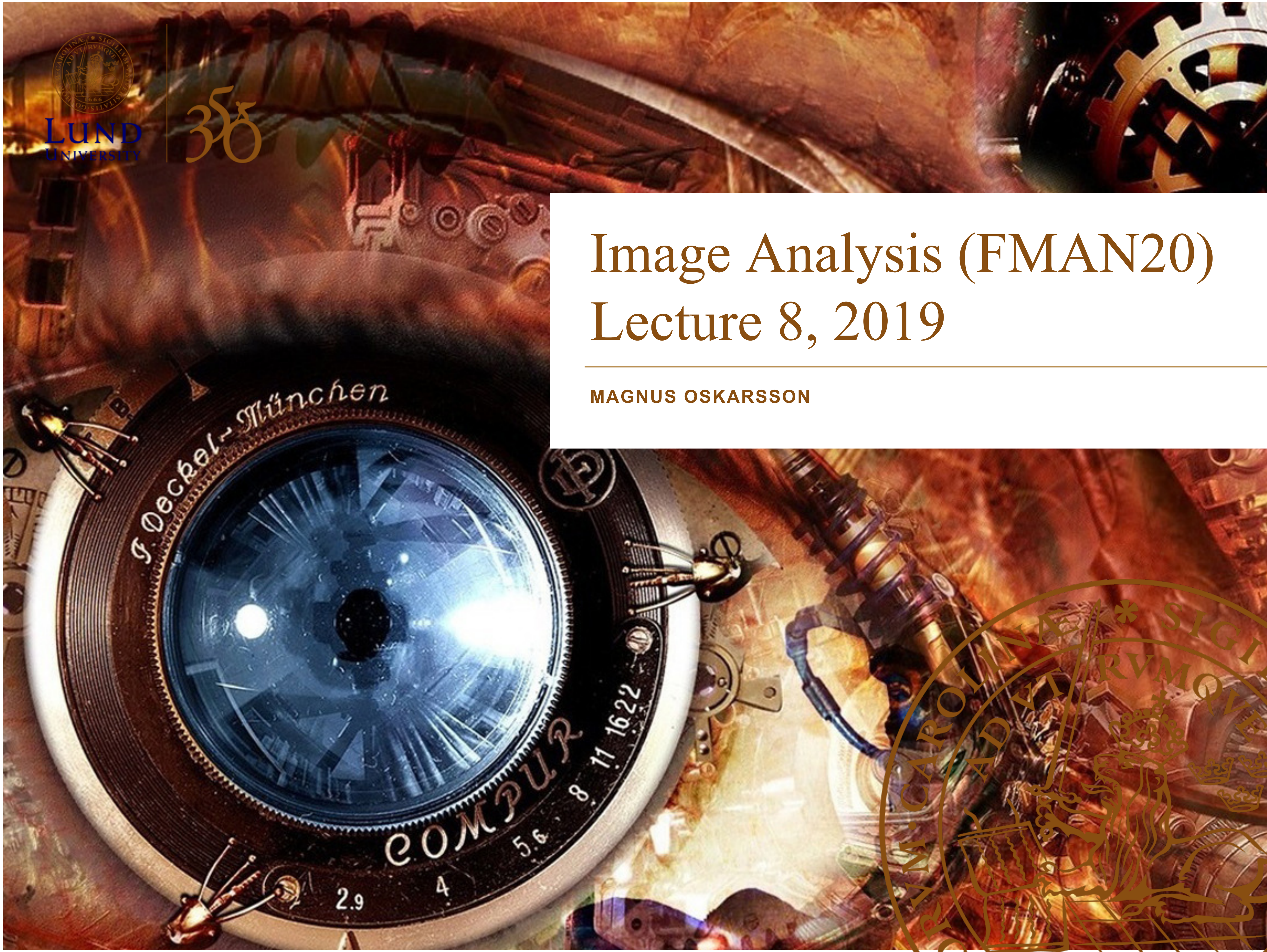


LUND UNIVERSITY

Image Analysis (FMAN20)
Lecture 8, 2019

MAGNUS OSKARSSON

Selective Attention Test

from Simons & Chabris (1999)

Overview – Systems & Segmentation

- Recap and outlook
- Computer and Segmentation. Does segmentation matter?
- System
  - Build
  - Test
- Segmentation principles
- Segmentation using tools from lectures 1-7
- Mathematical Morphology

