



Segmentation & Layout Taster Competitions

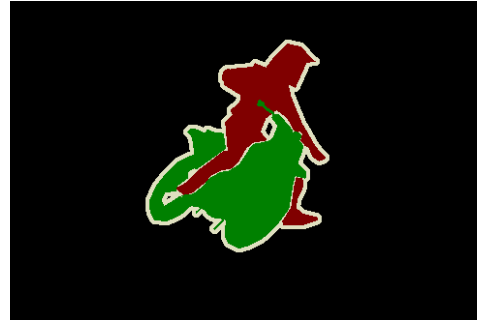
PASCAL Visual Object Classes 2007

Segmentation training data

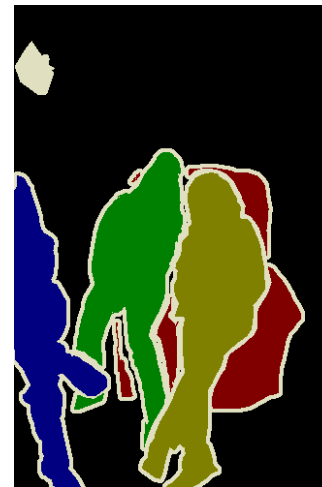
Image



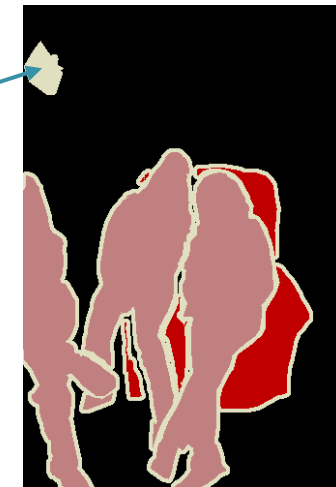
Object segmentation



Class segmentation



Difficult
objects
masked



Data set sizes:	Training	Validation	Total
Images	209	213	422
Objects	633	582	1215

Example segmentations

Image



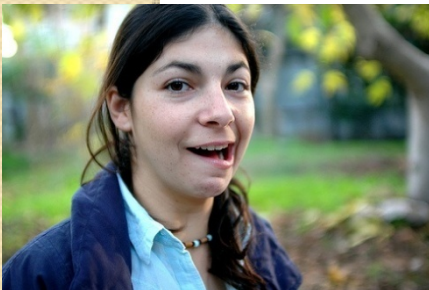
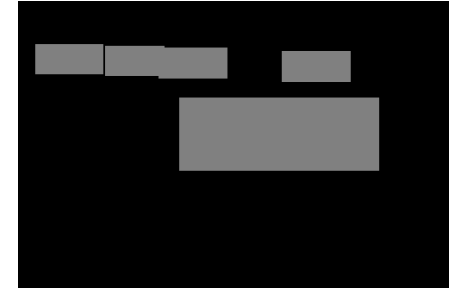
Ground truth



Entry #1



Entry #2
(auto-generated)



Segmentation taster results

Evaluation: average % accuracy across all classes.

	mean	back-ground	aero-plane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	dining table	dog	horse	motor-bike	person	potted plant	sheep	sofa	train	tv/monitor
Brookes	8.5	77.7	5.5	0.0	0.4	0.4	0.0	8.6	5.2	9.6	1.4	1.7	10.6	0.3	5.9	6.1	28.8	2.3	2.3	0.3	10.6	0.7
INRIA Normal*	7.7	2.8	1.2	1.8	8.3	1.5	52.4	0.3	12.4	5.3	3.7	0.0	18.3	4.1	0.1	0.1	3.6	28.7	0.3	6.1	0.4	10.6
INRIA PlusClass*	23.5	2.9	0.6	44.8	34.4	16.4	19.9	0.4	68.0	58.1	10.5	0.4	43.5	7.7	0.9	1.7	59.2	37.2	0.0	5.5	19.0	63.2
MPI Center*	17.5	56.6	11.8	31.2	0.0	11.0	0.0	0.1	35.7	51.4	7.3	20.3	0.0	1.0	2.3	2.6	60.0	0.0	2.6	5.0	43.4	25.3
MPI ESSOL*	27.8	2.6	29.7	30.8	9.5	41.4	6.7	8.0	72.9	55.7	37.1	11.1	19.4	2.2	14.9	23.8	66.8	25.9	8.6	3.2	58.1	55.1
TKK*	30.4	22.9	18.8	20.7	5.2	16.1	3.1	1.2	78.3	1.1	2.5	0.8	23.4	69.4	44.4	42.1	0.0	64.7	30.2	34.6	89.3	70.6
UoCTTI*	21.2	2.5	24.1	52.5	0.4	1.6	16.4	49.4	32.6	1.0	5.5	9.5	0.1	0.2	2.7	20.9	60.2	11.4	0.0	25.8	71.7	57.5

