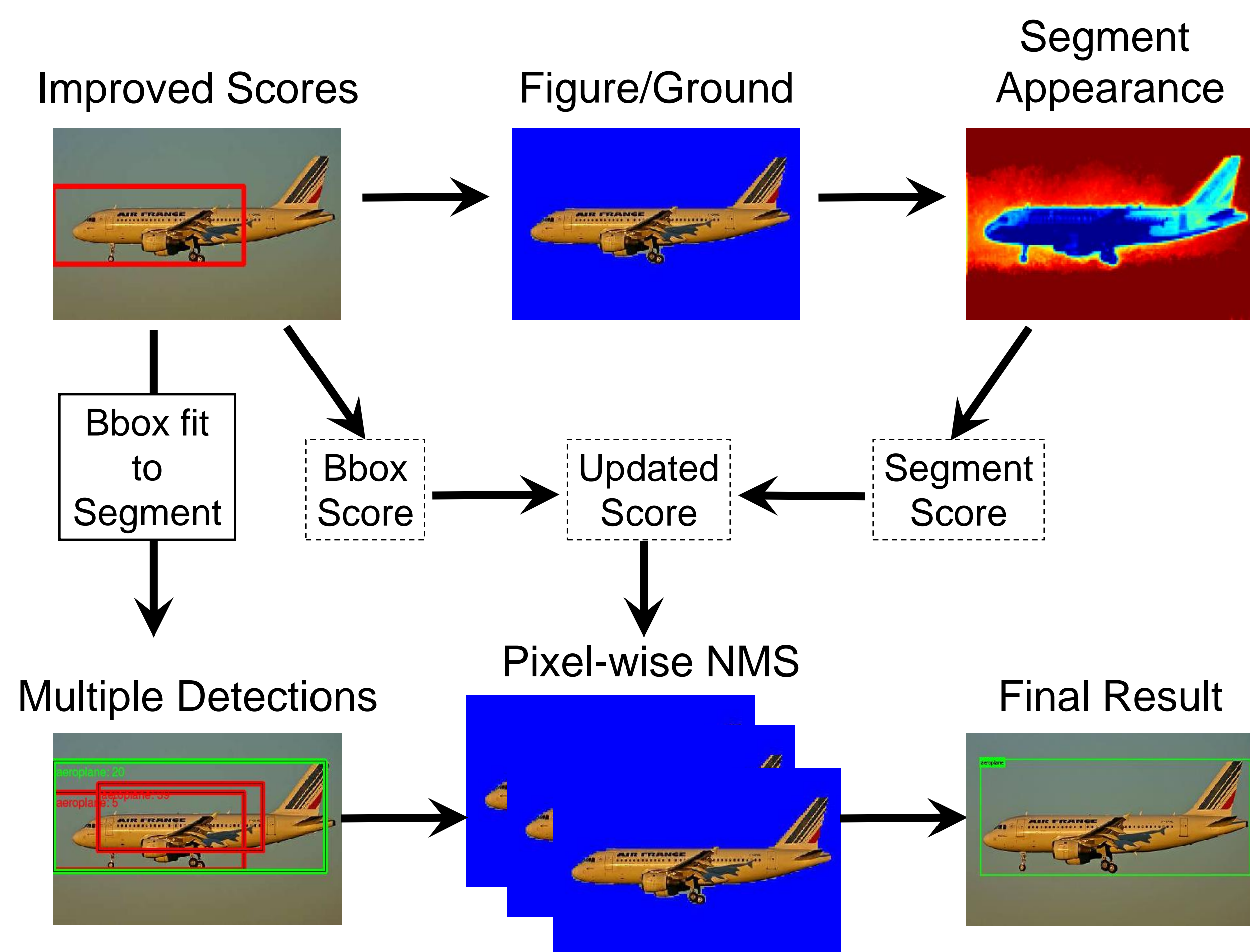


Photogrammetric, Gist, and Geometric

Scene Context Combination

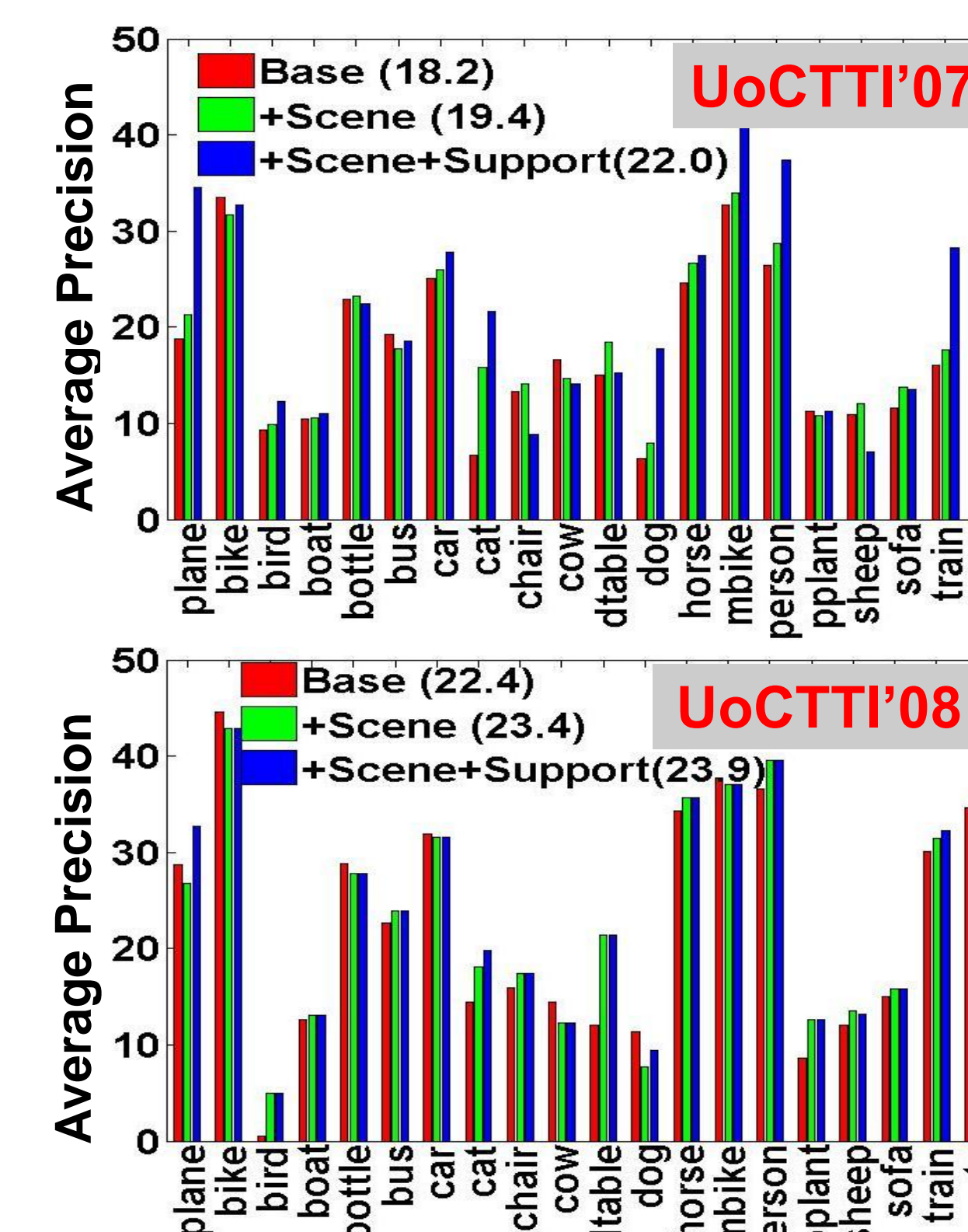