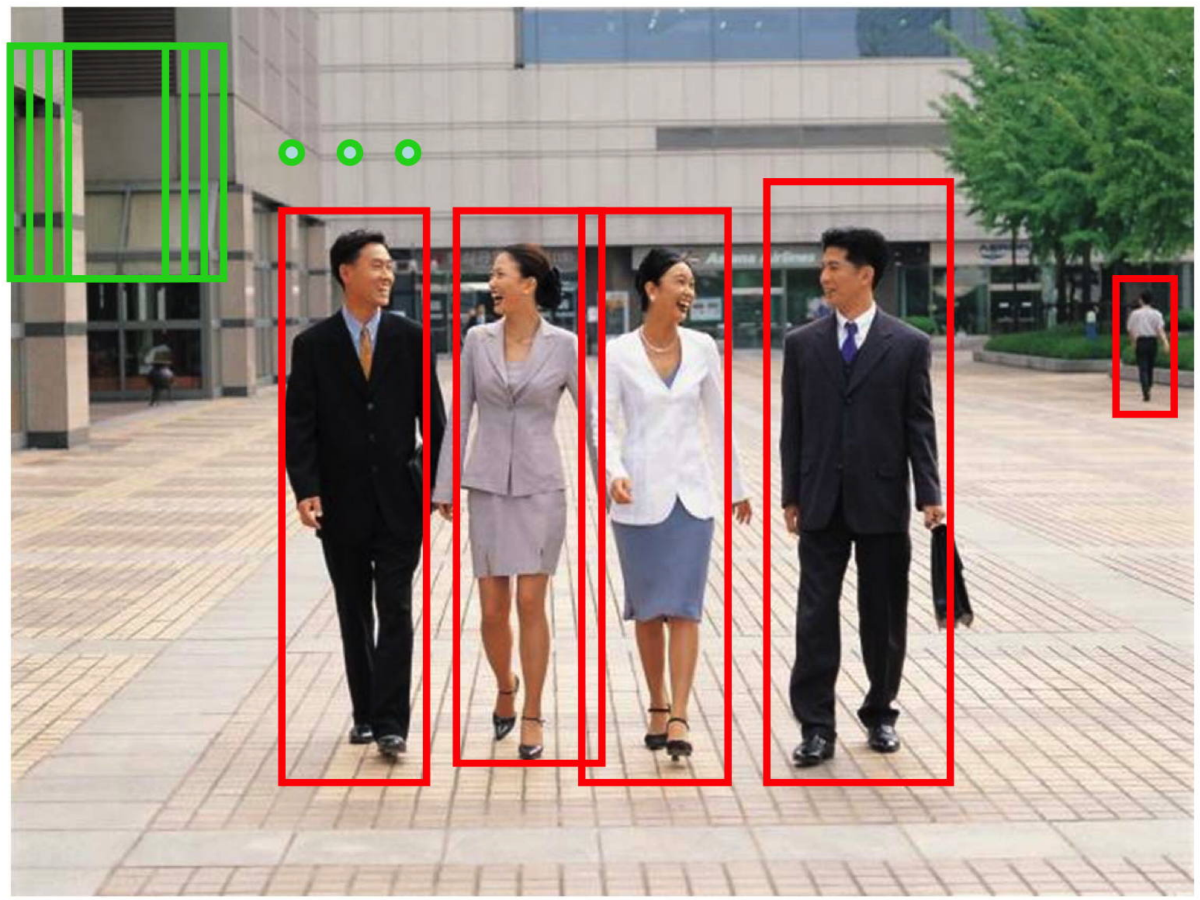
Overview –
Systems & Segmentation

- **Computer and Segmentation. Does segmentation matter?**
- System
  - Build
  - Test
- Segmentation principles
- Segmentation using tools from lectures 1-7
- Mathematical Morphology

Computer Vision Tasks

- Classifying images
- Estimating the spatial layout of structurally distinct scene elements
  - Segmentation
- Identifying geometric structure
  - 3d surface or volume of objects
  - Pose of people or other biological forms
- Recognizing objects and actions

