

# When Good Detectors Go Bad: A Study of Object Detection Error Automatically Generated Detector Report

July 31, 2014

## 1 Info

The CNN v7 bb detector is analyzed. This is an automatically created document.

## 2 Overall Detector Characteristics

### 3 aeroplane

```
Characteristics: ntotal=285 ntrunc=97
occlevel: None=238 Low=38 Med=5 High=4
side visible:
  bottom: Yes=79 No=206
  front: Yes=68 No=217
  rear: Yes=24 No=261
  side: Yes=221 No=64
  top: Yes=36 No=249
part visible:
  body: Yes=265 No=20
  head: Yes=255 No=30
  tail: Yes=243 No=42
  wing: Yes=267 No=18
```

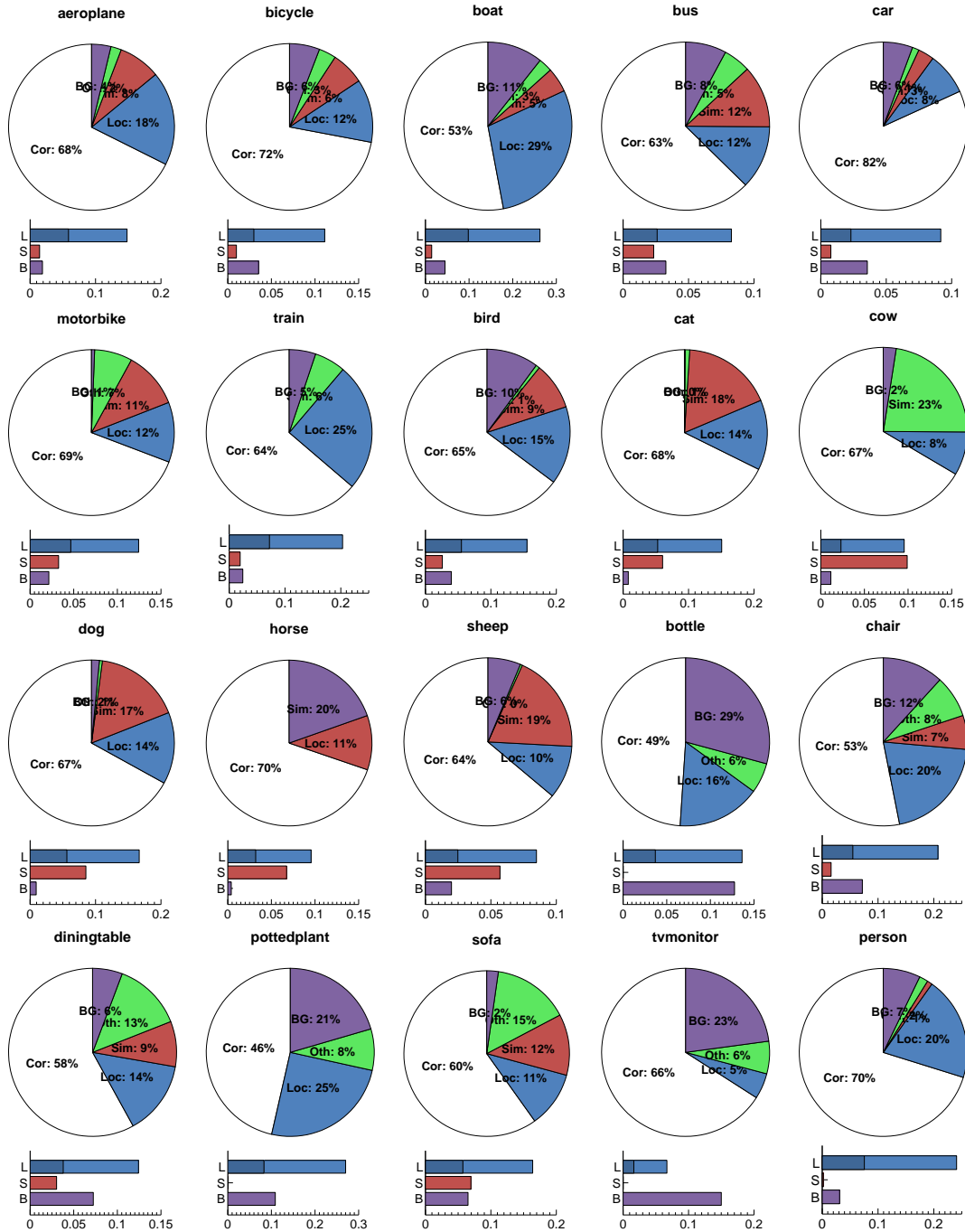


Figure 1: **Analysis of Top-Ranked Detections.** Pie charts: fraction of top N detections ( $N$ =num of objs in category) that are correct (Cor), or false positives due to poor localization (Loc), confusion with similar objects (Sim), confusion with other VOC objects (Oth), or confusion with background or unlabeled objects (BG). Bar graphs: absolute AP improvement by removing all false positives of one type. ‘B’: no confusion with background and non-similar objects. ‘L’: first bar segment is improvement if duplicate or poor localizations are removed; second bar is improvement if localization errors are corrected so that the false positives become true positives.

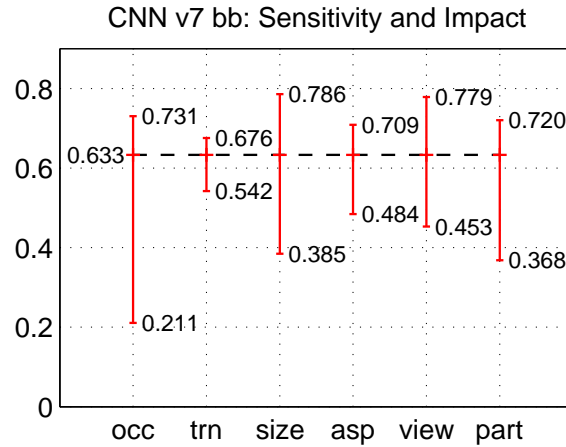


Figure 2: **Summary of Sensitivity and Impact of Object Characteristics:** We show the average (over categories) AP<sub>N</sub> performance of the highest performing and lowest performing subsets within each characteristic (occlusion, truncation, bounding box area, aspect ratio, viewpoint, part visibility). Overall AP<sub>N</sub> is indicated by the black dashed line. The difference between max and min indicates sensitivity; the difference between max and overall indicates the impact.

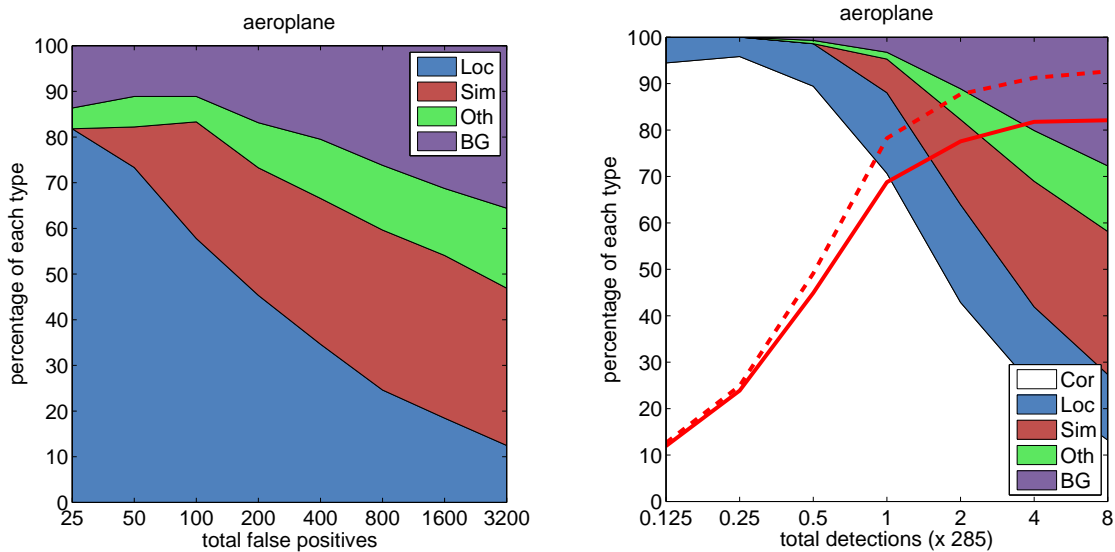


Figure 3: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

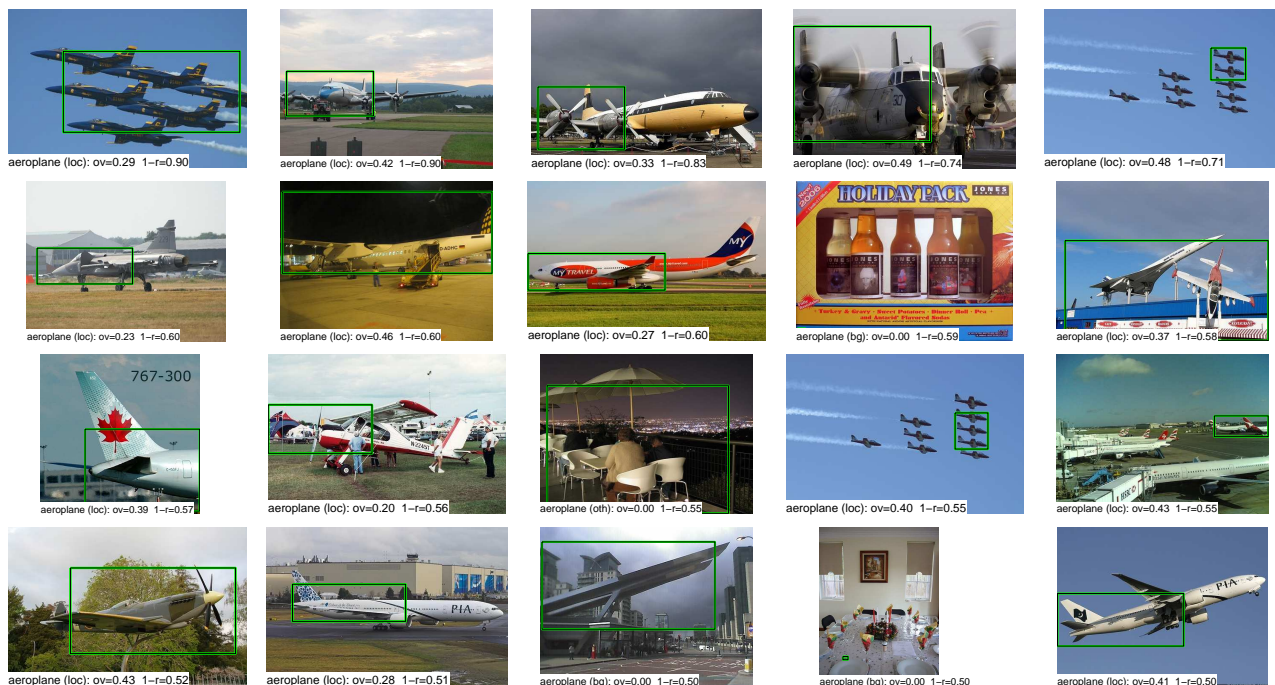


Figure 4: Examples of top aeroplane false positives

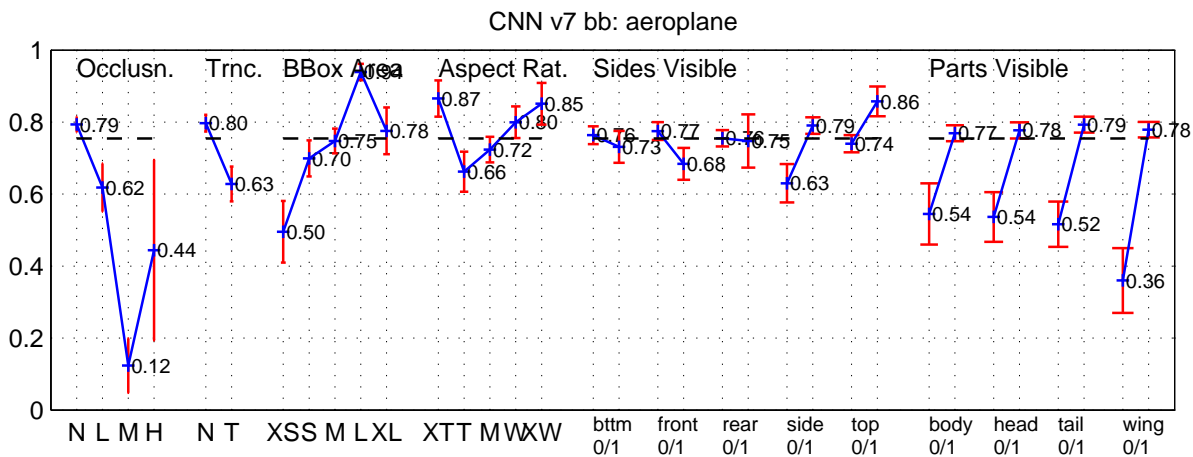


Figure 5: Analysis of aeroplane characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.

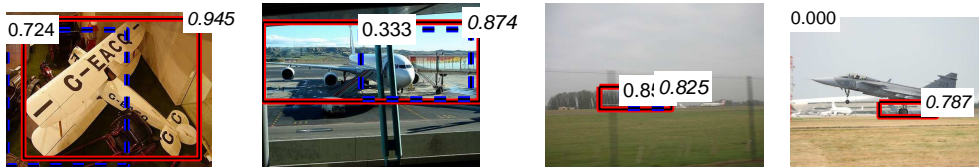


Figure 6: **Unexpectedly difficult aeroplane detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.

