# The PASCAL Visual Object Classes Challenge 2008 (VOC2008)

## Part 3 – Segmentation Taster

Mark Everingham
Luc Van Gool
Chris Williams
John Winn
Andrew Zisserman

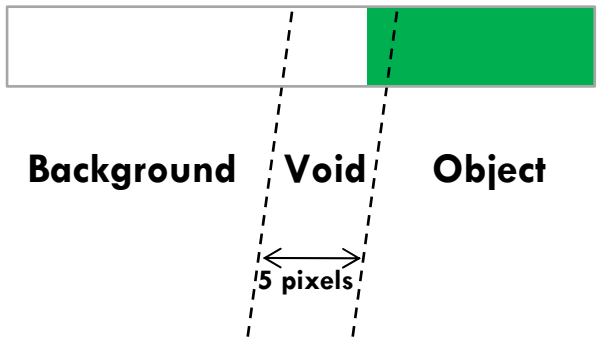
# Segmentation Taster

- For each pixel, predict the class of the object containing that pixel or 'background'.

- Competition 5: Train on the supplied data
  - Which methods perform best given specified training data?
- [Competition 6: Train on any (non-test) data]
  - Not an official competition!

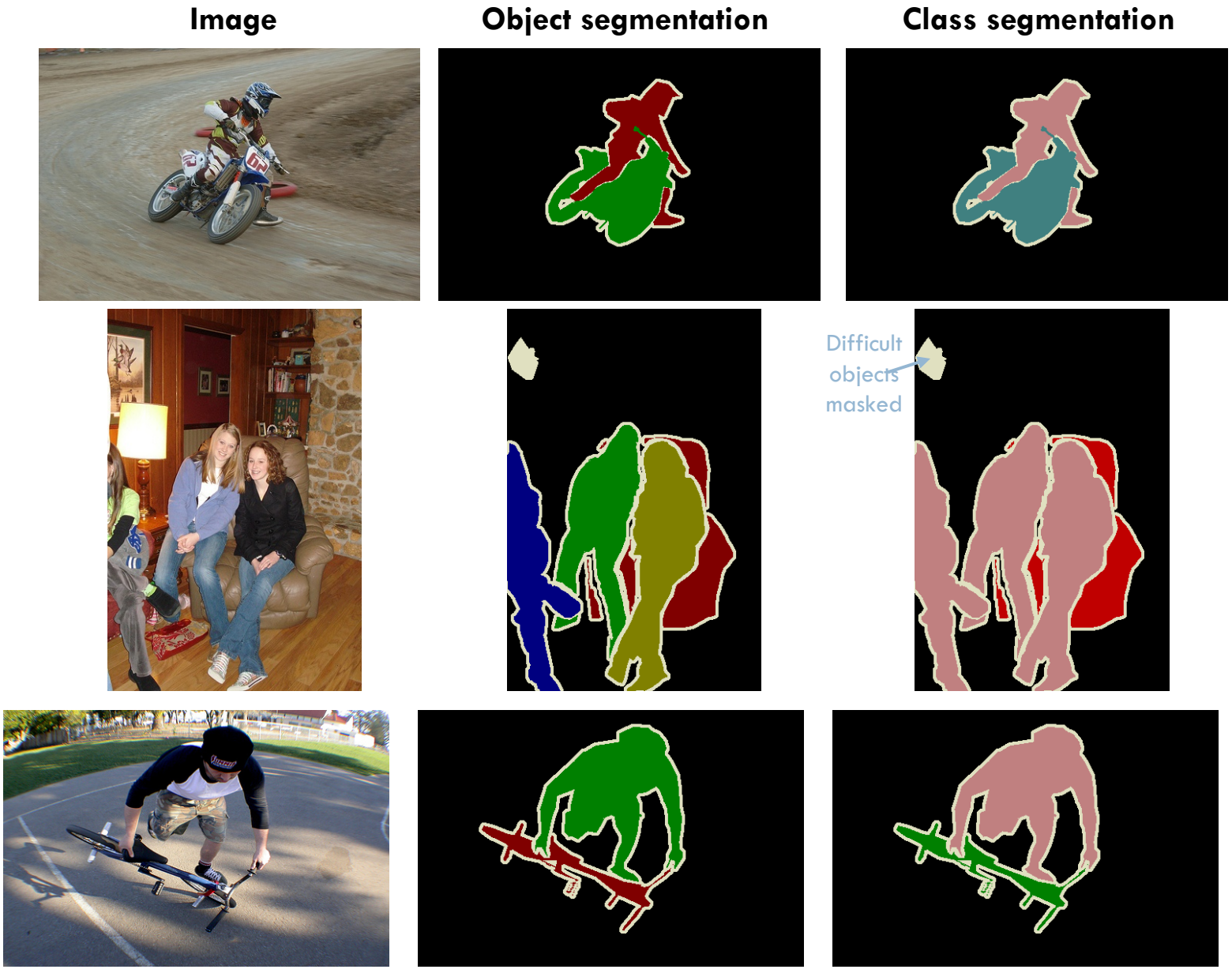
# Annotation

- Annotation in one session with written guidelines
  - Segmentation is ‘refinement’ of bounding box (but may go outside it)