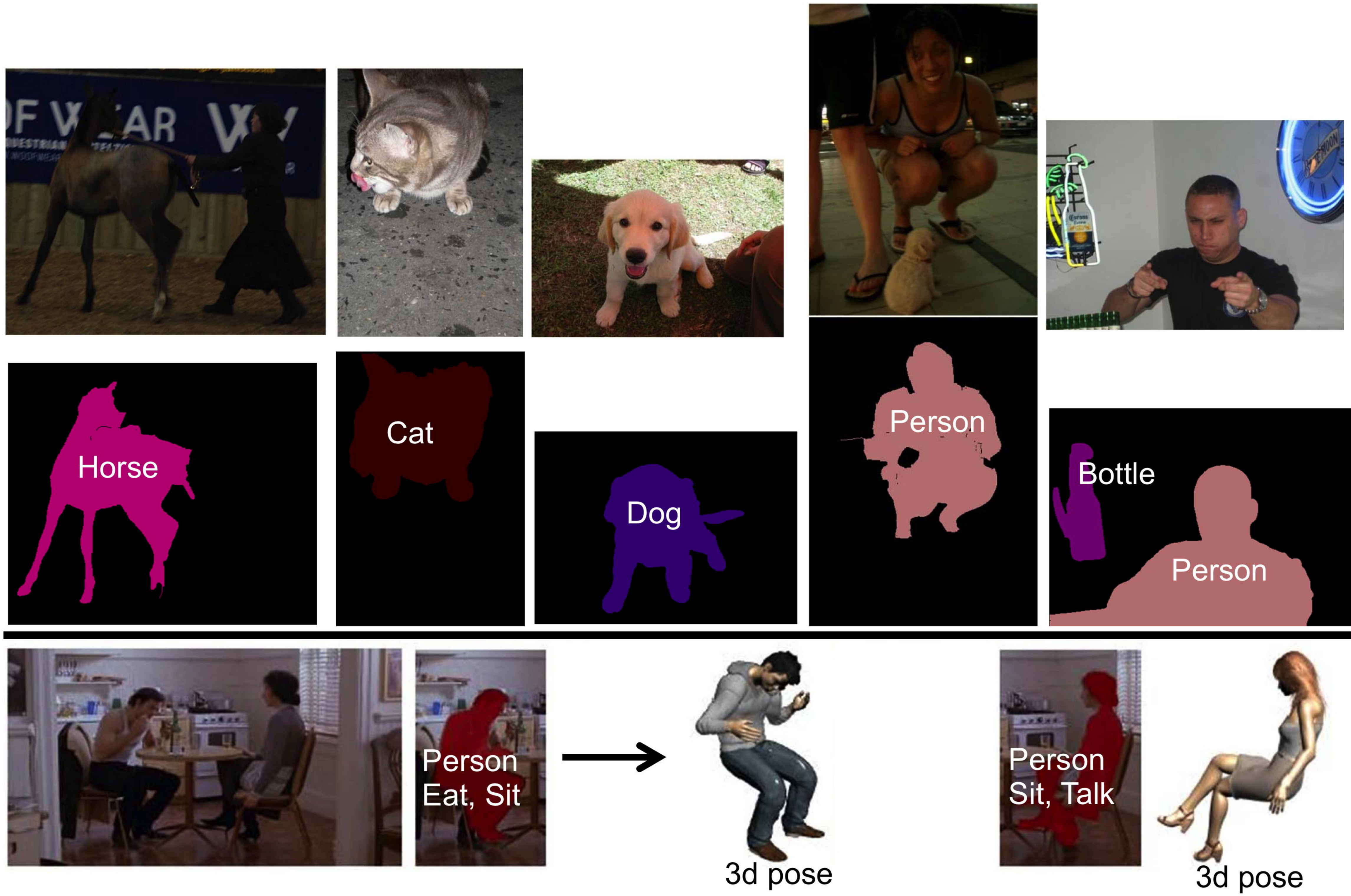
Recognition by Detection



Search at multiple locations, scales and for all object categories of interest

Rowley, Baluja & Kanade 1996 (face detection)