## 4 bicycle

Characteristics: ntotal=337 ntrunc=152

occlevel: None=107 Low=168 Med=57 High=5

side visible:

bottom: Yes=4 No=333

front: Yes=83 No=254

rear: Yes=41 No=296

top: Yes=103 No=234

side: Yes=218 No=119

part visible:

body: Yes=305 No=32

handlebars: Yes=308 No=29

seat: Yes=165 No=172

wheel: Yes=329 No=8

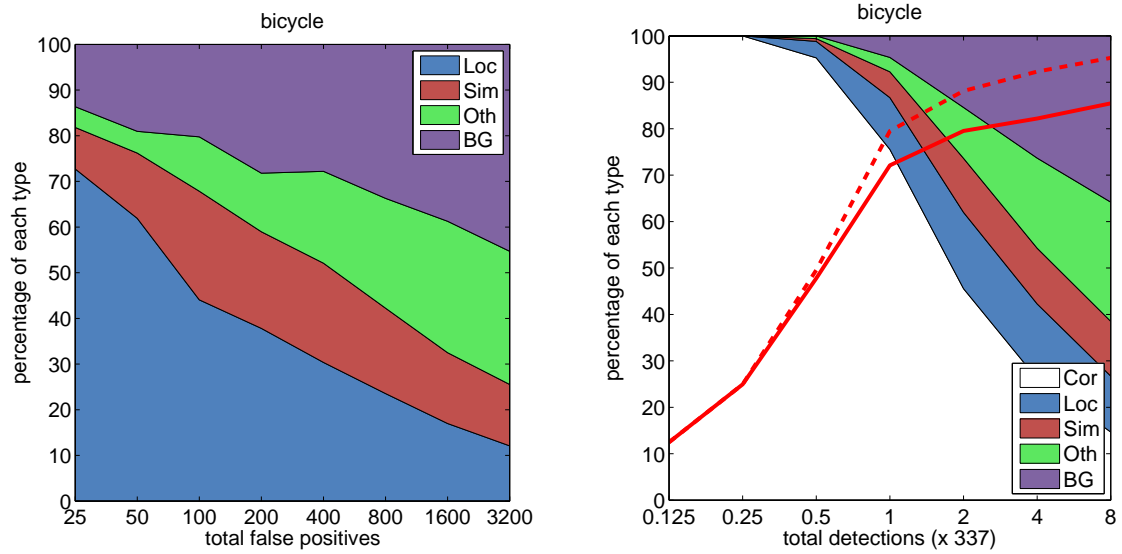


Figure 7: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

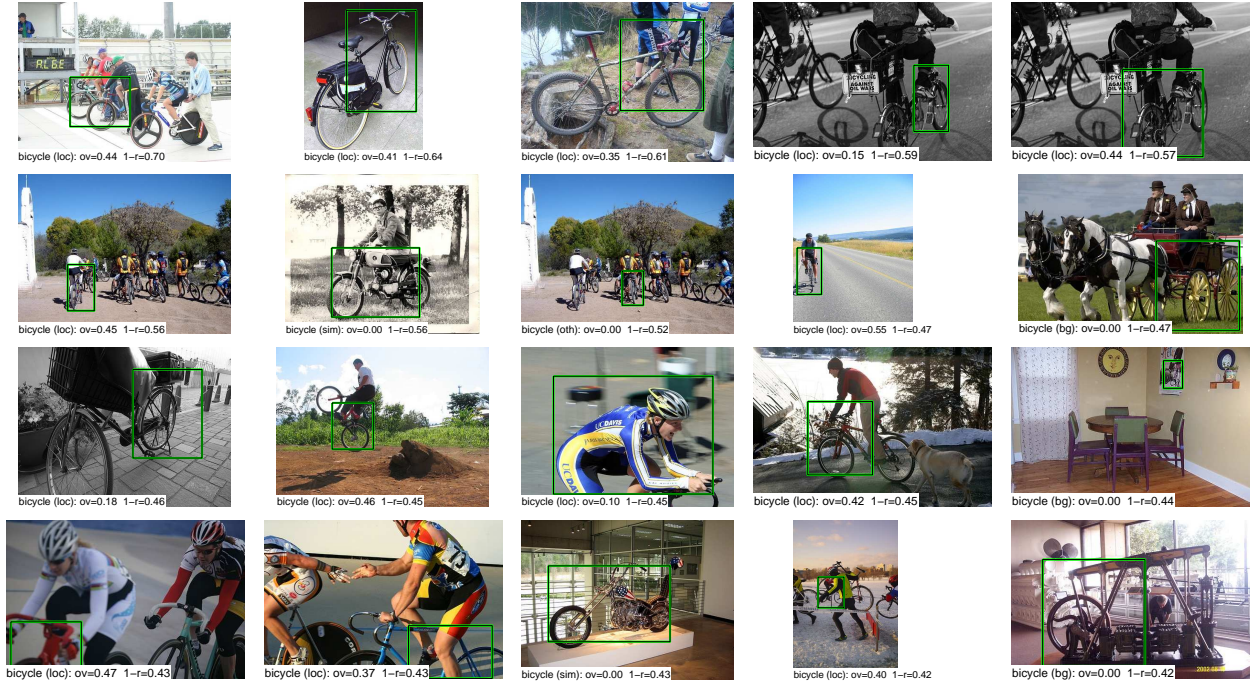


Figure 8: Examples of top bicycle false positives

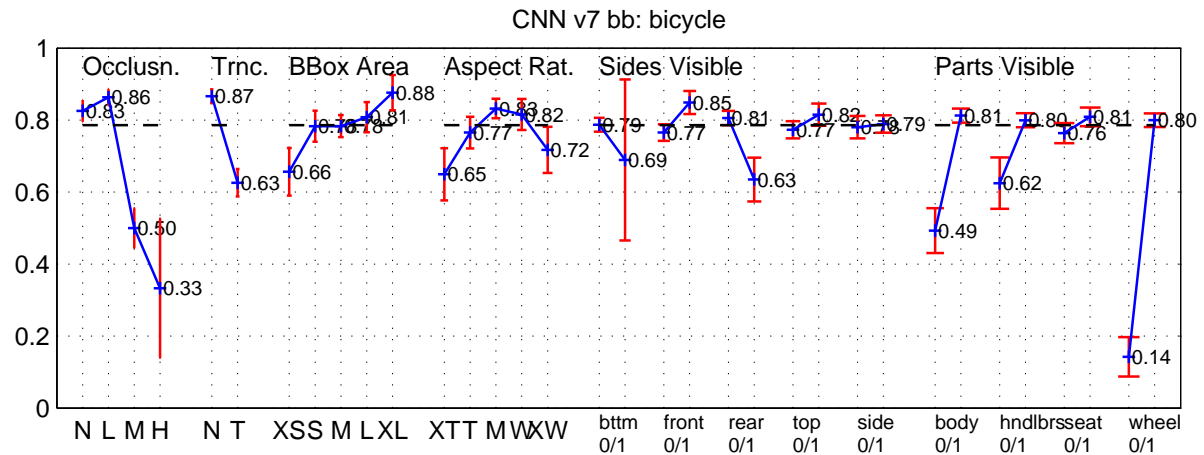


Figure 9: Analysis of bicycle characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.



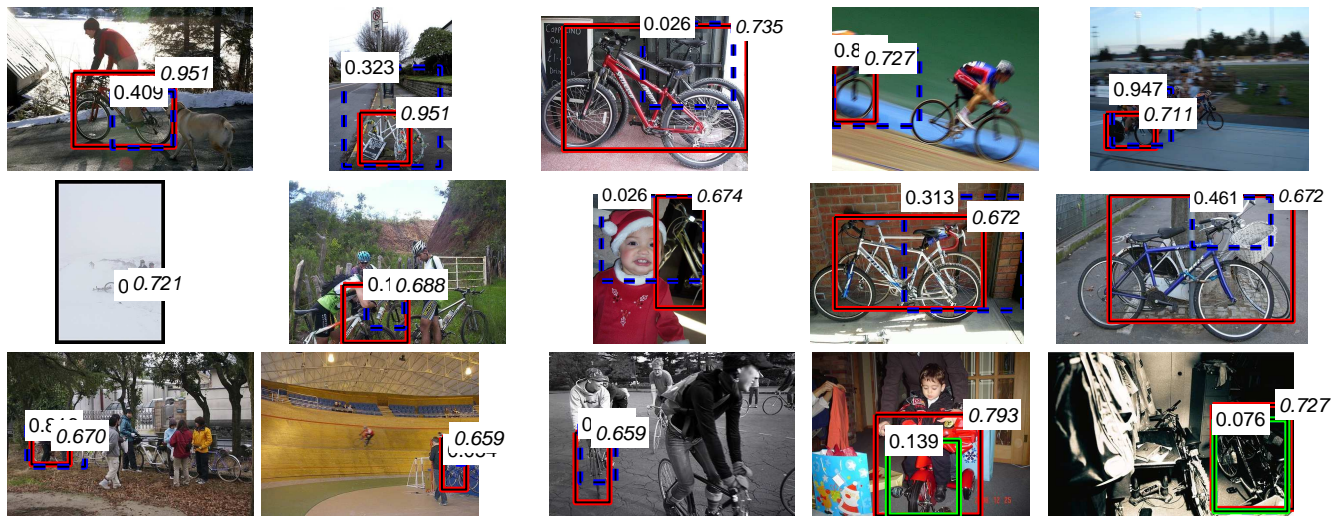


Figure 10: **Unexpectedly difficult bicycle detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.

## 5 boat

Characteristics: ntotal=263 ntrunc=98  
occlevel: None=187 Low=61 Med=14 High=1  
side visible:  
  bottom: Yes=5 No=258  
  front: Yes=91 No=172  
  rear: Yes=25 No=238  
  side: Yes=241 No=22  
  top: Yes=32 No=231  
part visible:  
  body: Yes=261 No=2  
  cabin: Yes=85 No=178  
  mast: Yes=76 No=187  
  paddle: Yes=2 No=261  
  sail: Yes=51 No=212  
  window: Yes=28 No=235

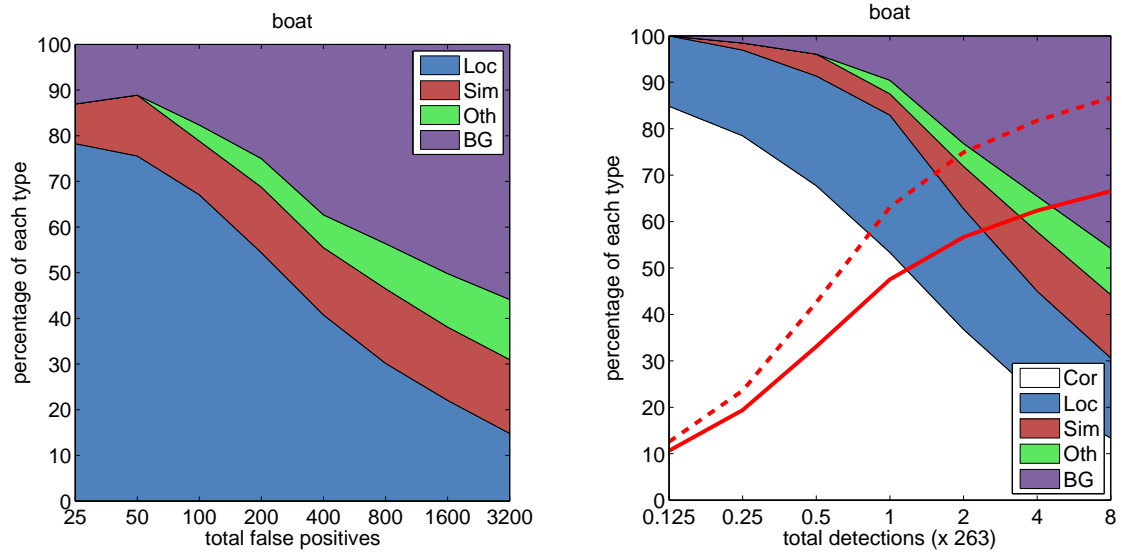


Figure 11: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

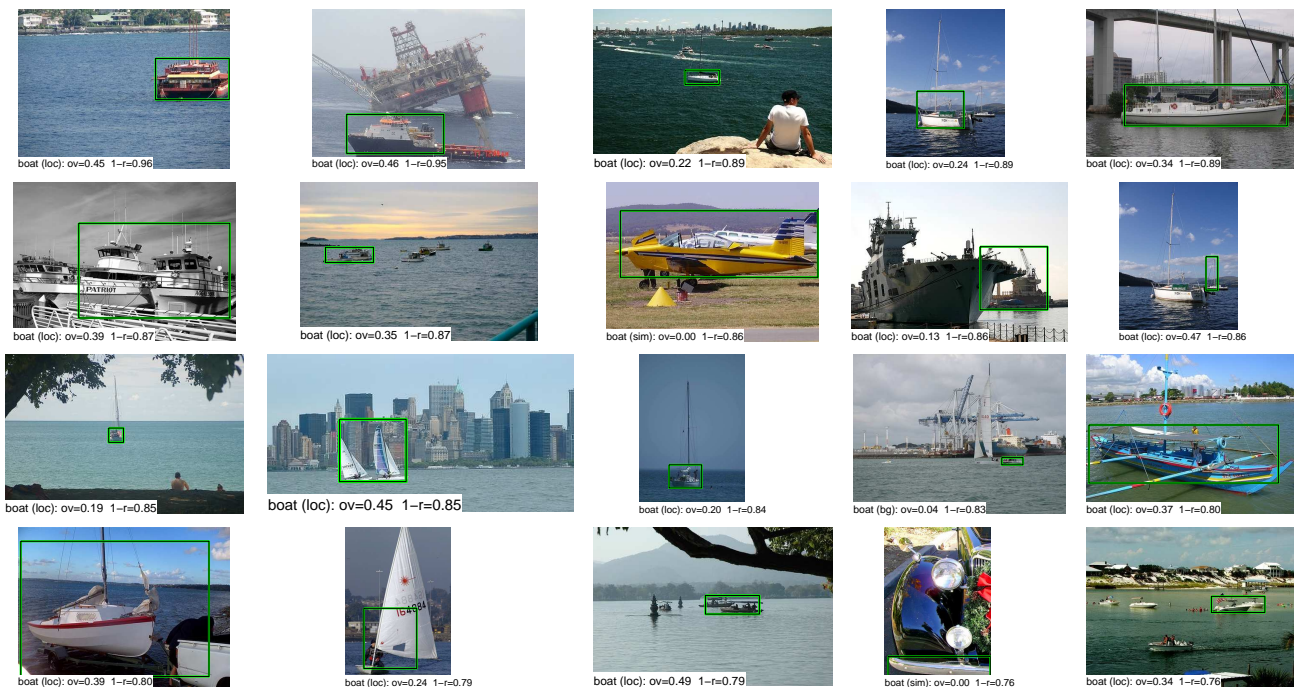


Figure 12: Examples of top boat false positives

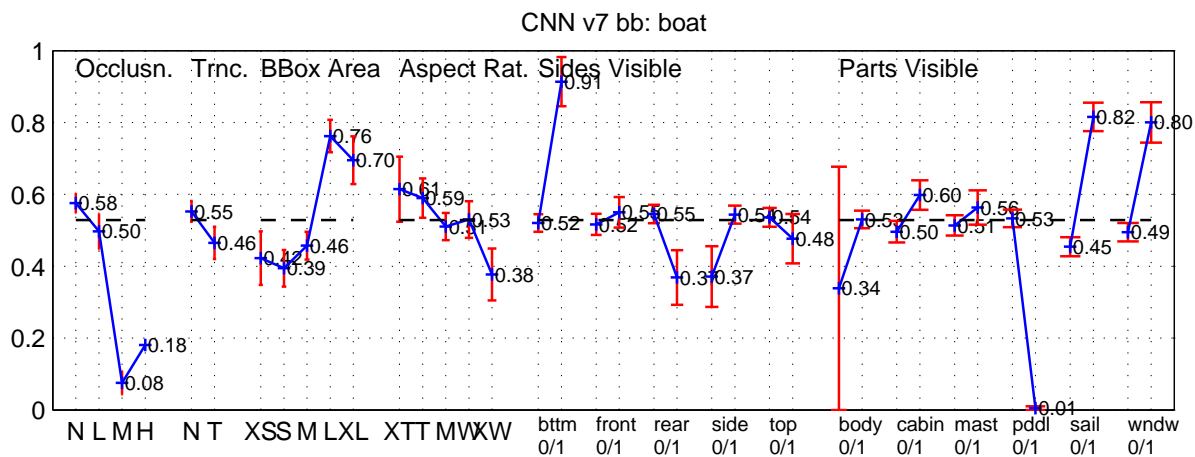


Figure 13: Analysis of boat characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.