  - Segmentation accurate to within 5-pixel boundary region which is marked ‘void’

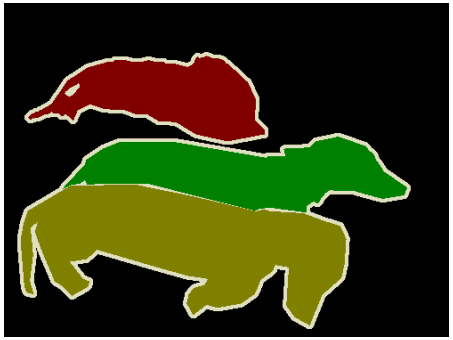
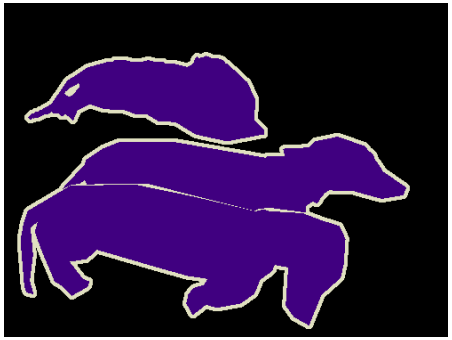
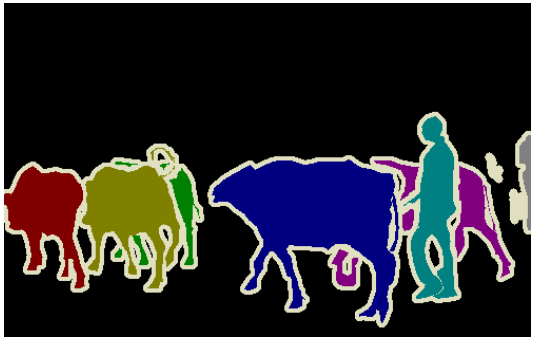
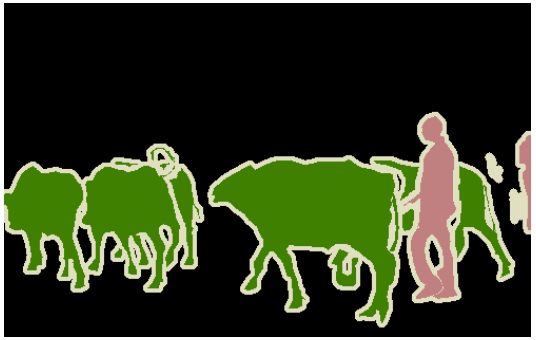



  - 1-pixel wide structures (whiskers, wires) can be ignored
  - Surface objects considered part of the object (e.g. items on a table)

# Example annotations



# Example annotations

| Image | Object segmentation | Class segmentation |
| --- | --- | --- |
| | | |
| |  |  |
| |  |  |

# Training/validation data sets

- 2008 data training and validation sets include and extend 2007 data set
- More than double the number of annotated images:

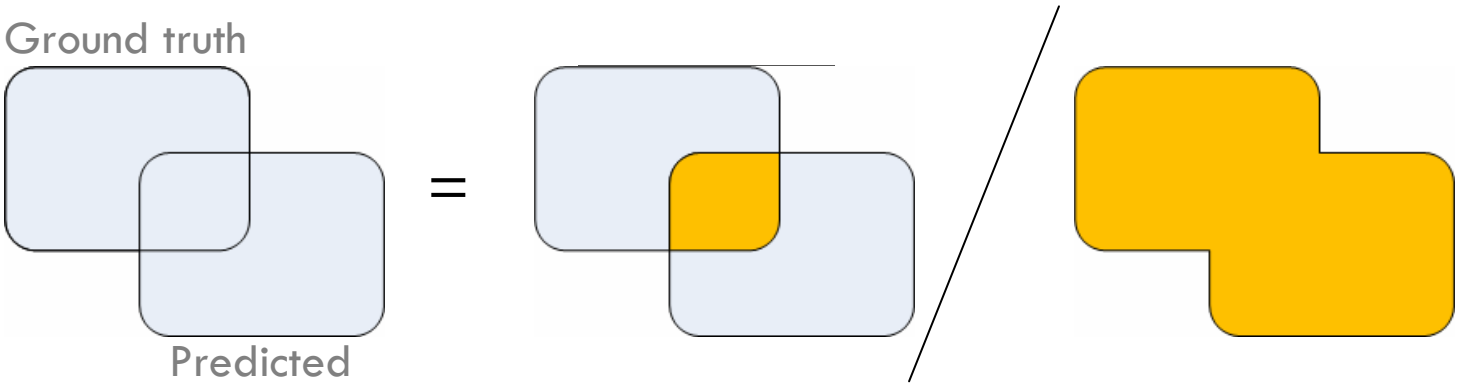
| Number of: | Training | Validation | Total |
|---|---|---|---|
| **Images** | 511 (209) | 512 (213) | **1023** (422) |
| **Objects** | 1166 (633) | 1203 (582) | **2369** (1215) |

VOC 2007 totals shown in brackets

# New evaluation metric for VOC 2008

Intersection/union of **class** labels

$$= \frac{\text{true pos. class}}{\text{true pos.} + \text{false pos.} + \text{false neg.}}$$




- Metric chosen because:
  - Allows per-class participation
  - Penalises both over- and under-estimates
- Overall evaluation metric is average over all classes (including background)

# Methods

- 6 direct and 5 ‘automatic’ entries using a variety of methods
- Features:
  - SIFT, RGB, Textons, randomized forests
- Segmentation method:
  - Bottom-up oversegmentation (superpixels)
  - MRF approaches (including high order cliques)
  - Refinement of detections (also provides instance labels)
- Use of image-level classifiers

# Results

| | 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/ MSRC | 20.1 | 75.0 | **36.9** | 4.8 | **22.2** | 11.2 | 13.7 | 13.8 | 20.4 | 10.0 | **8.7** | 3.6 | **28.3** | 6.6 | 17.1 | 22.6 | **30.6** | 13.5 | 26.8 | 12.1 | 20.1 | 24.8 |
| *CASIA_def** | 13.5 | 44.7 | 8.5 | 9.4 | 10.7 | 7.5 | 13.3 | **37.3** | 10.3 | 14.4 | 2.0 | 5.4 | 4.0 | 8.0 | 9.0 | 15.9 | 10.5 | 8.2 | 16.8 | 4.1 | 21.6 | 22.1 |
| Jena | 8.0 | 47.8 | 7.2 | 3.1 | 4.6 | 5.6 | 2.2 | 0.6 | 13.4 | 0.0 | 0.7 | 7.5 | 0.7 | 5.7 | 4.4 | 8.9 | 8.7 | 5.0 | 9.2 | 3.4 | 12.2 | 17.8 |
| *LEAR PlusClass** | 3.7 | 5.5 | 4.1 | 2.2 | 0.0 | 3.5 | 4.9 | 3.9 | 9.6 | 0.8 | 1.5 | 0.1 | 0.4 | 0.9 | 1.2 | 2.5 | 7.4 | 0.2 | 0.2 | 0.3 | 4.0 | 24.7 |
| MPI_norank | 7.0 | 66.3 | 6.7 | 1.2 | 2.1 | 3.1 | 2.5 | 5.8 | 2.6 | 2.9 | 1.1 | 1.7 | 4.0 | 2.6 | 3.7 | 5.1 | 10.5 | 0.8 | 5.8 | 2.1 | 8.4 | 8.0 |
| MPI_single | 12.9 | 75.4 | 19.1 | 7.7 | 6.1 | 9.4 | 3.8 | 11.0 | 12.1 | 5.6 | 0.7 | 3.7 | 15.9 | 3.6 | 12.2 | 16.1 | 15.9 | 0.6 | 19.7 | 5.9 | 14.7 | 12.5 |
| *MPI_struct** | 12.9 | 59.2 | 15.1 | 5.5 | 8.2 | 19.1 | 8.6 | 13.0 | 11.0 | 11.9 | 6.8 | 0.0 | 3.8 | 11.0 | 14.3 | 14.9 | 12.1 | 0.1 | 7.2 | 5.1 | 19.6 | 23.5 |
| *UoCTTIUCI** | 11.6 | 0.9 | 9.1 | 12.8 | 1.0 | 6.7 | 11.8 | 24.7 | 11.4 | 8.9 | 1.6 | 9.0 | 1.3 | 7.9 | 14.4 | 18.8 | 13.1 | 9.5 | 24.3 | 11.0 | 17.1 | 27.8 |
| *XRCE_Def** | 18.9 | 64.8 | 11.1 | 12.7 | 15.4 | 10.9 | 7.8 | 17.7 | 18.4 | 19.4 | 5.6 | 14.0 | 8.5 | 16.8 | 16.0 | 30.9 | 25.3 | 19.1 | 29.7 | 5.5 | 18.0 | 29.0 |
| XRCE_Seg | **25.4** | **75.9** | 25.8 | **15.7** | 19.2 | **21.6** | **17.2** | 27.3 | **25.5** | **24.2** | 7.9 | **25.4** | 9.9 | **17.8** | **23.3** | **34.0** | 28.8 | **23.2** | **32.1** | **14.9** | **25.9** | **37.3** |