*Automatic entry from detection competition entry

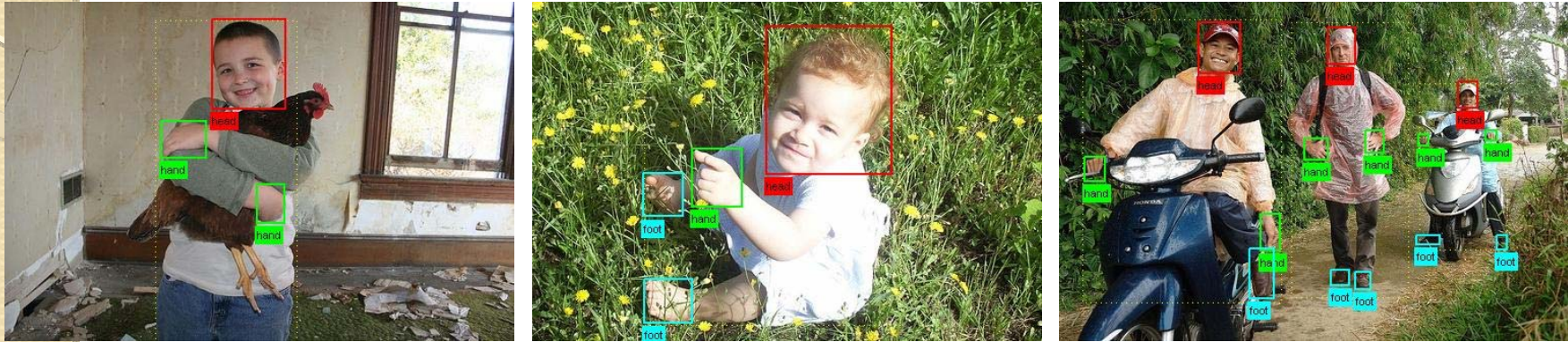


- Winner: Helsinki University of Technology (TKK)
- Honourable mention: Oxford Brookes

Discussion questions

- Evaluation metric
 - Percentage accuracy overall/average per-class
 - Intersection/union measure
$$= \frac{\text{true pos.}}{\text{true pos.} + \text{false pos.} + \text{false neg.}}$$
- Selection of classes
 - Same as for detection challenge (allows entry for both)
 - Fewer classes (more training data per class)
 - Also allow 'stuff' e.g. sky, grass
- Other suggestions?