Figure 14: **Unexpectedly difficult boat detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.

## 6 bus

Characteristics: ntotal=213 ntrunc=120

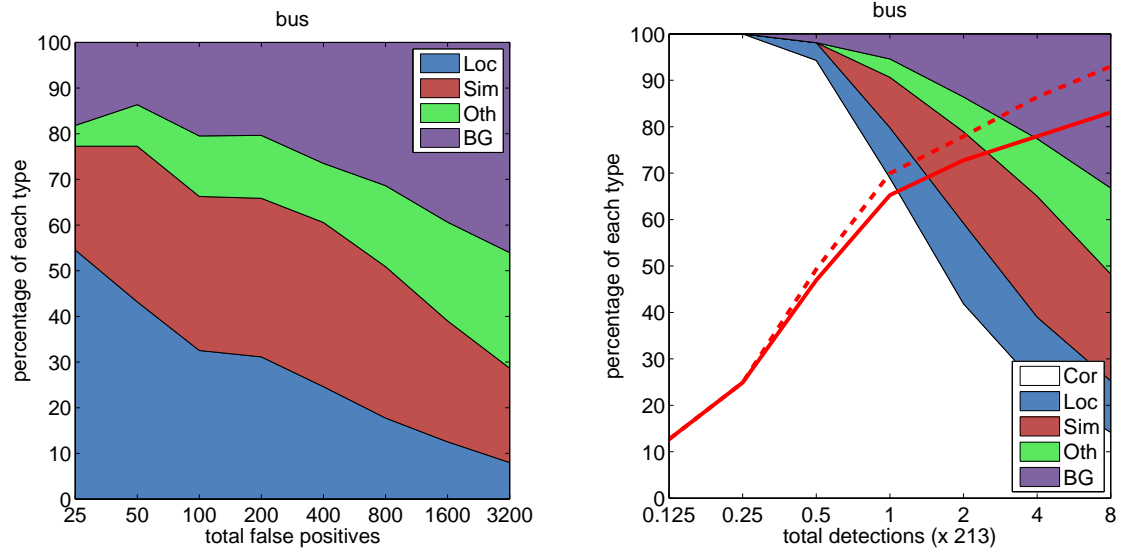


Figure 15: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).



Figure 16: Examples of top bus false positives



## 7 car

Characteristics: ntotal=1201 ntrunc=782

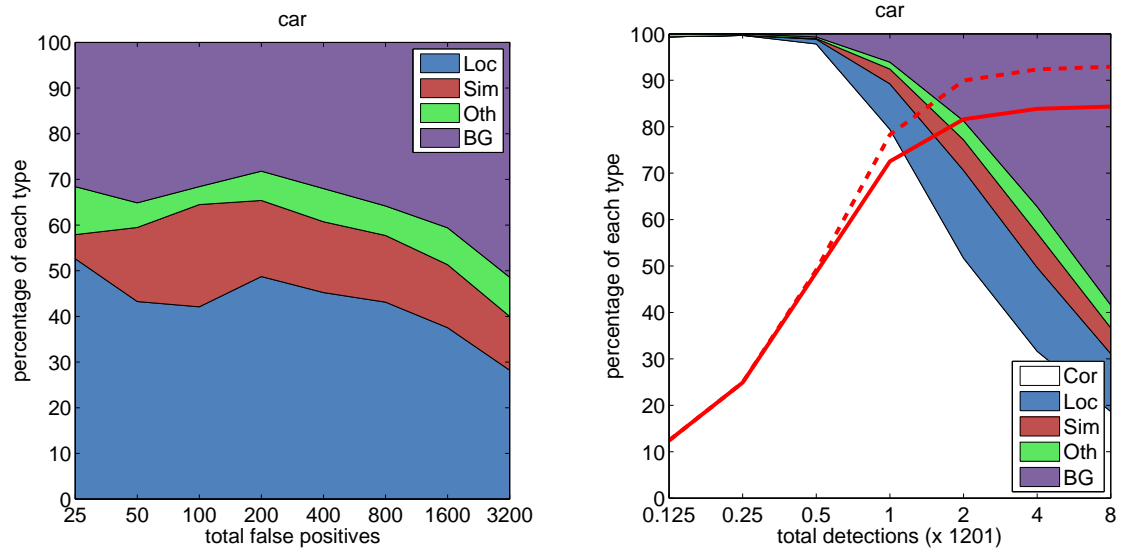


Figure 17: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

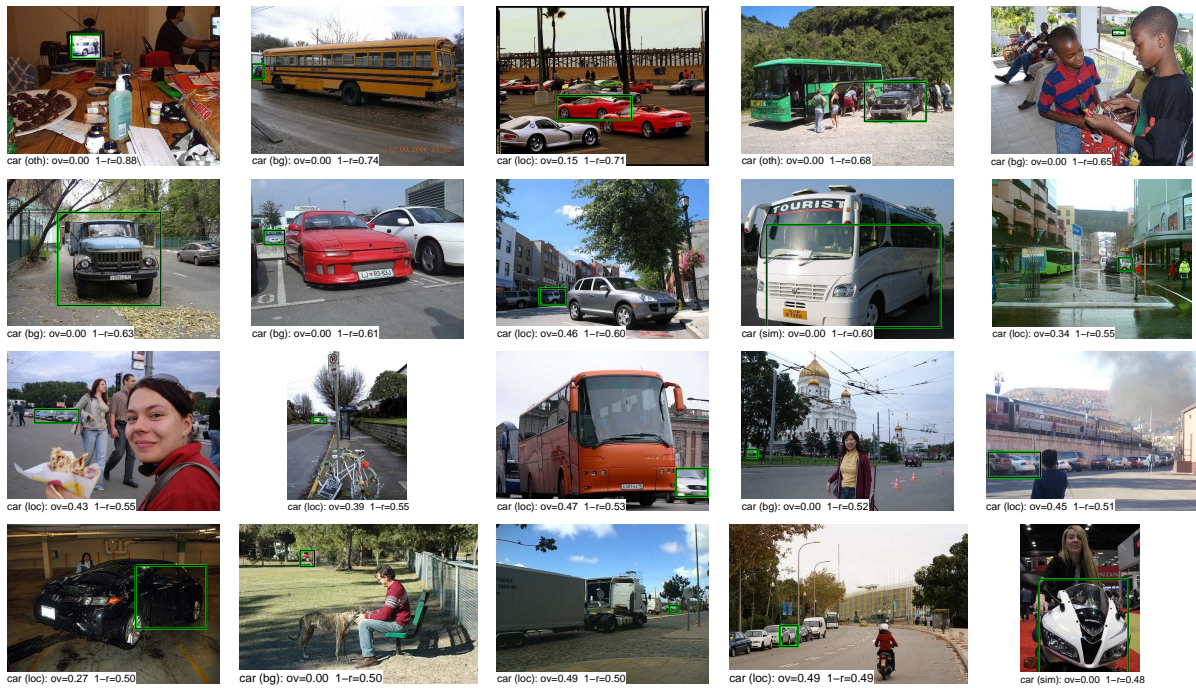


Figure 18: Examples of top car false positives

## 8 motorbike

Characteristics: ntotal=325 ntrunc=144

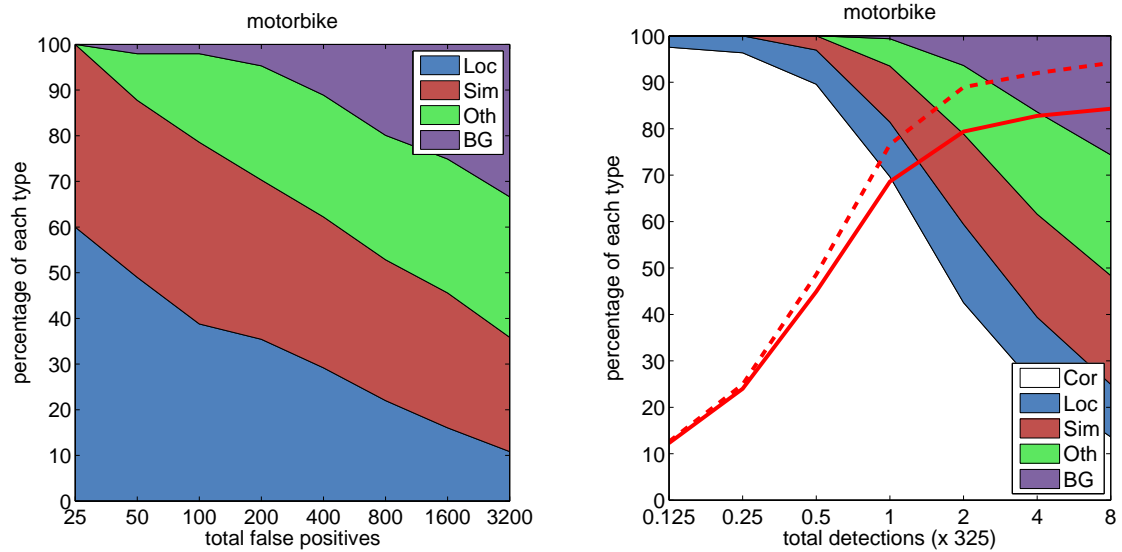


Figure 19: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

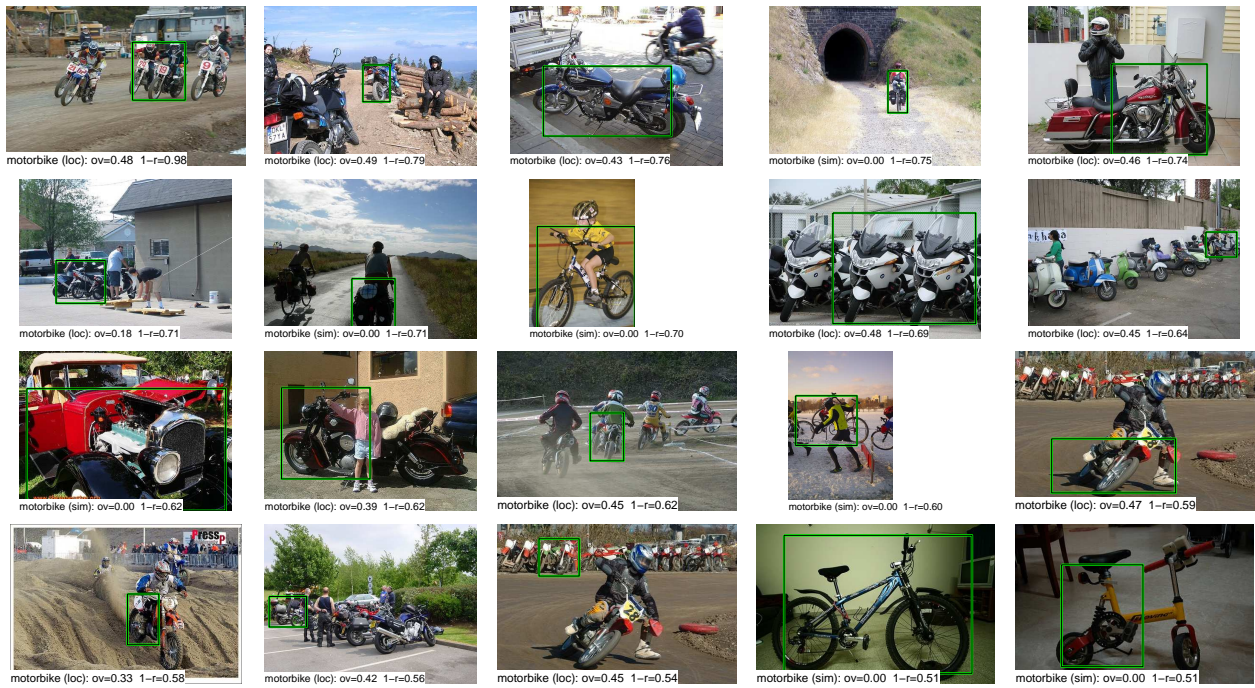


Figure 20: Examples of top motorbike false positives

## 9 train

Characteristics: ntotal=282 ntrunc=173

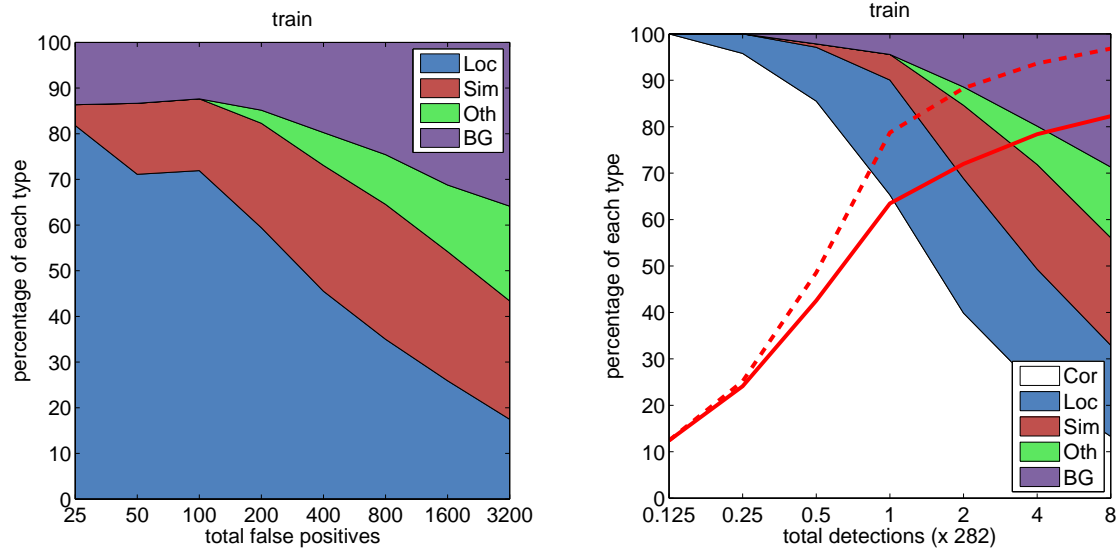


Figure 21: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).