Spatial Support



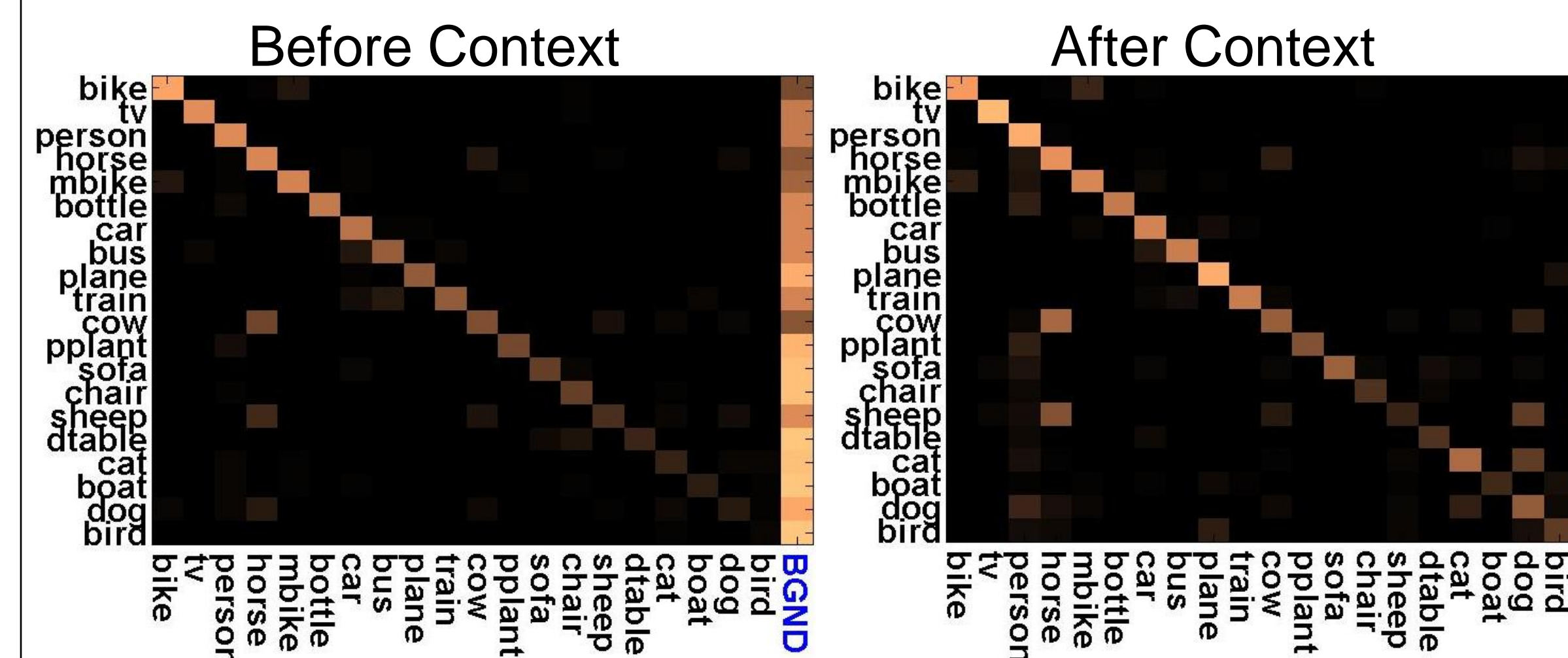
Experimental Results

Overall Results

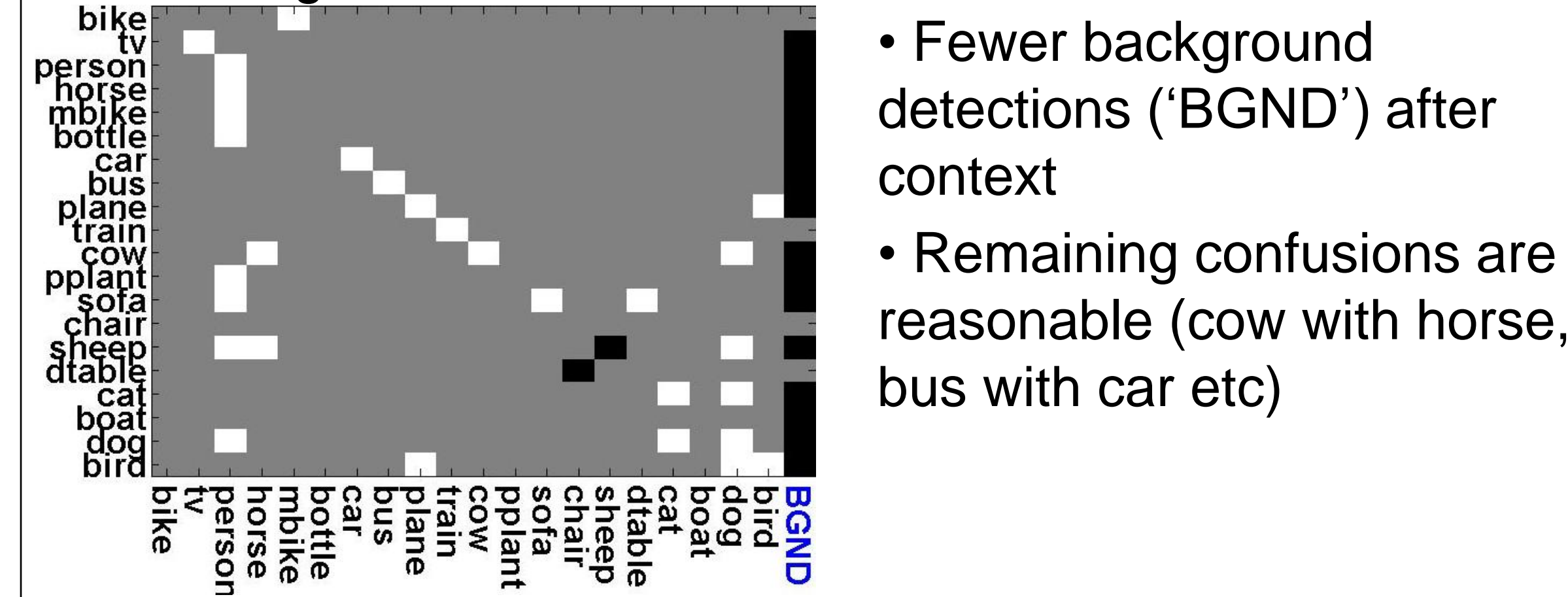


Considerable gains for airplane, cat, person, train, small drop for bicycle, cow

Confusion Matrices



Change in Confusions



- Fewer background detections ('BGND') after context
- Remaining confusions are reasonable (cow with horse, bus with car etc)