*Automatic entry from detection competition entry

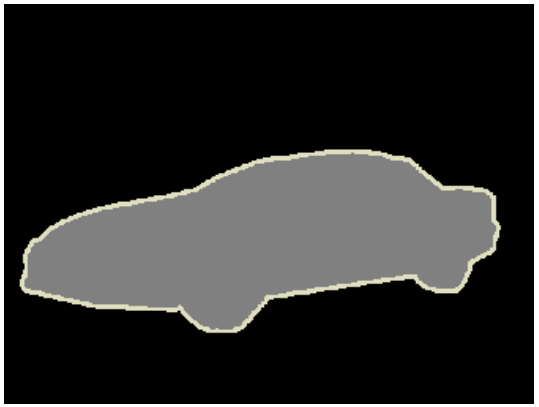
Trained on external data:

| | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| UIUC_CMU | 19.5 | **79.3** | 31.9 | **21.0** | 8.3 | 6.5 | **34.3** | 15.8 | 22.7 | 10.4 | 1.2 | 6.8 | 8.0 | 10.2 | 22.7 | 24.9 | 27.7 | 15.9 | 4.3 | 5.5 | 19.0 | 32.1 |

# Example segmentations

Image

GroundTruth



Brookes MSRC

MPI single

UIUC CMU

XRCE

# Example segmentations

Image

GroundTruth

Brookes MSRC

MPI single

UIUC CMU

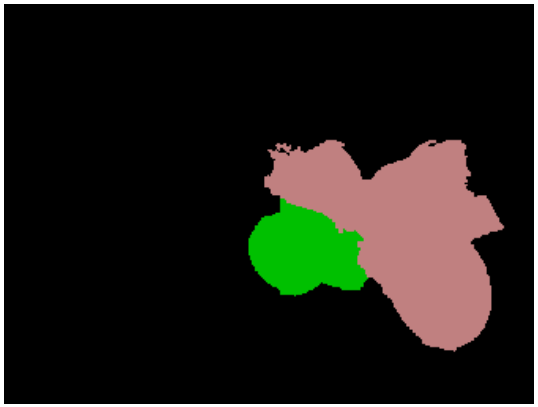
XRCE

# Example segmentations

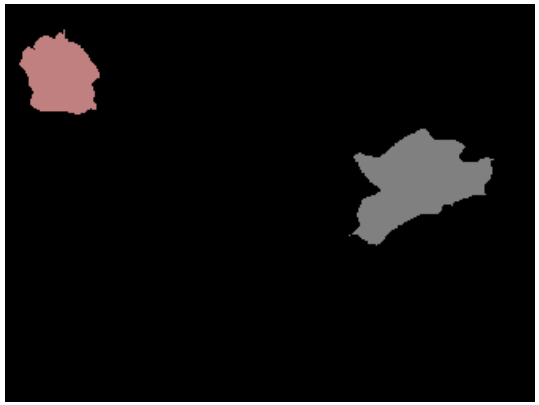