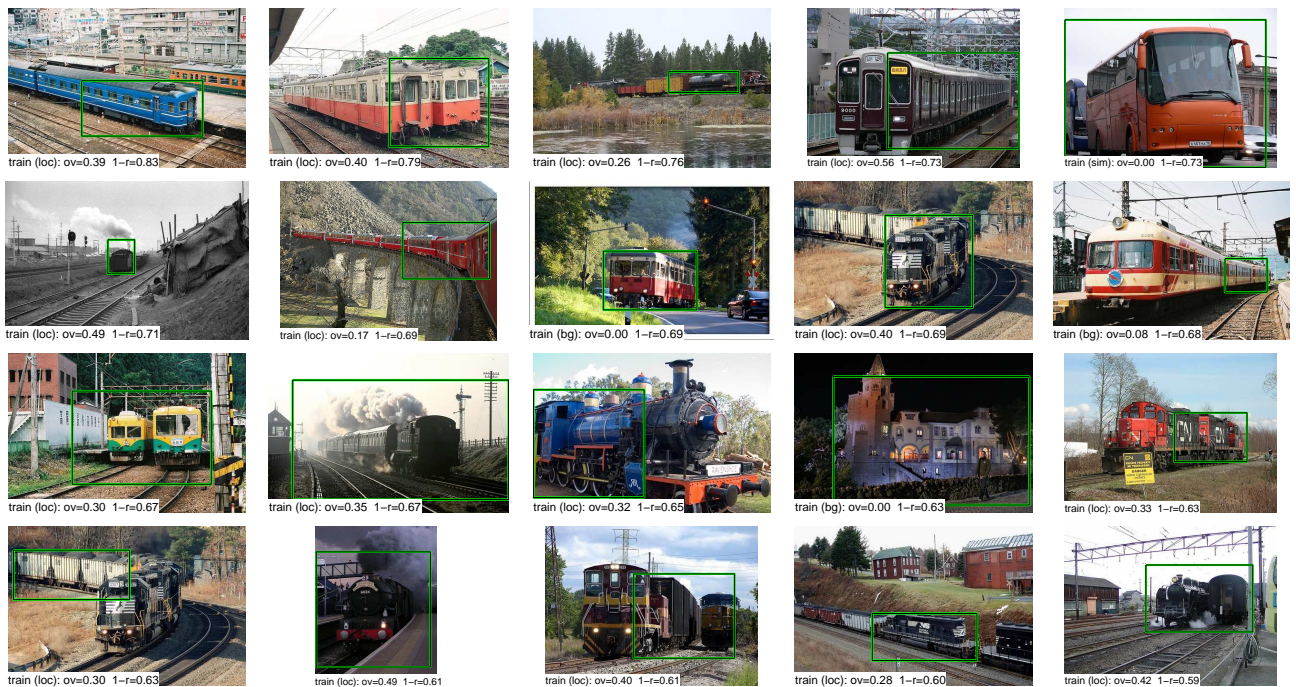


Figure 22: Examples of top train false positives

## 10 bird

Characteristics: ntotal=459 ntrunc=112

occlevel: None=328 Low=109 Med=19 High=3

side visible:

bottom: Yes=53 No=406

front: Yes=126 No=333

rear: Yes=63 No=396

side: Yes=387 No=72

top: Yes=53 No=406

part visible:

body: Yes=432 No=27

face: Yes=328 No=131

beak: Yes=371 No=88

leg: Yes=265 No=194

tail: Yes=328 No=131

wing: Yes=366 No=93

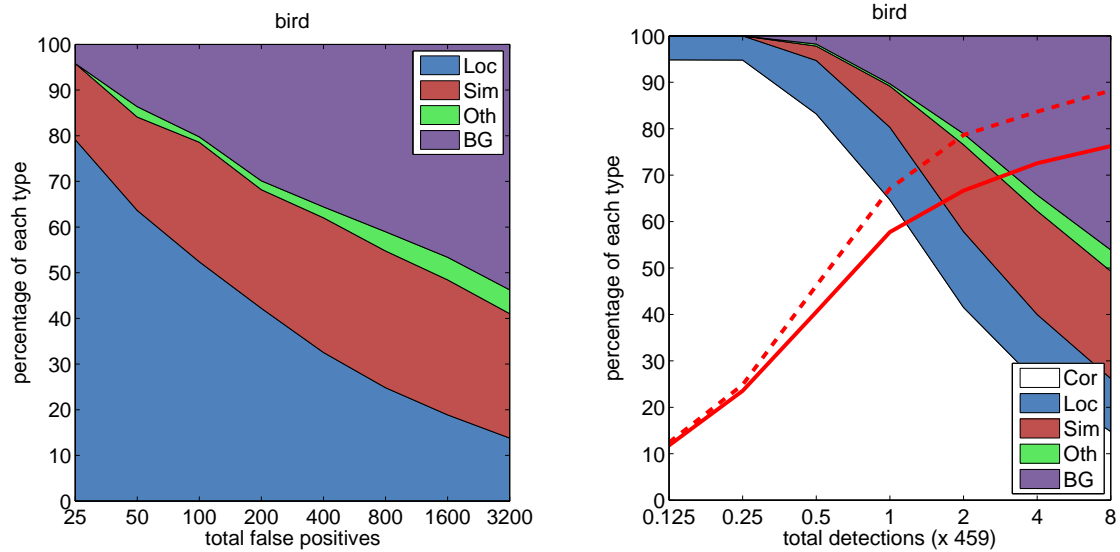


Figure 23: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

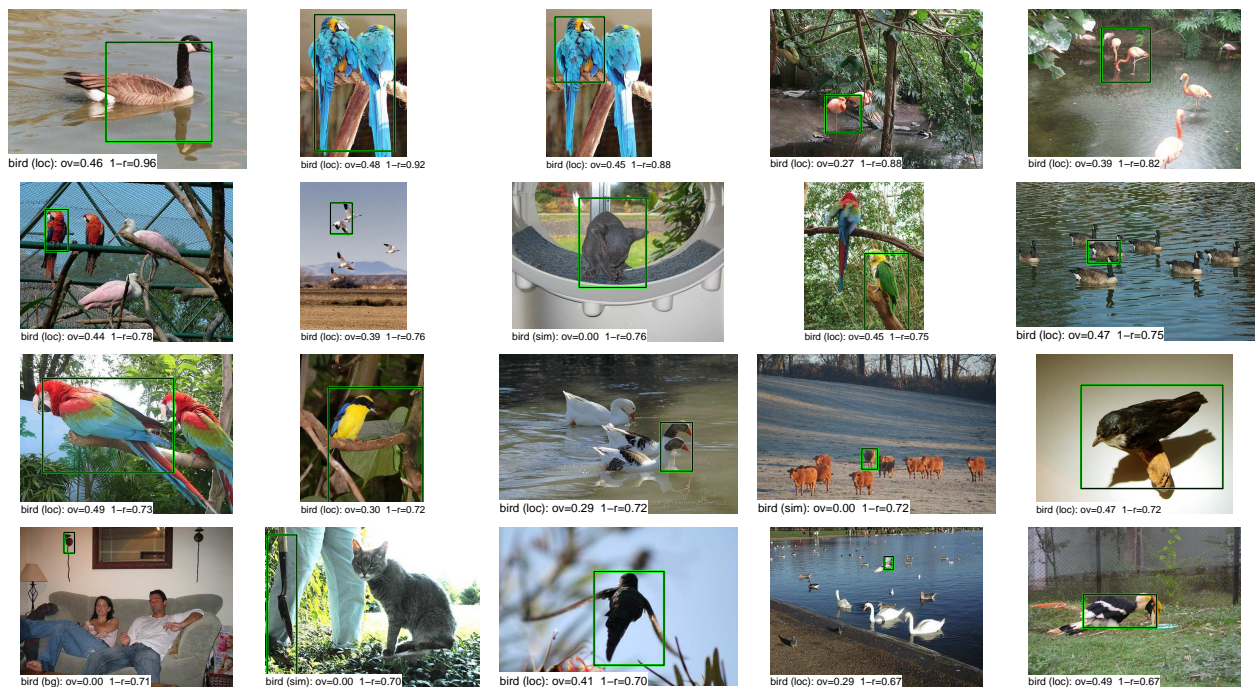


Figure 24: Examples of top bird false positives

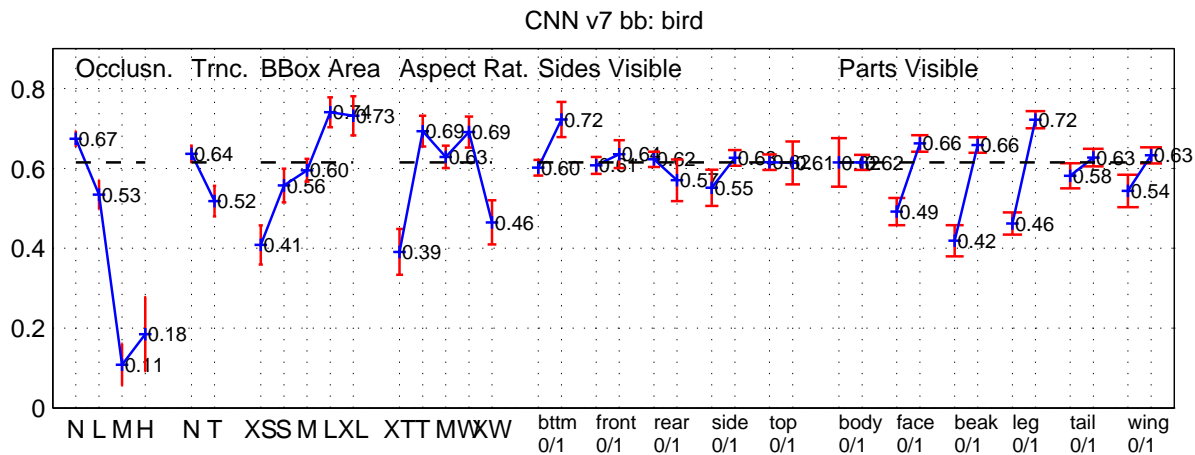


Figure 25: Analysis of bird characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.

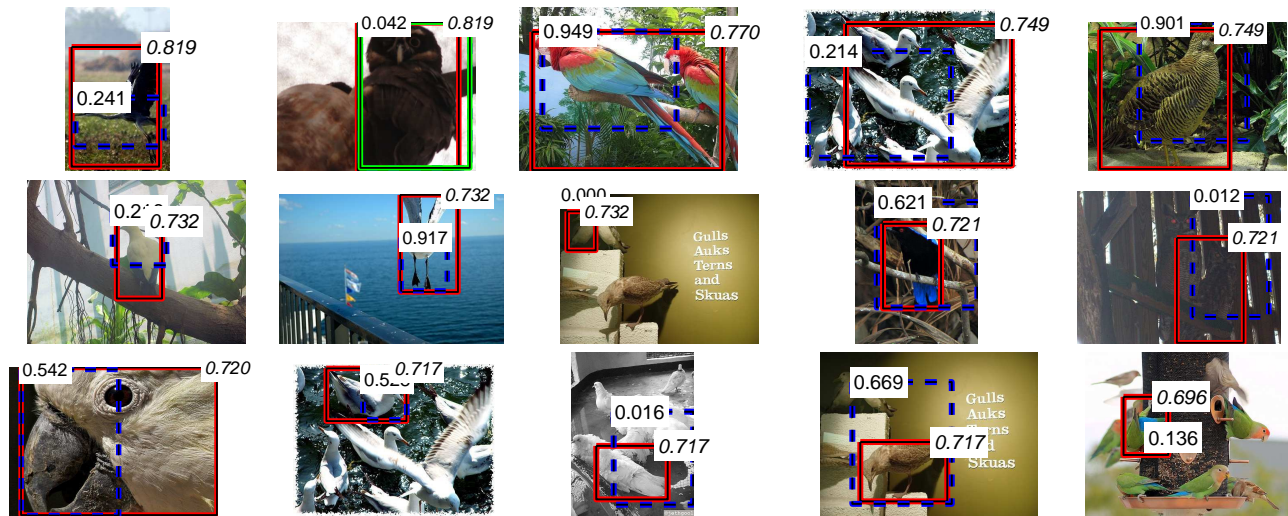


Figure 26: **Unexpectedly difficult bird detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.

## 11 cat

Characteristics: ntotal=358 ntrunc=156  
occlevel: None=269 Low=53 Med=33 High=3  
side visible:  
  bottom: Yes=41 No=317  
  front: Yes=150 No=208  
  rear: Yes=20 No=338  
  side: Yes=275 No=83  
  top: Yes=38 No=320  
part visible:  
  body: Yes=330 No=28  
  ear: Yes=342 No=16  
  face: Yes=290 No=68  
  leg: Yes=225 No=133  
  tail: Yes=135 No=223