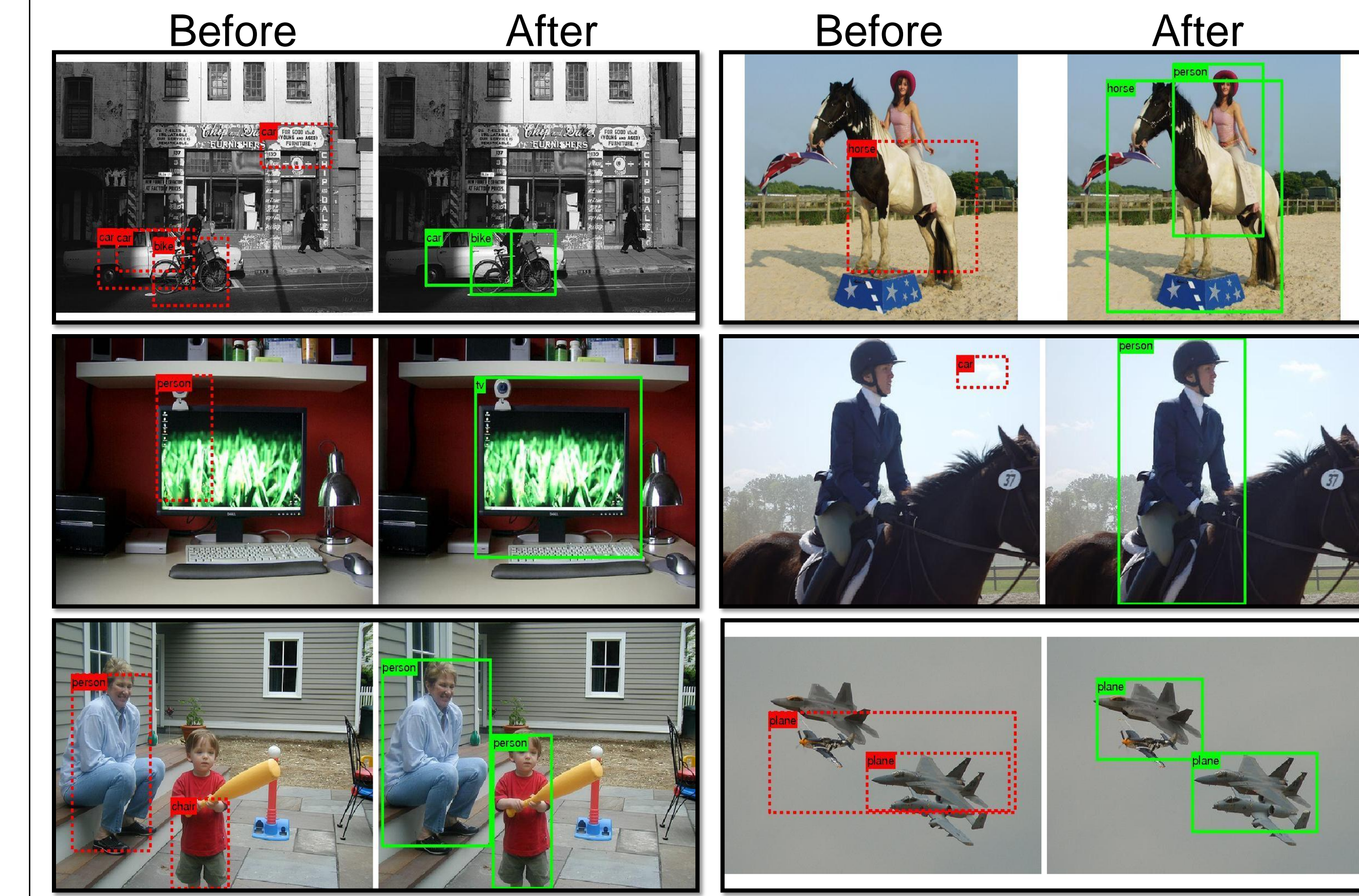
Analysis w.r.t Different Subsets

Subsets	Small	Large	Occluded	Non-Occluded	Difficult
Mean A.P.					
No Context	6.7%	9.3%	4.8%	10.4%	0.2%
Context	12.0%	9.7%	7.5%	11.5%	0.3%

Context particularly helps when objects have impoverished appearance

Experimental Results

Some Qualitative Results



(Only detections above 0.5 precision are displayed)

Analysis of Context Sources & Uses

Sources \ Uses	Uses			Mean A.P.
	Presence A.P.	Location A.P.	Size Avg. Prediction Error	
All				20.5%
Scene Gist	23.9%	3.0%	1.16	20.2%
Geometric	21.5%	2.5%	1.18	
Geographic	15.1%	-	-	19.8%
Semantic	25.6%	-	-	
Photogrammetric	-	-	1.08	19.2%
No Size				
All (Baseline)	31.2%	6.5%	1.08	18.5%

Each source of context is useful; size is the strongest use of context while location is the weakest

Challenging Scenarios



Objects out of their typical context Objects sharing similar contexts

Conclusions

1. Context provides modest improvement on average
2. Detection errors after context are more reasonable
3. Each type of context helps