Image

GroundTruth

Brookes MSRC



MPI single



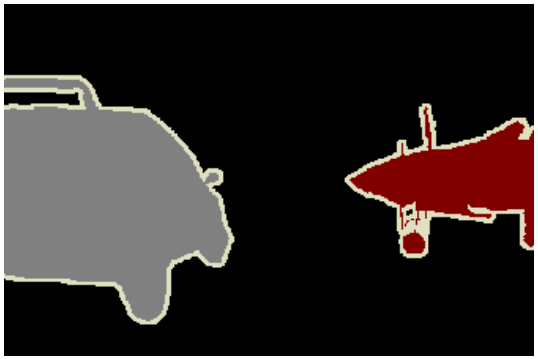
UIUC CMU

XRCE

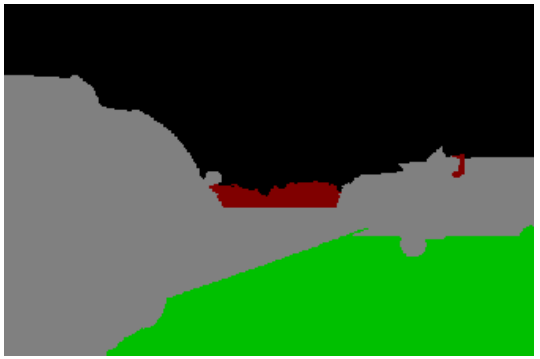
# Example segmentations

Image

GroundTruth



Brookes MSRC



MPI single

UIUC CMU

XRCE

# Example segmentations

Image

GroundTruth

Brookes MSRC

MPI single

UIUC CMU

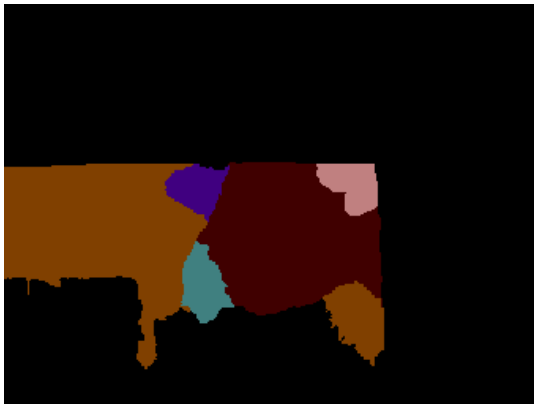
XRCE

# Example segmentations

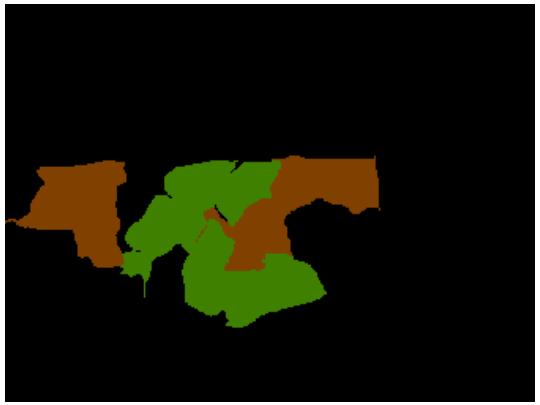
Image

GroundTruth

Brookes MSRC



MPI single



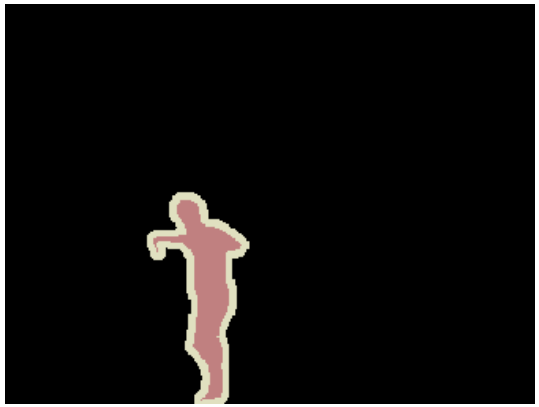
UIUC CMU

XRCE

# Example segmentations

Image