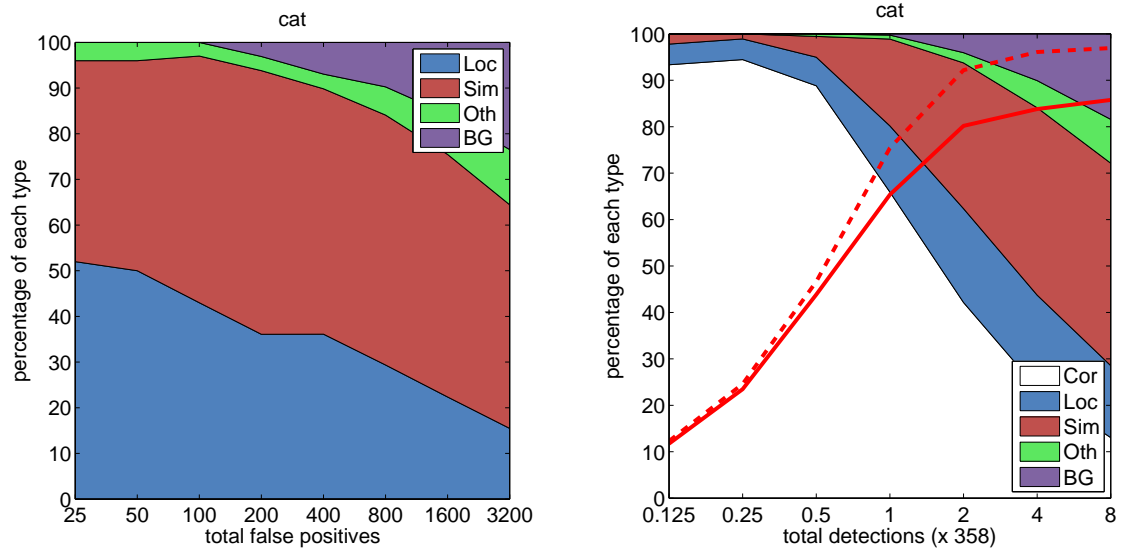


Figure 27: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

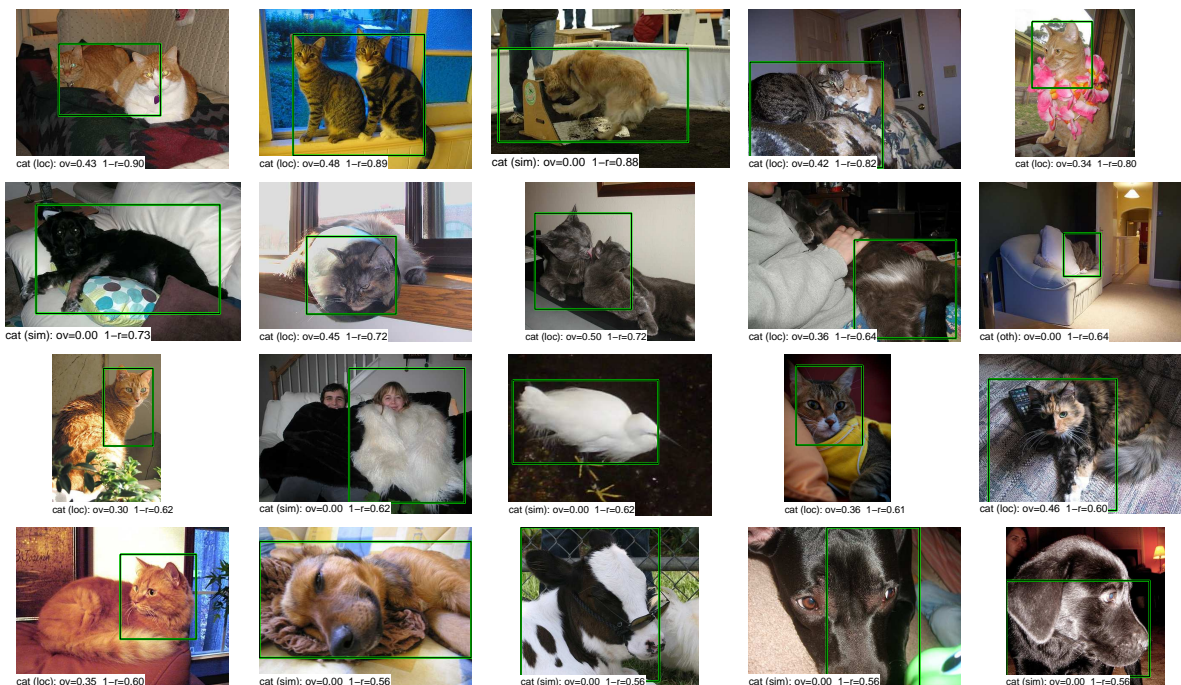


Figure 28: Examples of top cat false positives

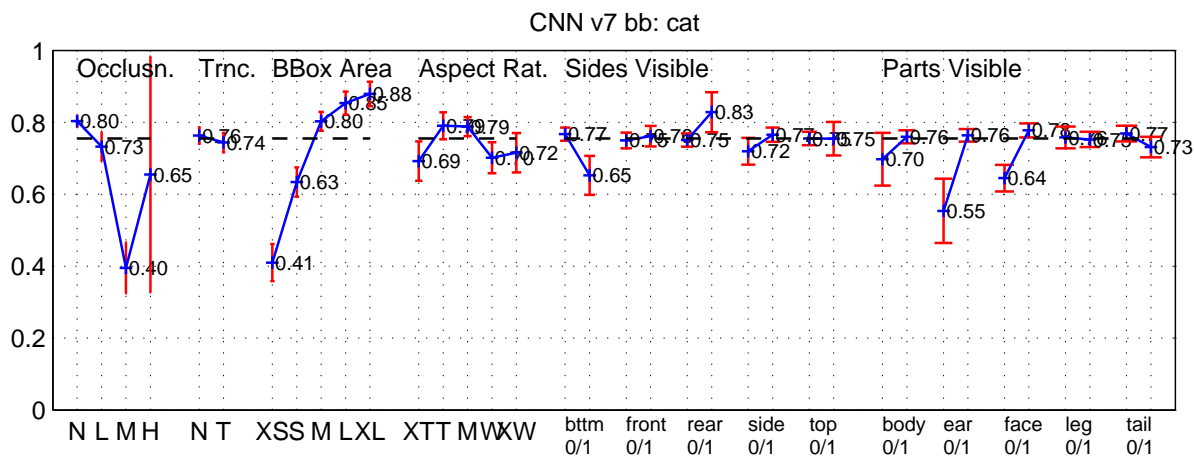


Figure 29: Analysis of cat characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.



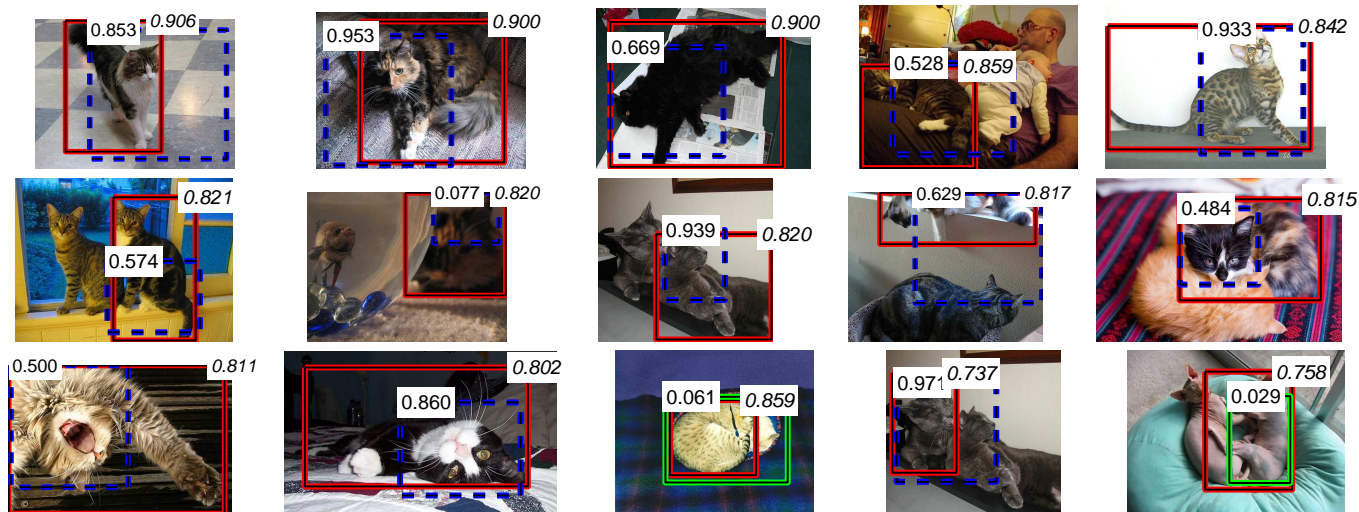


Figure 30: **Unexpectedly difficult cat detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.

## 12 cow

Characteristics: ntotal=244 ntrunc=108

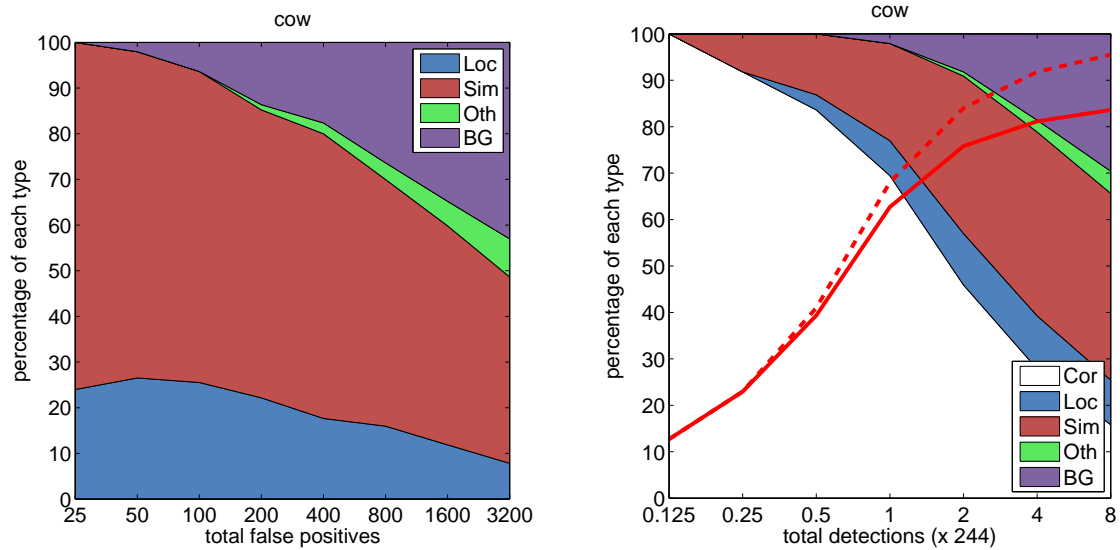


Figure 31: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).



Figure 32: Examples of top cow false positives

## 13 dog

Characteristics: ntotal=489 ntrunc=230

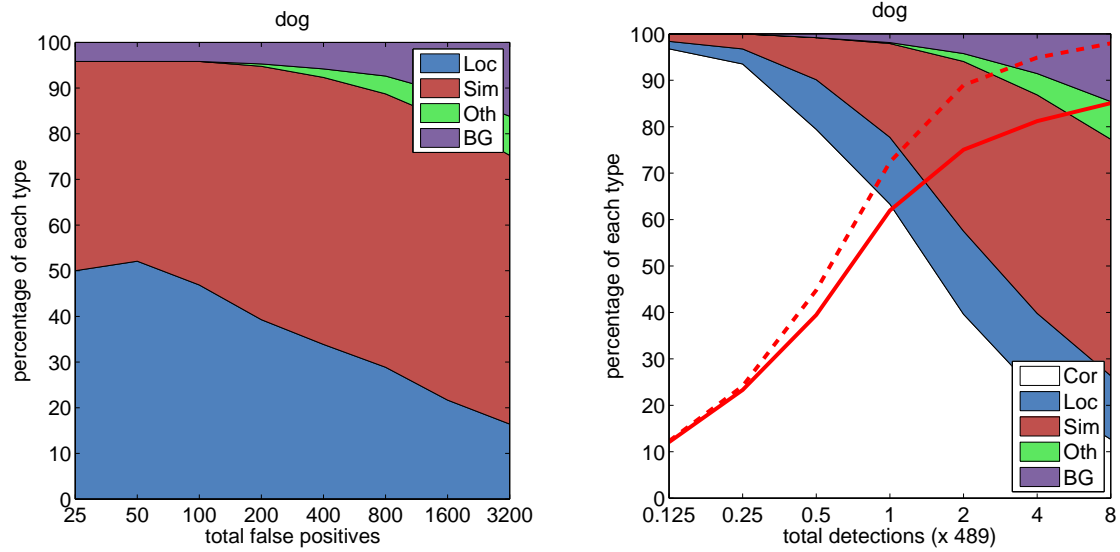


Figure 33: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

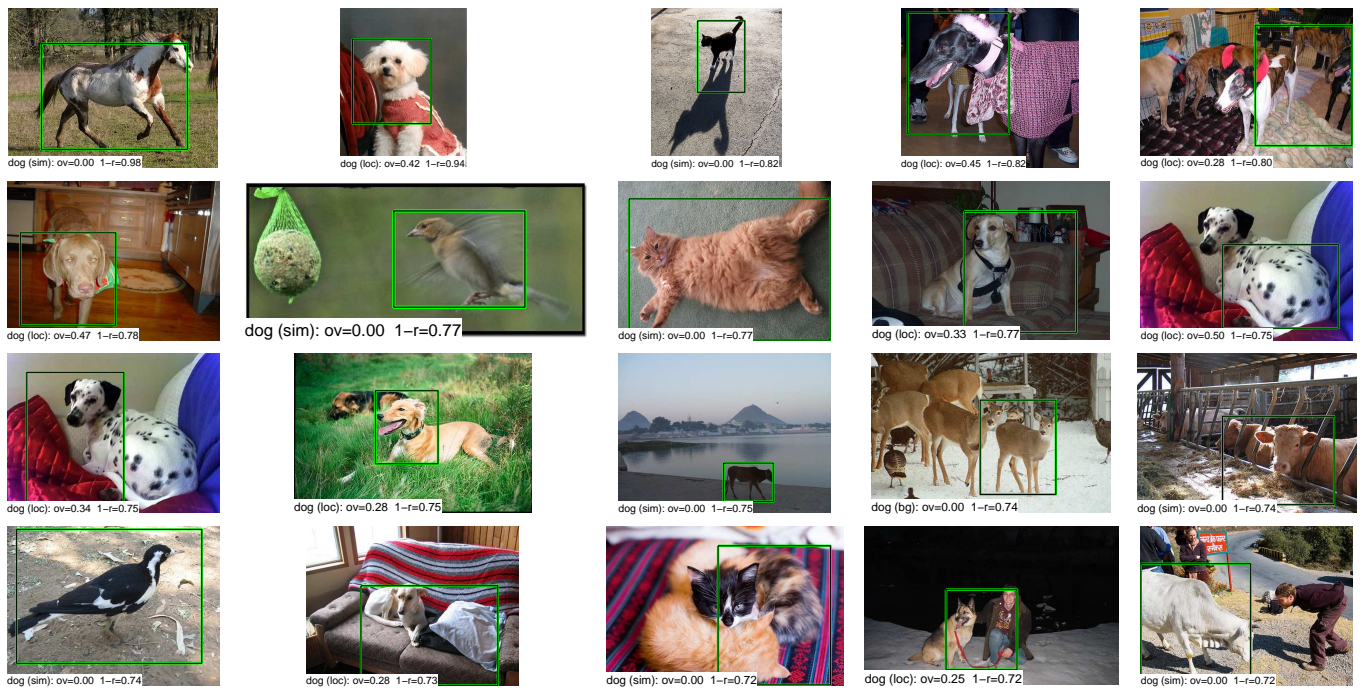


Figure 34: Examples of top dog false positives

## 14 horse

Characteristics: ntotal=348 ntrunc=145



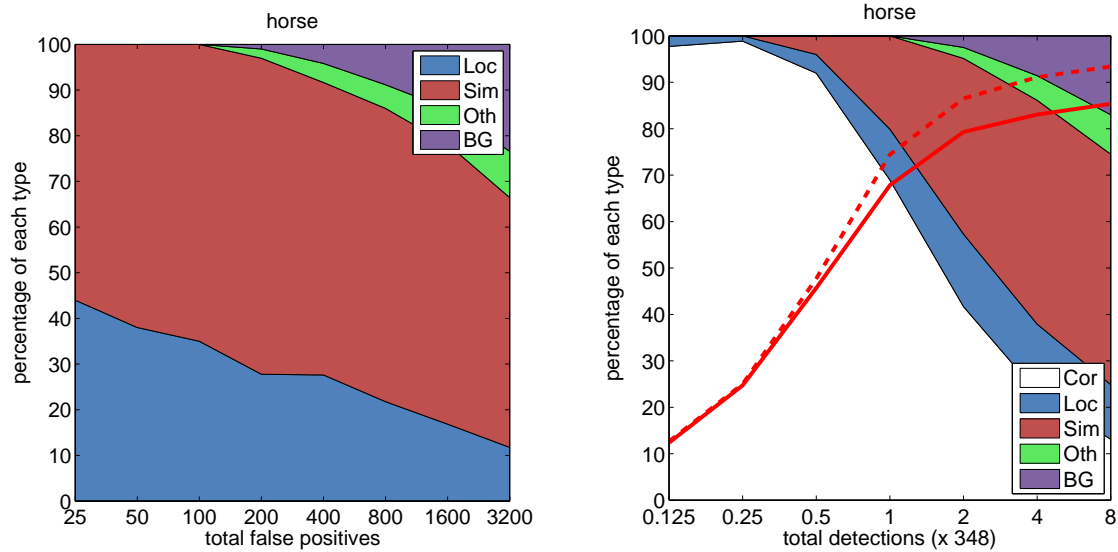


Figure 35: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).