Recognizing objects, poses, actions



Dynamic Scenes, 3D Reasoning

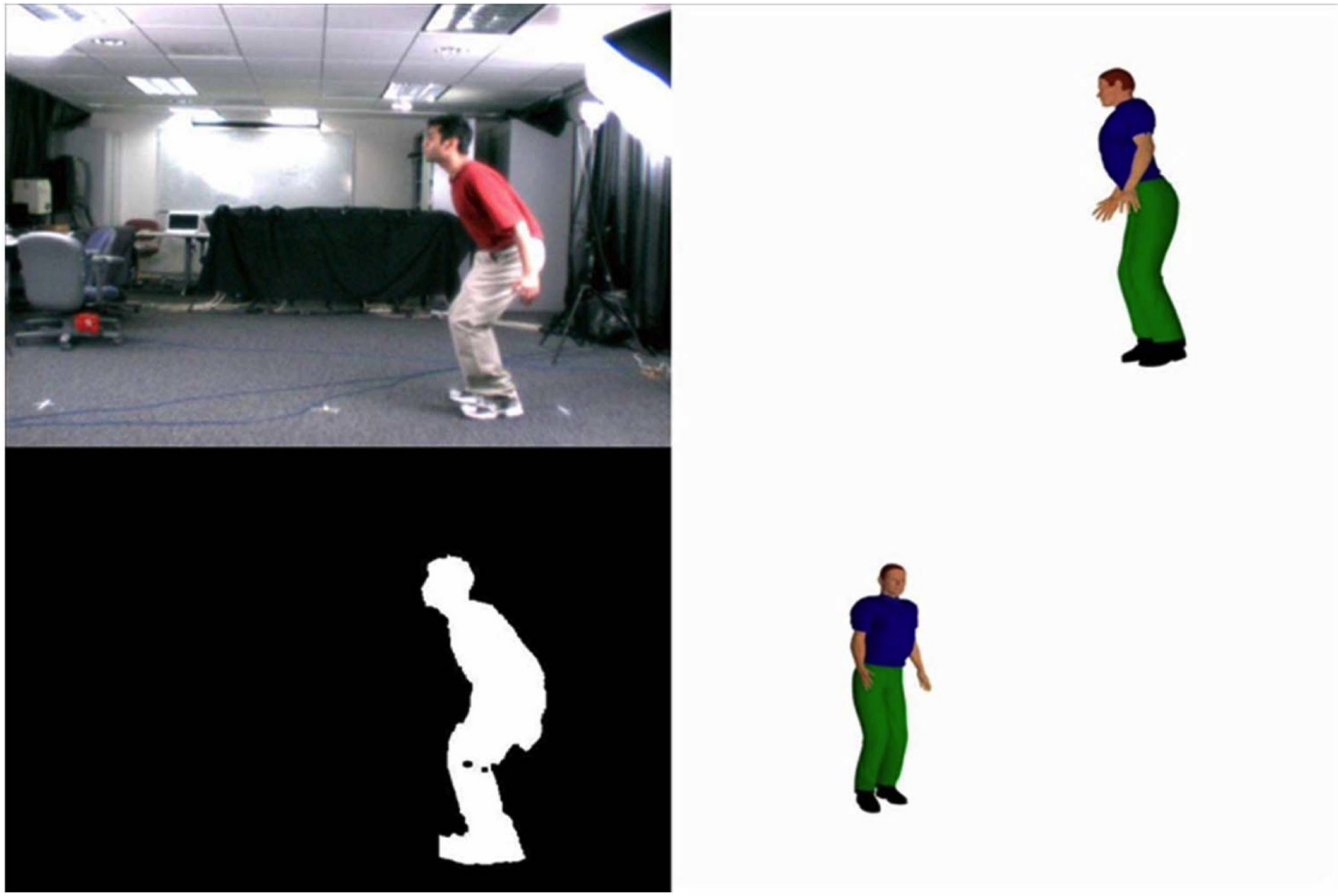
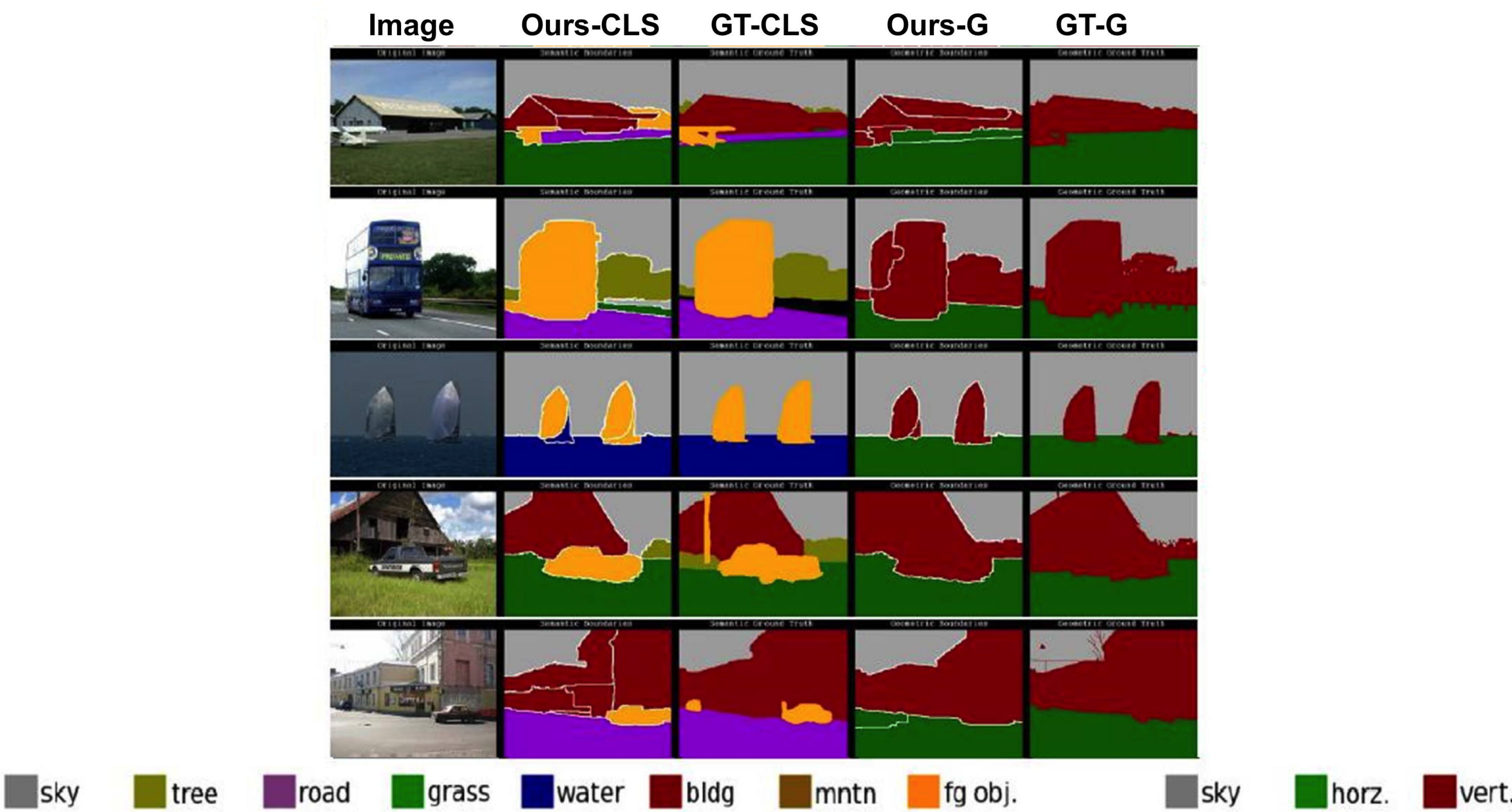