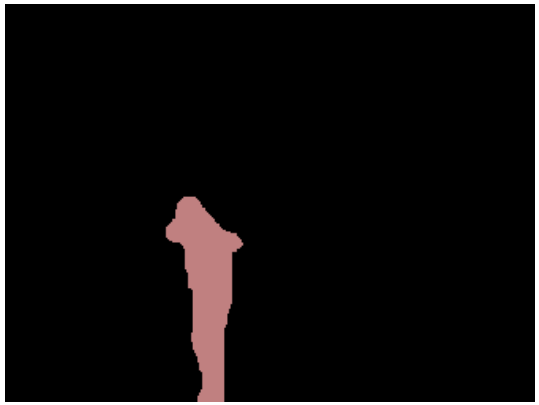
GroundTruth



Brookes MSRC

MPI single

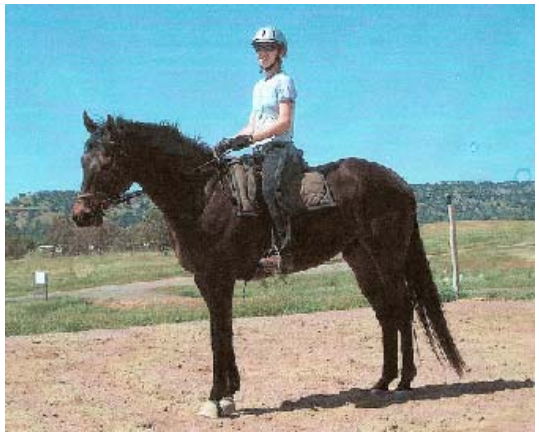
UIUC CMU



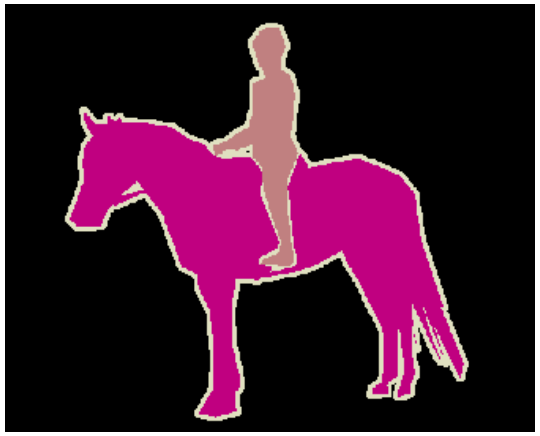
XRCE

# Example segmentations

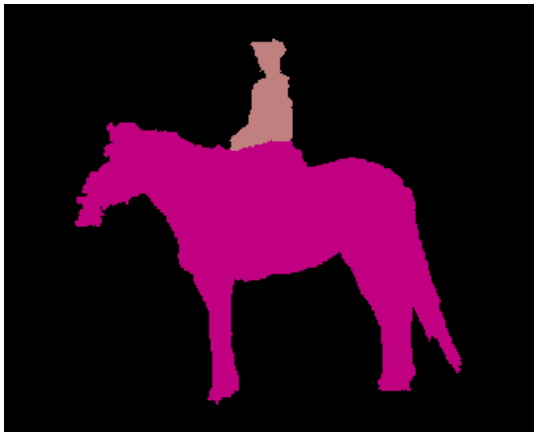
Image



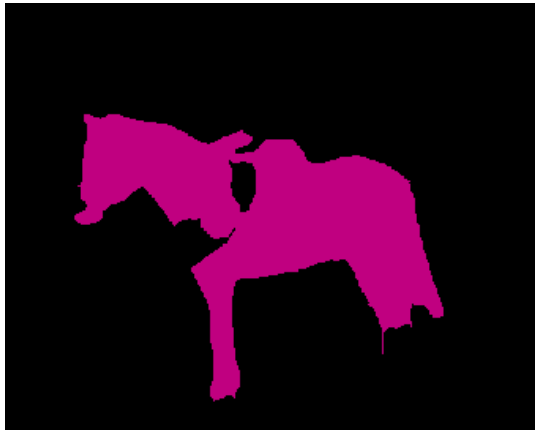
GroundTruth



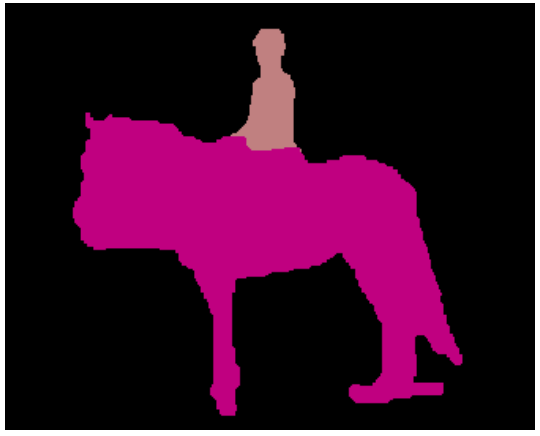
Brookes MSRC



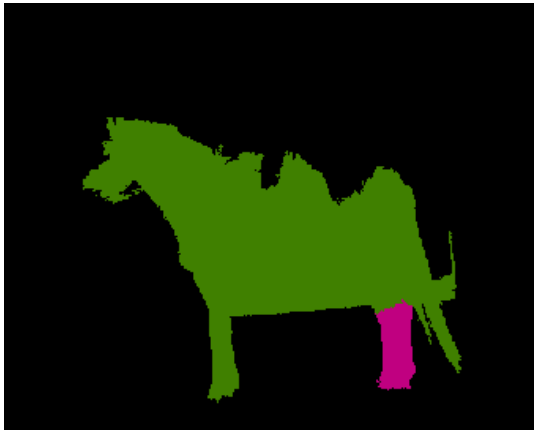
MPI single



UIUC CMU



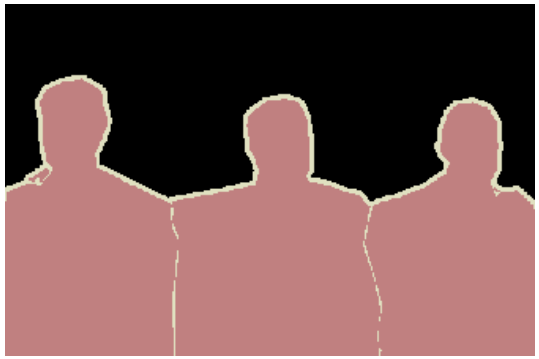
XRCE



# Example segmentations

Image

GroundTruth



Brookes MSRC

MPI single

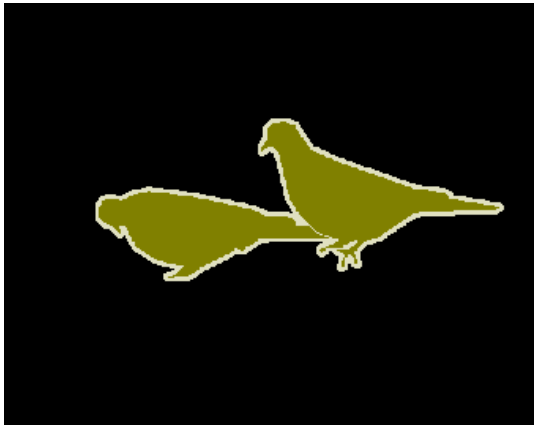
UIUC CMU

XRCE