Figure 36: Examples of top horse false positives

## 15 sheep

Characteristics: ntotal=242 ntrunc=111

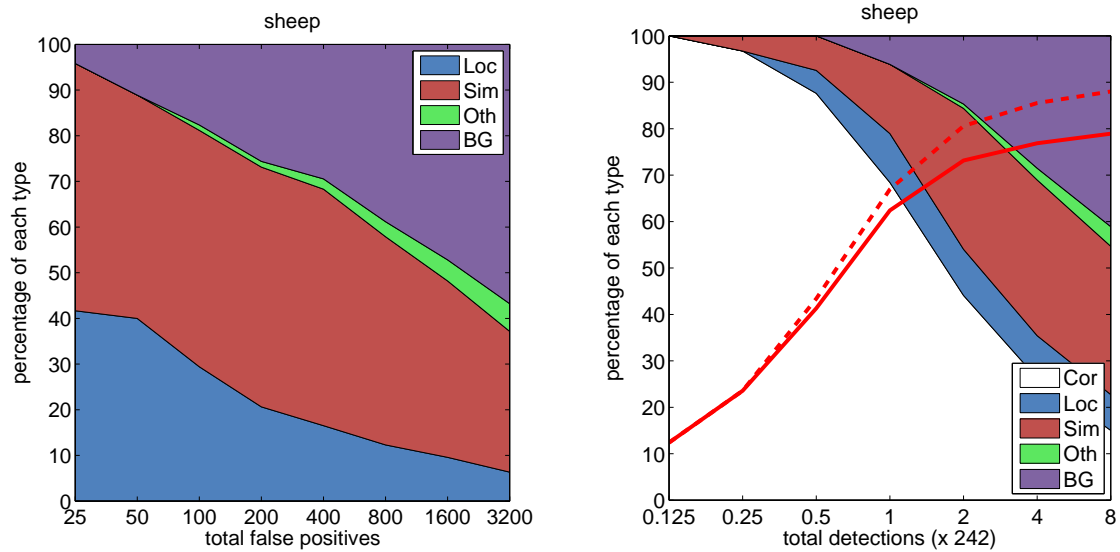


Figure 37: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).



Figure 38: Examples of top sheep false positives

## 16 bottle

Characteristics: ntotal=469 ntrunc=251

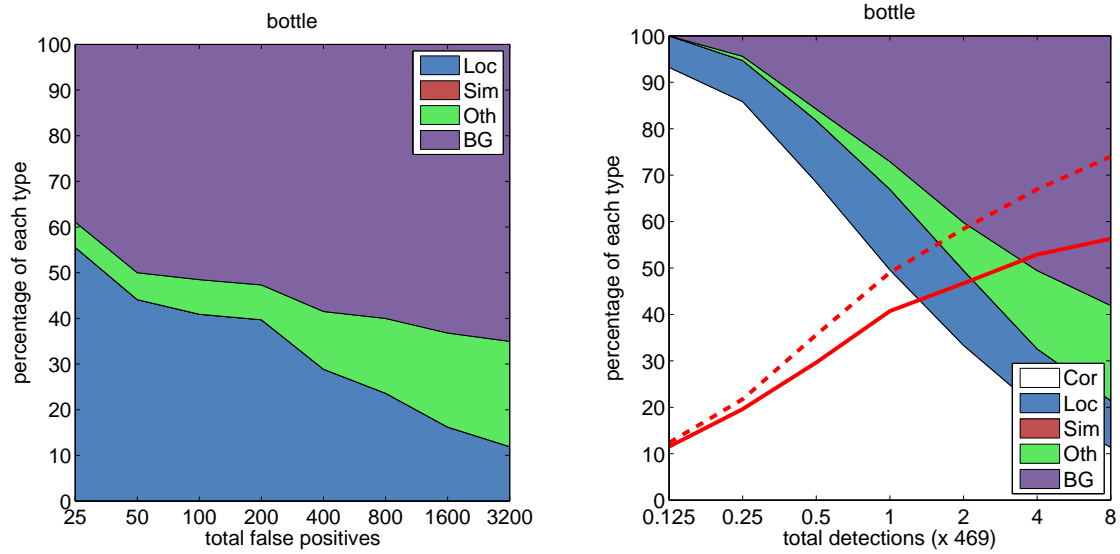


Figure 39: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

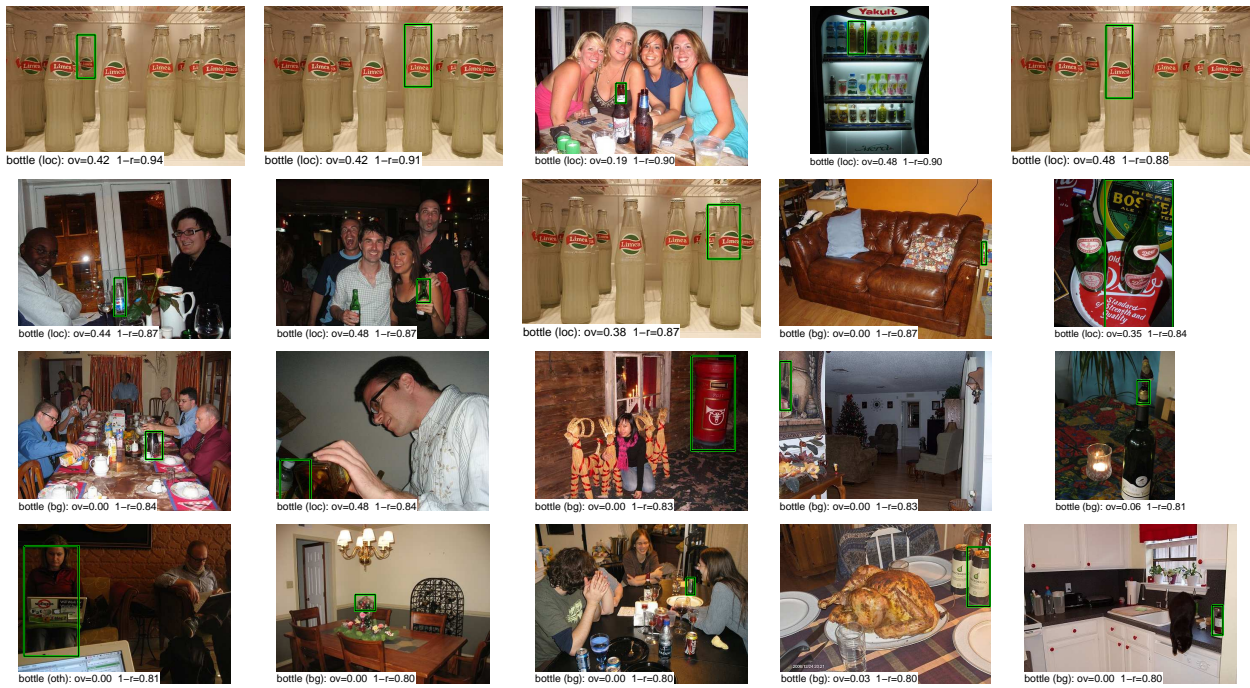


Figure 40: Examples of top bottle false positives



## 17 chair

Characteristics: ntotal=756 ntrunc=956

occlevel: None=233 Low=311 Med=174 High=38

side visible:

bottom: Yes=6 No=750

front: Yes=366 No=390

rear: Yes=239 No=517

side: Yes=568 No=188

top: Yes=367 No=389

part visible:

backrest: Yes=731 No=25

cushion: Yes=641 No=115

handrest: Yes=278 No=478

leg: Yes=507 No=249

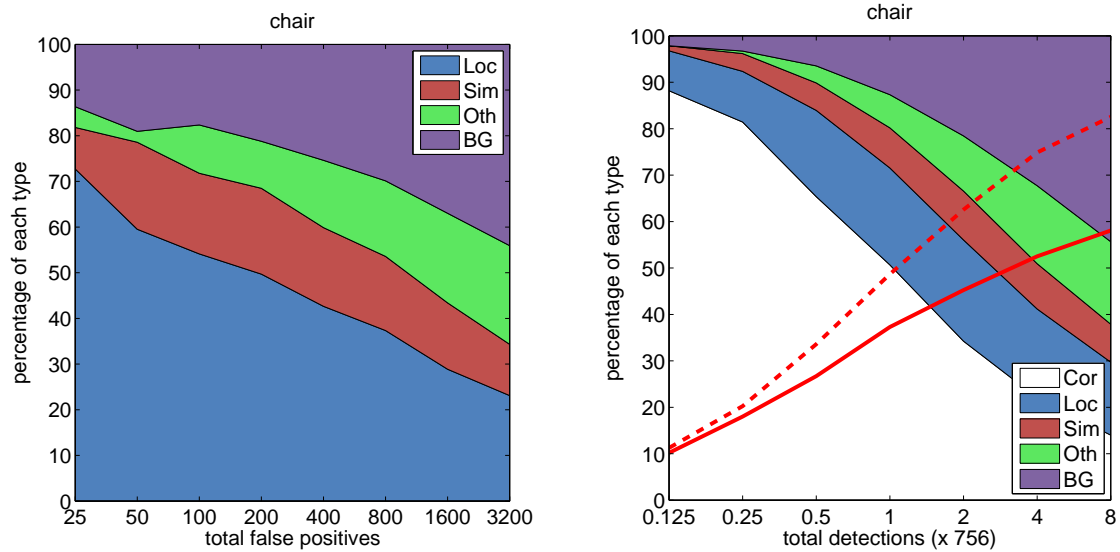


Figure 41: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

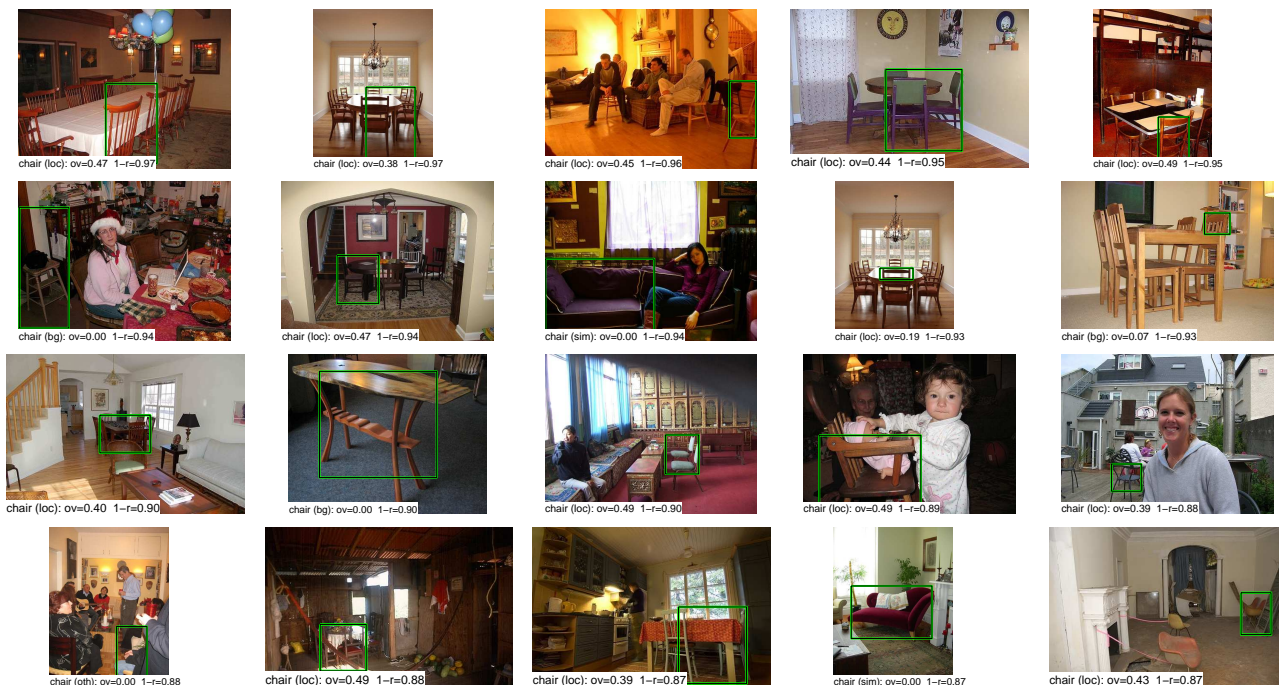


Figure 42: Examples of top chair false positives

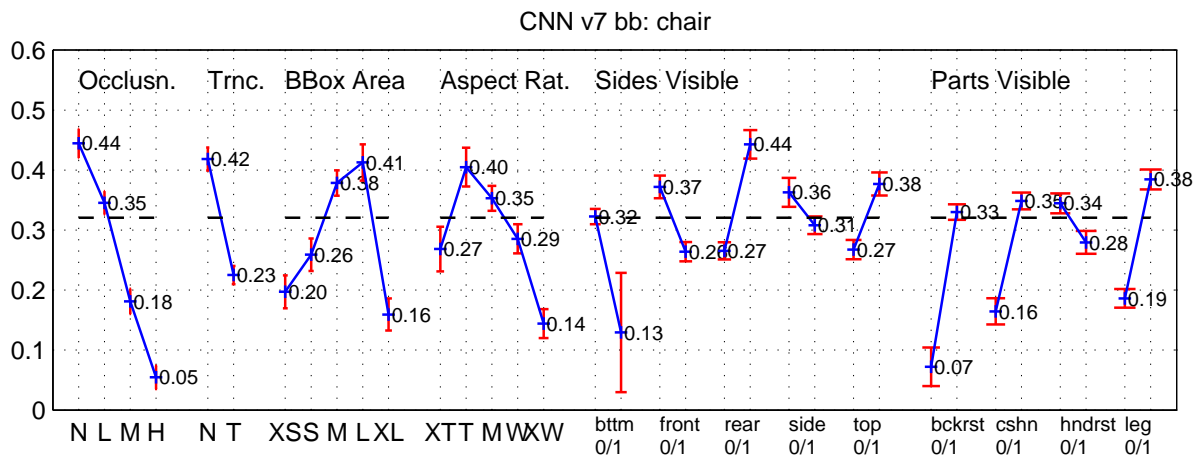


Figure 43: Analysis of chair characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.

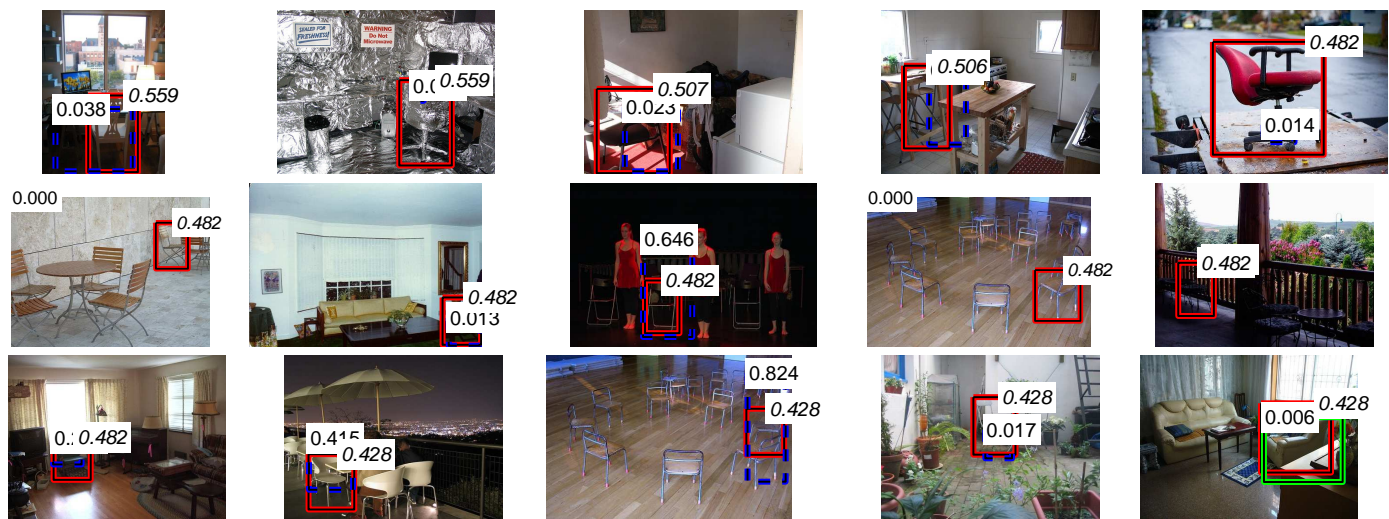


Figure 44: **Unexpectedly difficult chair detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.