Person layout training data



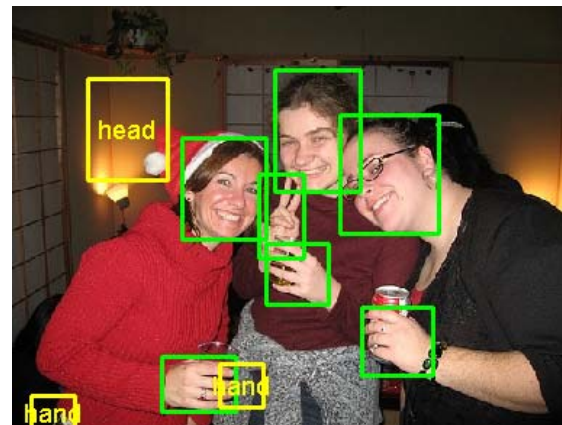
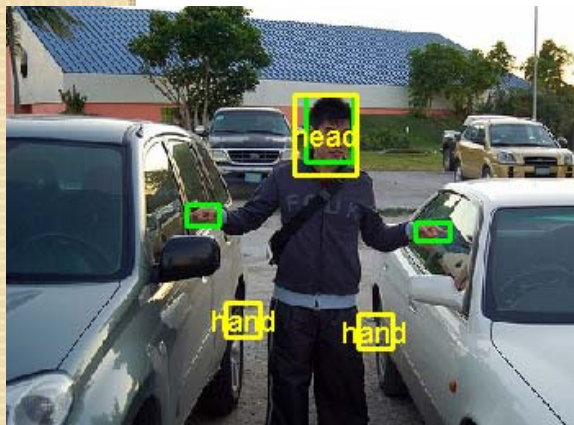
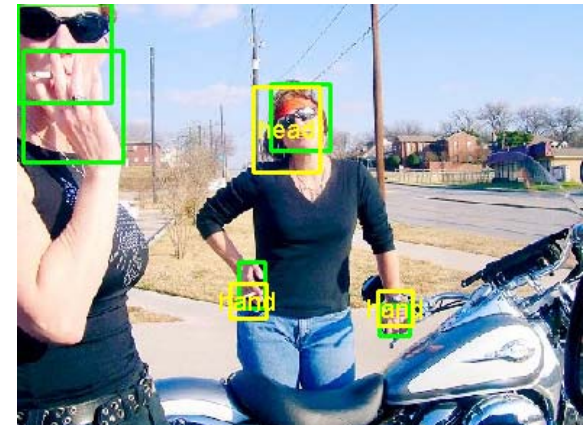
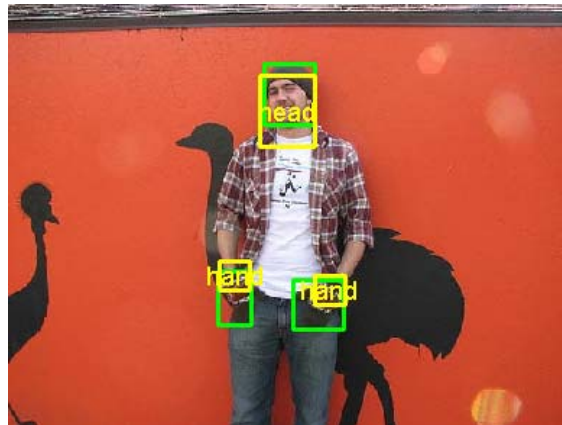
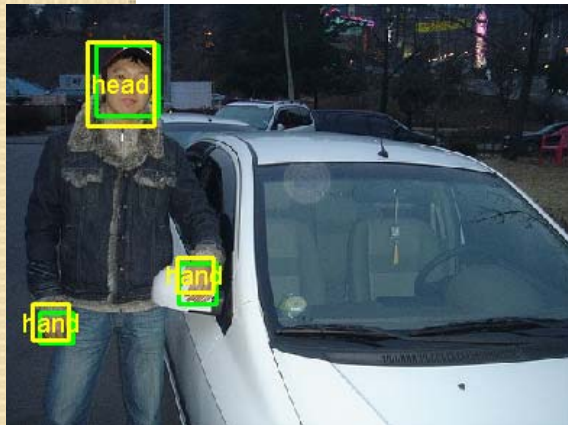
- Subset of “person” objects annotated with bounding boxes of head, hands, feet

Data set sizes:	Training	Validation	Total
Images	166	156	322
Objects	220	219	439

Example results



- One submission:
University of Mannheim





Discussion questions

- Evaluation metric
 - Average precision (AP)
 - Correct detection required correct prediction of all parts' presence/absence and bounding boxes
 - Submitted method had $AP=0$: Is this is a reasonable/sensitive evaluation measure?
- Participation
 - Only one result submitted – is there interest in this task, is data seen as too difficult...?
- Other suggestions?