# Example segmentations

Image

GroundTruth



Brookes MSRC

MPI single

UIUC CMU

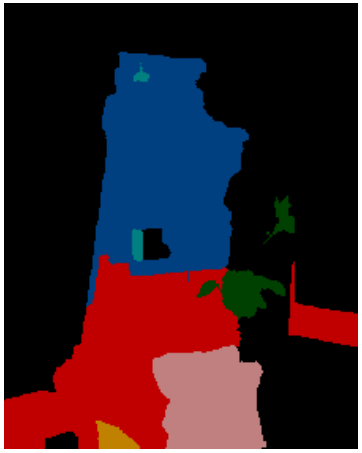
XRCE

# Example segmentations

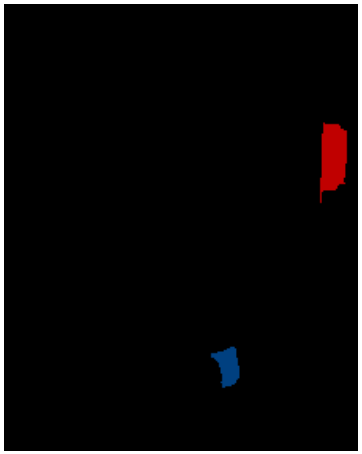
Image

GroundTruth

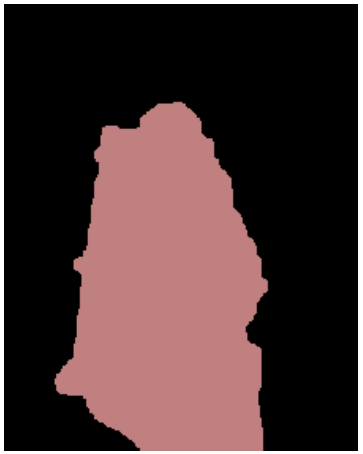
Brookes MSRC



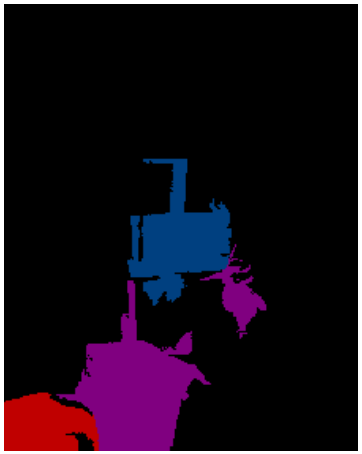
MPI single



UIUC CMU



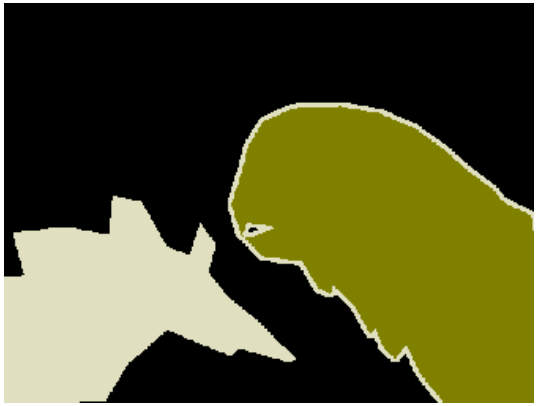
XRCE



# Example segmentations

Image

GroundTruth



Brookes MSRC

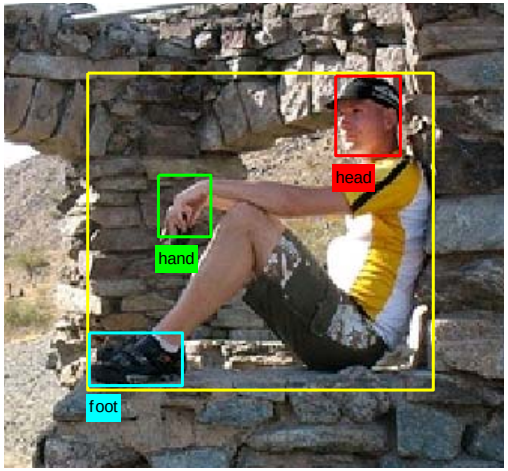
MPI single

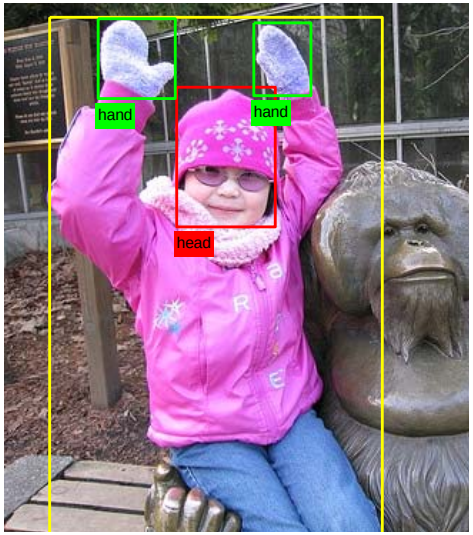
UIUC CMU

XRCE

# Layout Taster

- Given the bounding box of a person, predict the visibility and positions of head, hands and feet.
  - ~300 training examples
  - Please give it a try!





# Prizes

- **Winner: XRCE**

  Gabriela Czurka, Florent Perronnin, Yan Liu

  *Xerox Research Centre Europe (XRCE), Textual and Visual Pattern Analysis Group*

- **Runner up: Brookes/MSRC**

  Lubor Ladicky[1], Phil Torr[1], Pushmeet Kohli[2]

  *[1]Oxford Brookes University, [2]Microsoft Research Cambridge*

- **Honorable Mention: UIUC_CMU**

  Derek Hoiem[1], Santosh Divvala[2], James H. Hays[2]

  *[1]University of Illionois Urbana-Champaign; [2]Carnegie Mellon University*