## 18 diningtable

Characteristics: ntotal=206 ntrunc=251  
occlevel: None=12 Low=89 Med=90 High=15  
side visible:  
  side: Yes=132 No=74  
  top: Yes=196 No=10  
part visible:  
  tableleg: Yes=83 No=123  
  tabletop: Yes=205 No=1

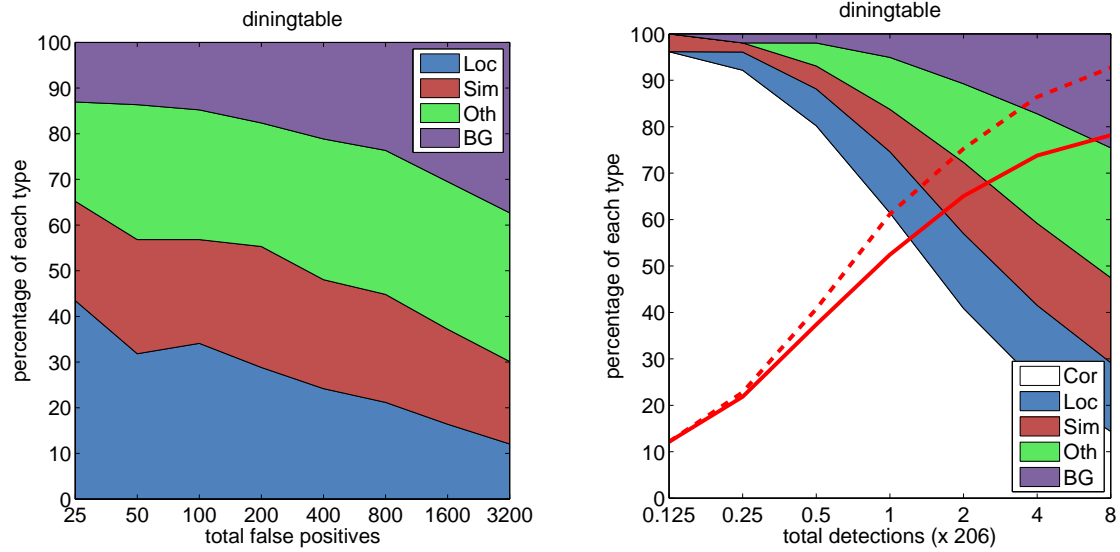


Figure 45: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

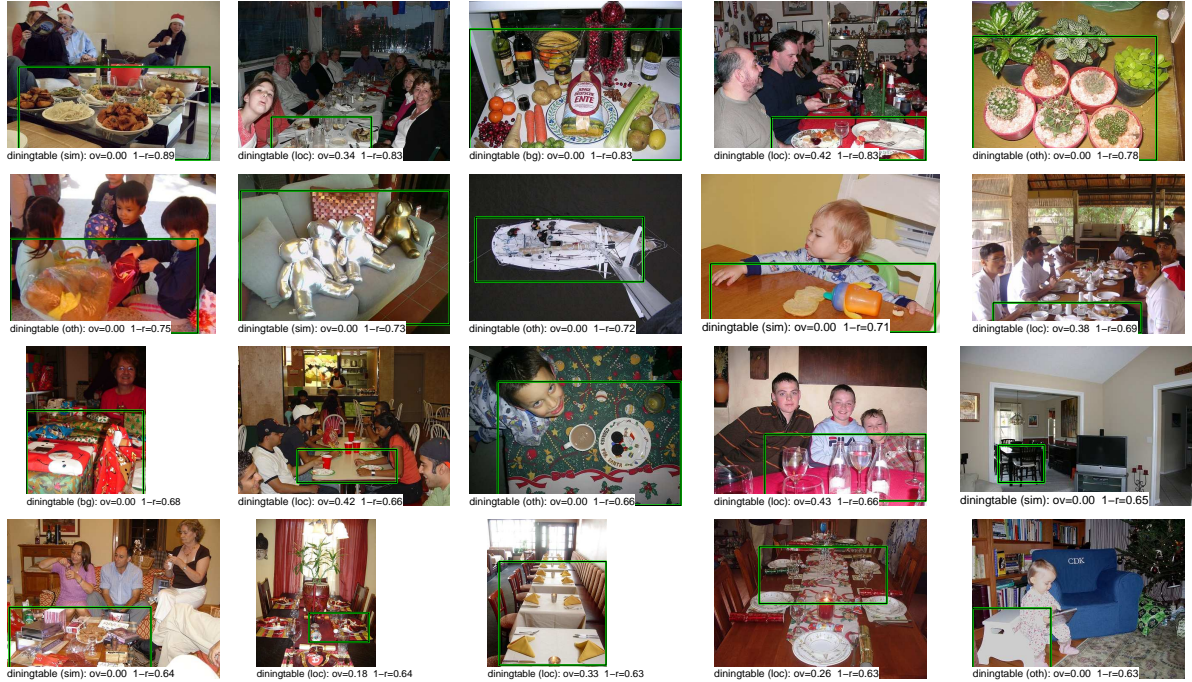


Figure 46: Examples of top diningtable false positives

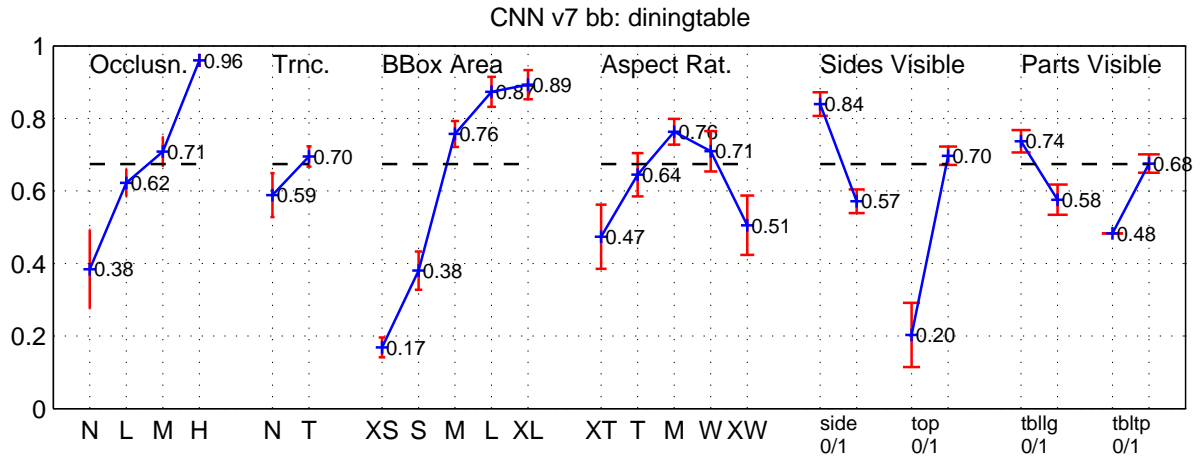


Figure 47: Analysis of diningtable characteristics: APn (+) with standard error bars (red). Black dashed lines indicate overall APn. See paper for further details.

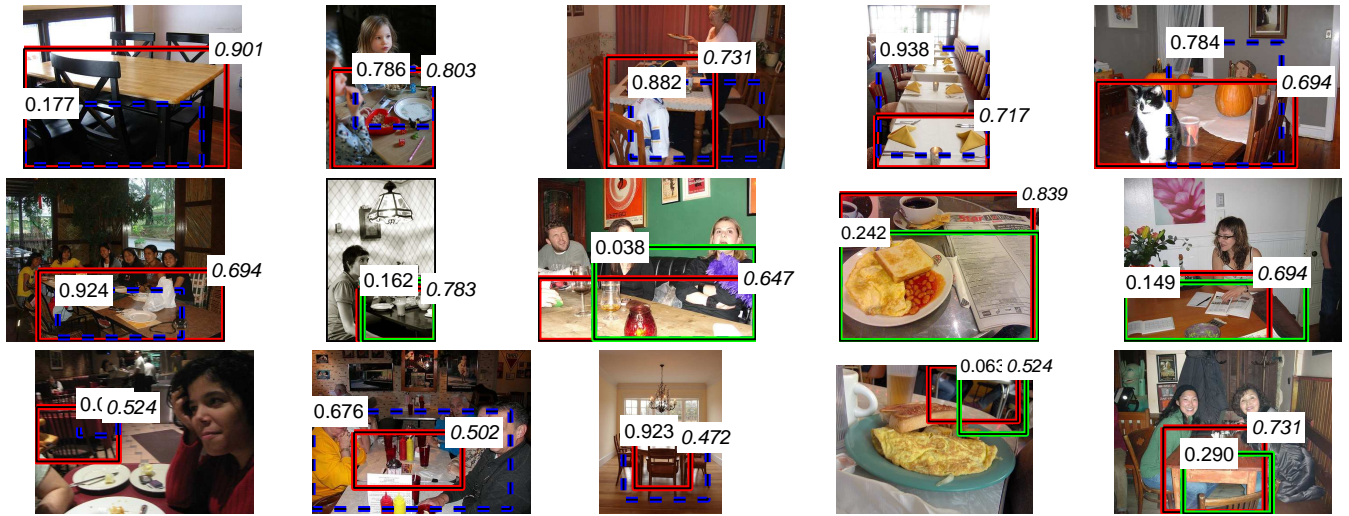


Figure 48: **Unexpectedly difficult diningtable detections:** Ground truth object is red; predicted confidence in italics; green box is highest scoring detection; blue box is highest scoring with overlap; detection confidence in upper-left corner.



## 19 pottedplant

Characteristics: ntotal=480 ntrunc=246

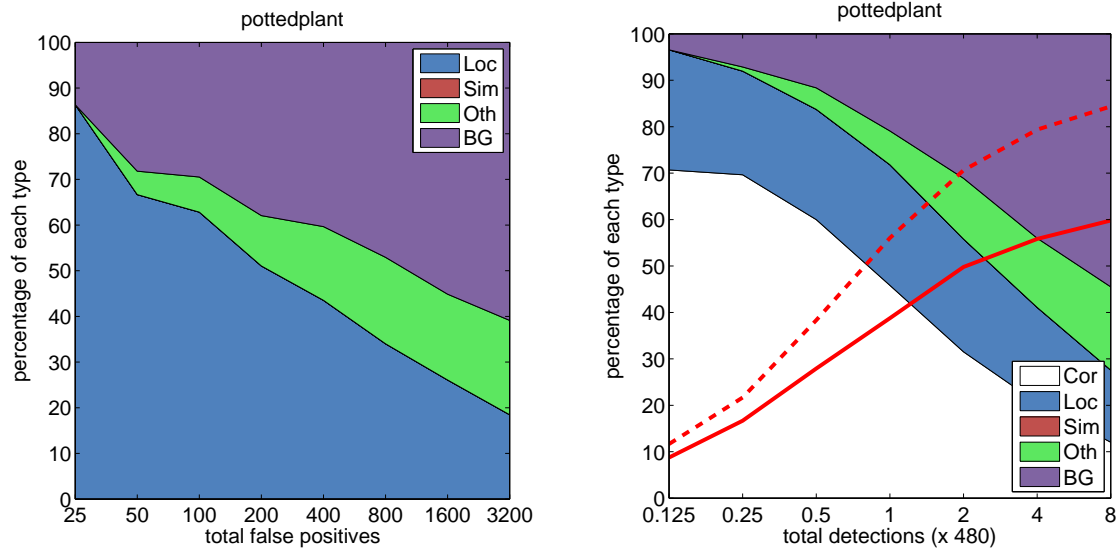


Figure 49: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

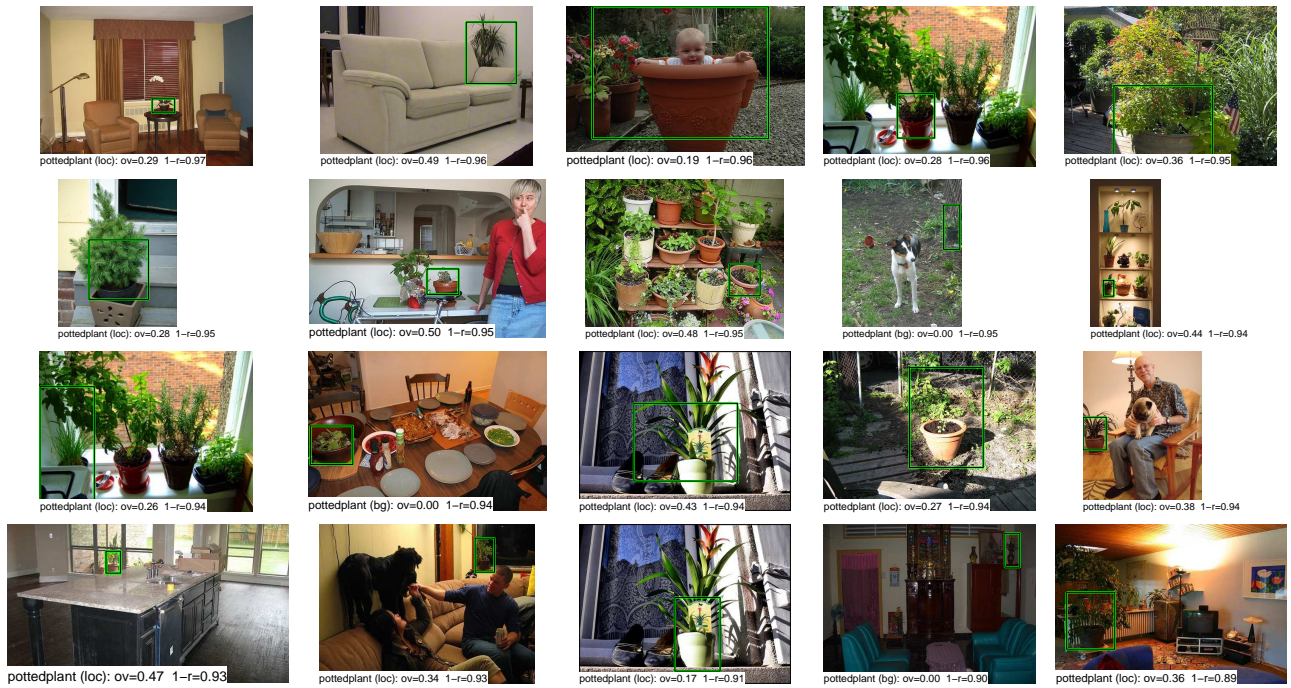


Figure 50: Examples of top pottedplant false positives

## 20 sofa

Characteristics: ntotal=239 ntrunc=267

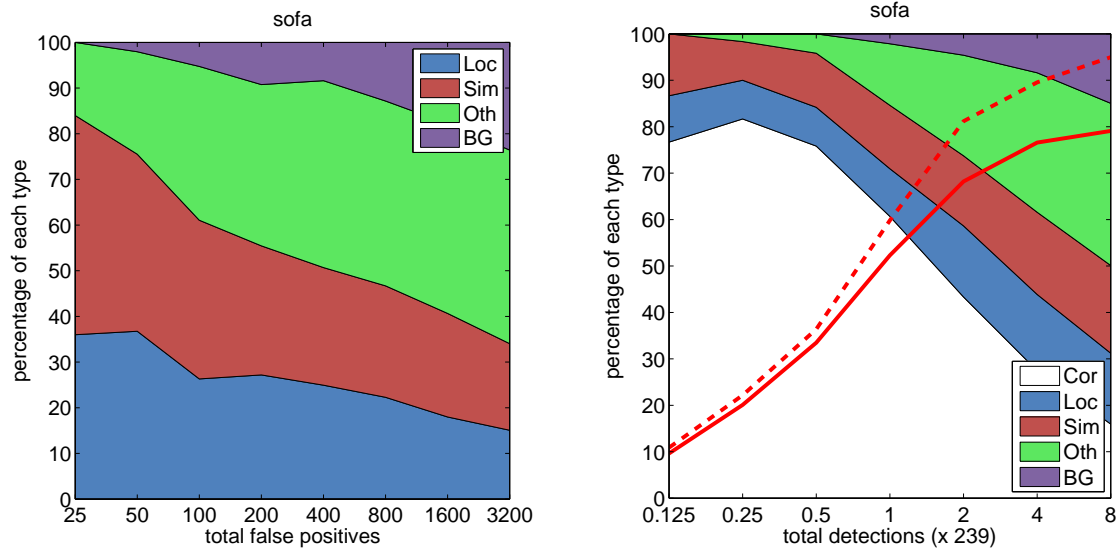


Figure 51: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

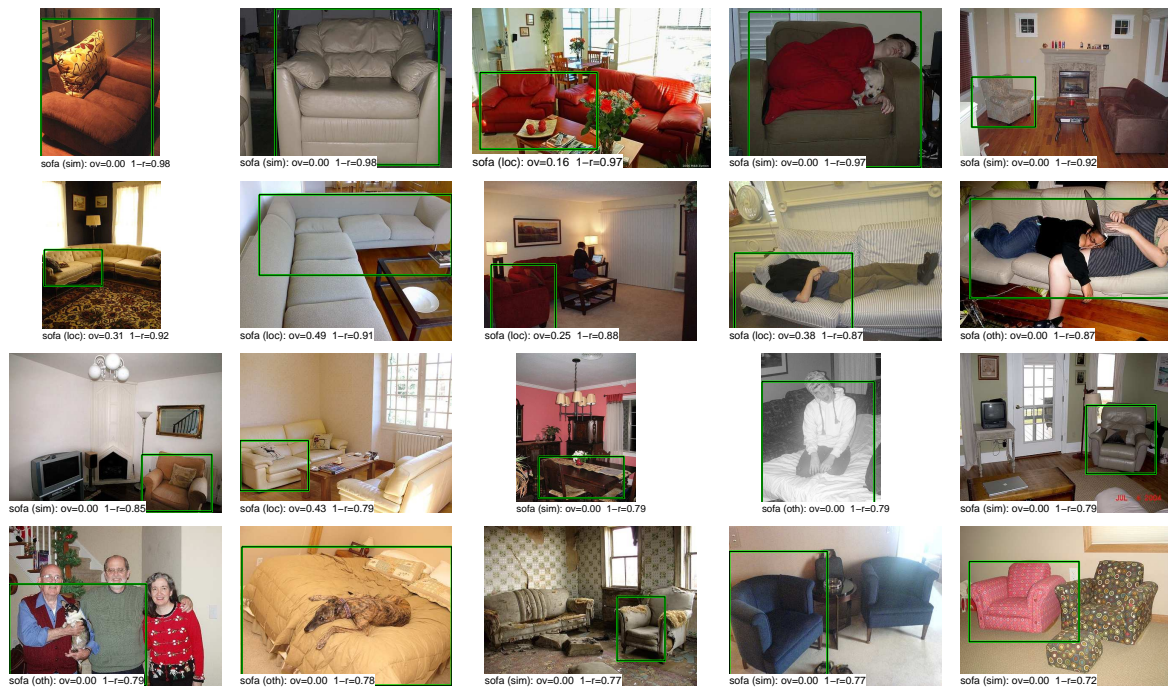


Figure 52: Examples of top sofa false positives

## 21 tvmonitor

Characteristics: ntotal=308 ntrunc=104

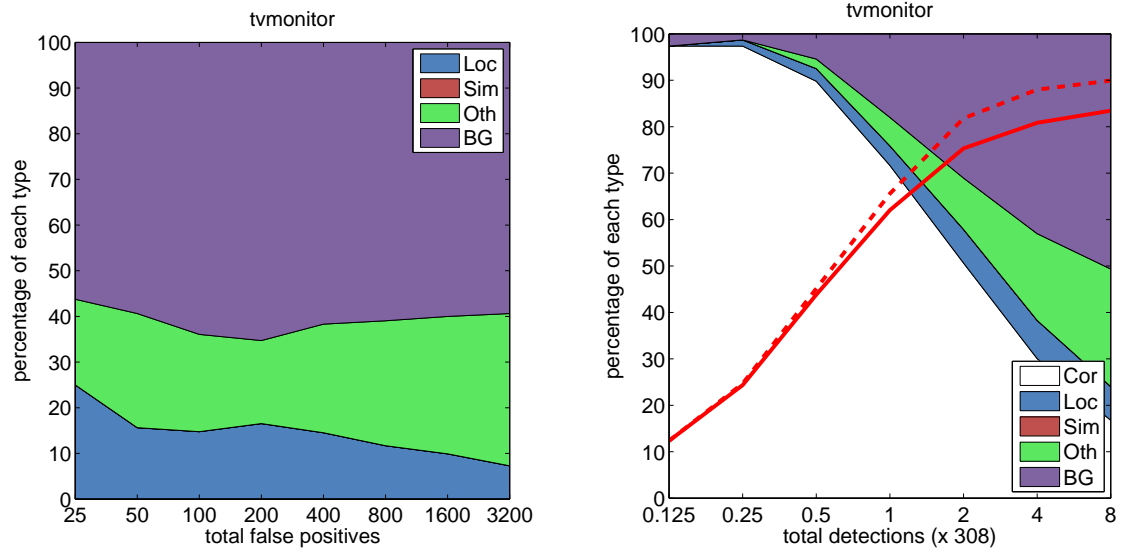


Figure 53: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).



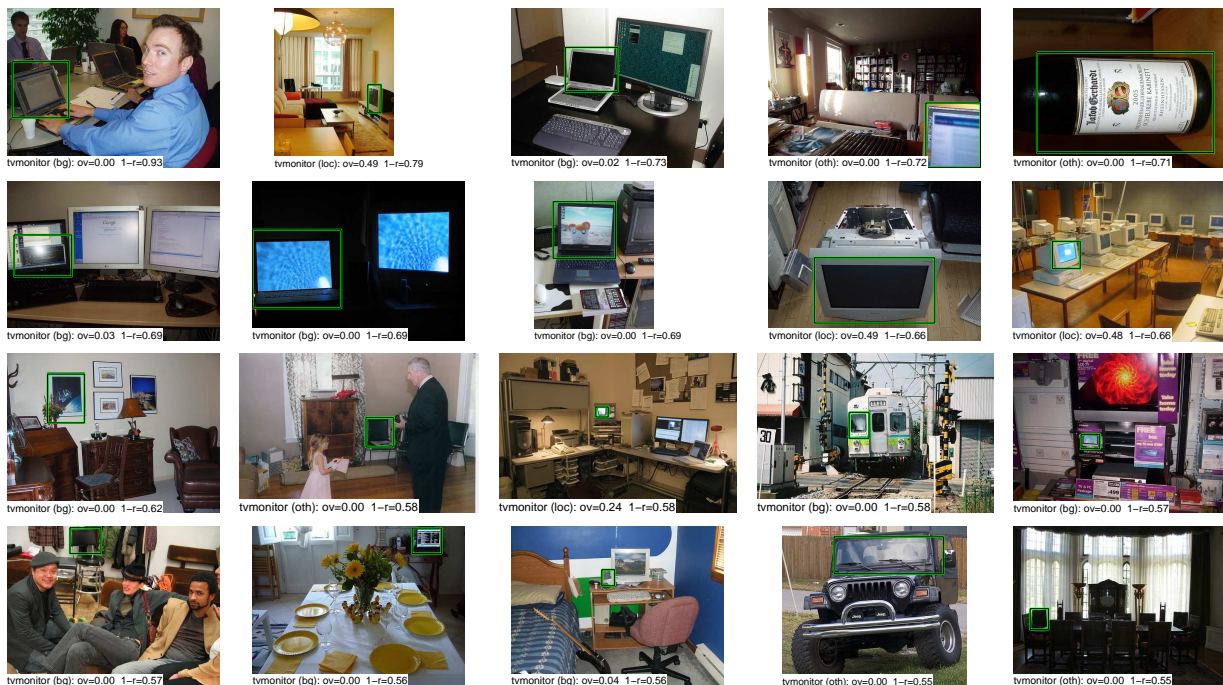


Figure 54: Examples of top tvmonitor false positives

## 22 person

Characteristics: ntotal=4528 ntrunc=3321

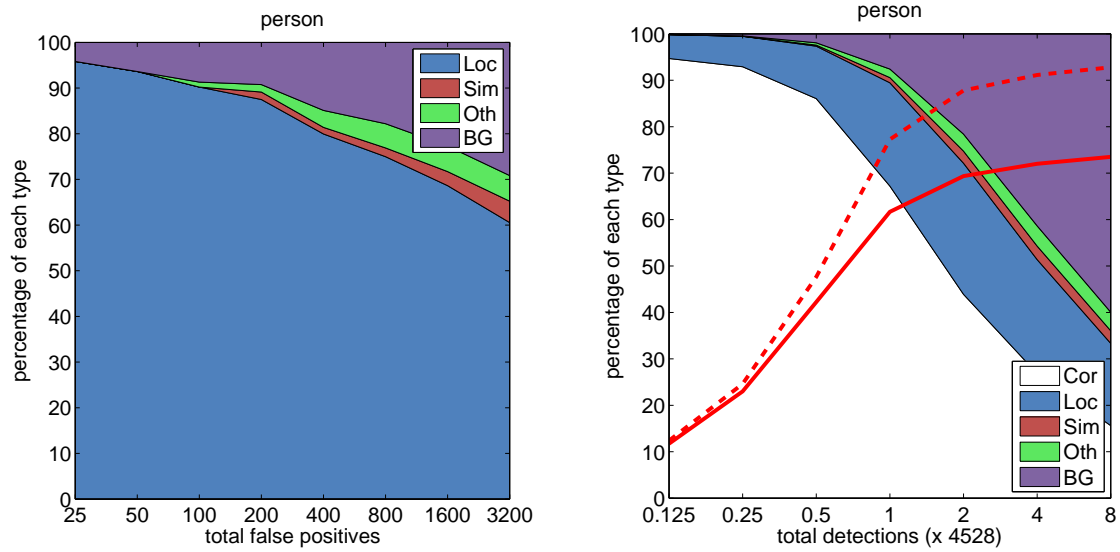


Figure 55: **False positive/detection trends with rank.** Left: stacked area plot showing fraction of FP of each type as the total number of FP increase. Right: type of detection as number of detections increases; line plots show recall as function of the number of objects (dashed=weak localization, solid=strong localization).

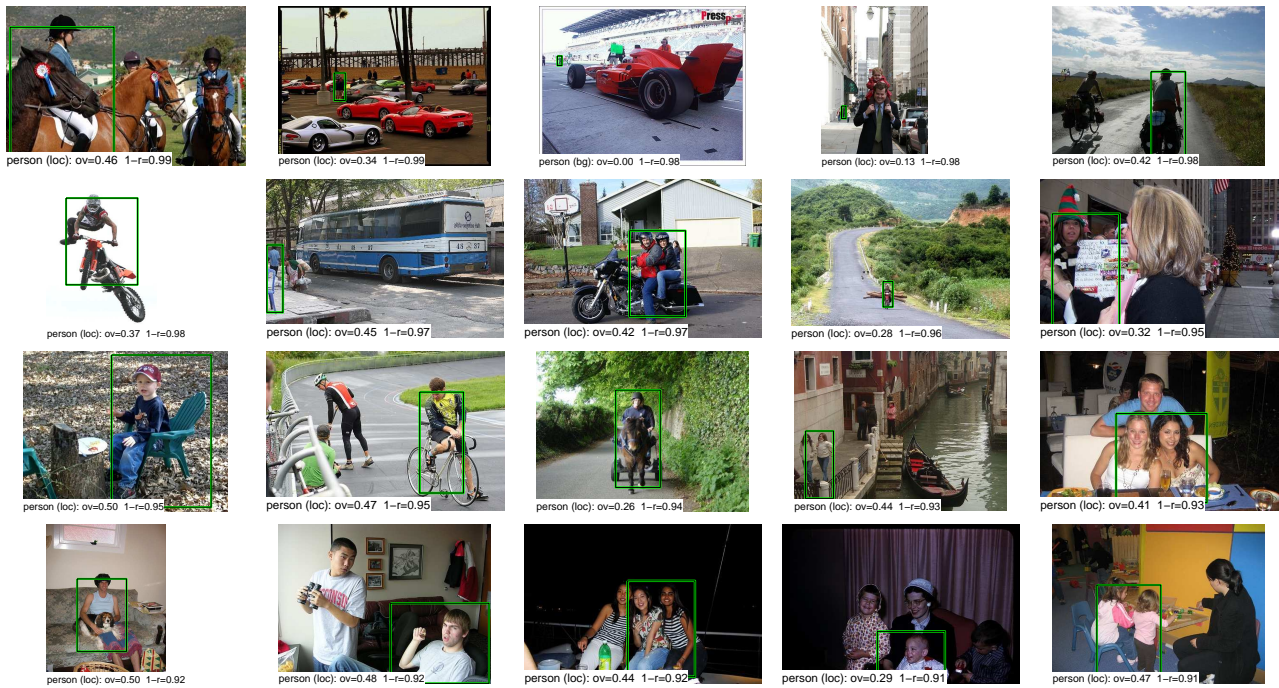


Figure 56: Examples of top person false positives