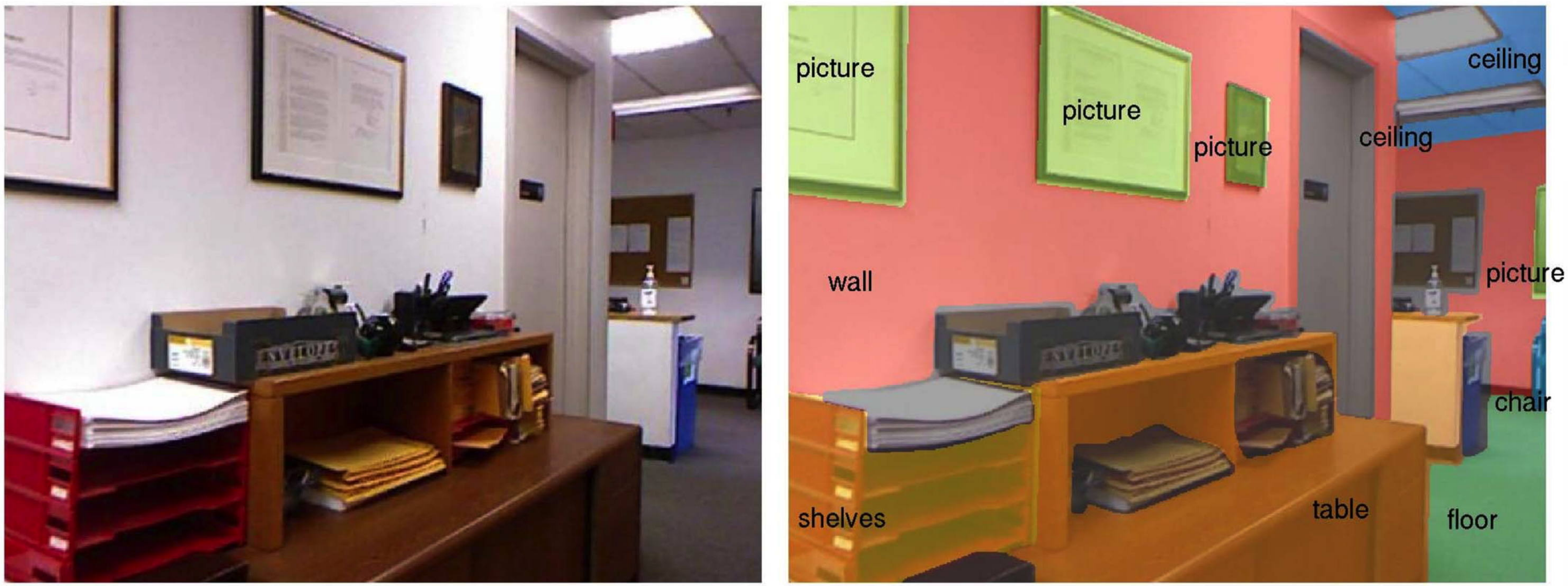
Figure-ground segmentation essential

Sminchisescu , Kanaujia, Metaxas, PAMI 2007

Scene Understanding



Scene Understanding



Ideally, we would want a framework that `uniformly' accommodates color, depth and video analysis

Problems

- **Region generation**
  - Systematic, figure-ground hypotheses, combinatorial
  - Boundaries from RGB, depth, motion
- **Region selection, hypothesis set compression**
  - Object-like = Class-independent = Objectness
  - Maximum marginal diversification
- **Region description**
  - Second-order methods
- **Complete scene recognition by composition**
  - Re-combination, re-segmentation of figure-ground
  - Sequential vs. simultaneous

Segmentation

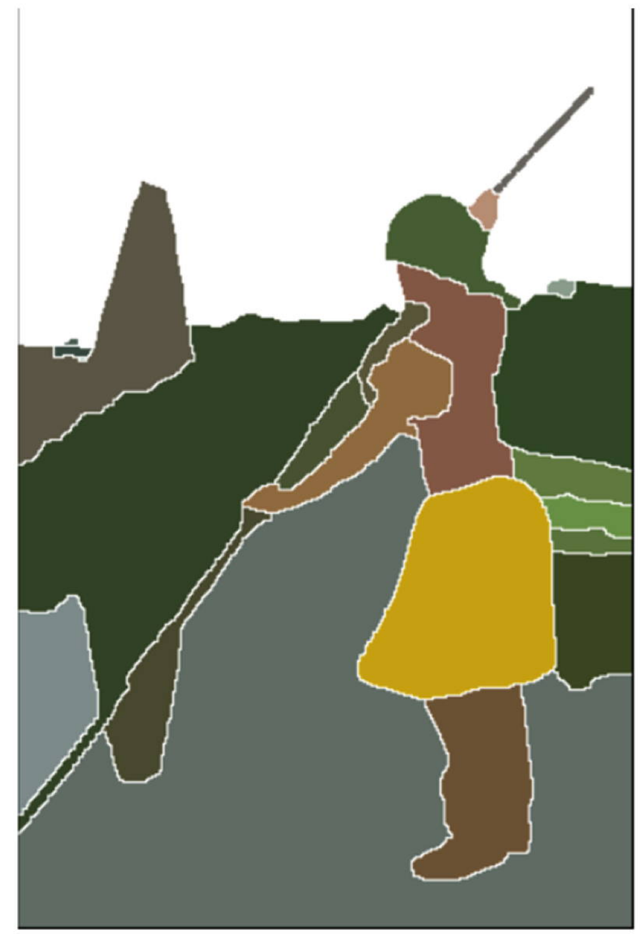
- **Image segmentation**: breaking the pixels or tokens of an image into regions (groups) that share some property
- **Semantic segmentation**: attach category labels to groups

Mid-level Image Segmentation

- Google Scholar returns over 1.000.000 hits for search terms “Image segmentation”.
- 50 years after the first segmentation algorithm.
  - “Experimental evaluation of techniques for automatic segmentation of objects in a complex scene”, J. Muerle and D. Allen,1968.
- Modern well know techniques still aim to segment homogeneous regions, not objects:
  - Normalized Cuts
  - Mean Shift
  - Hierarchical clustering

Mid-level Image Segmentation

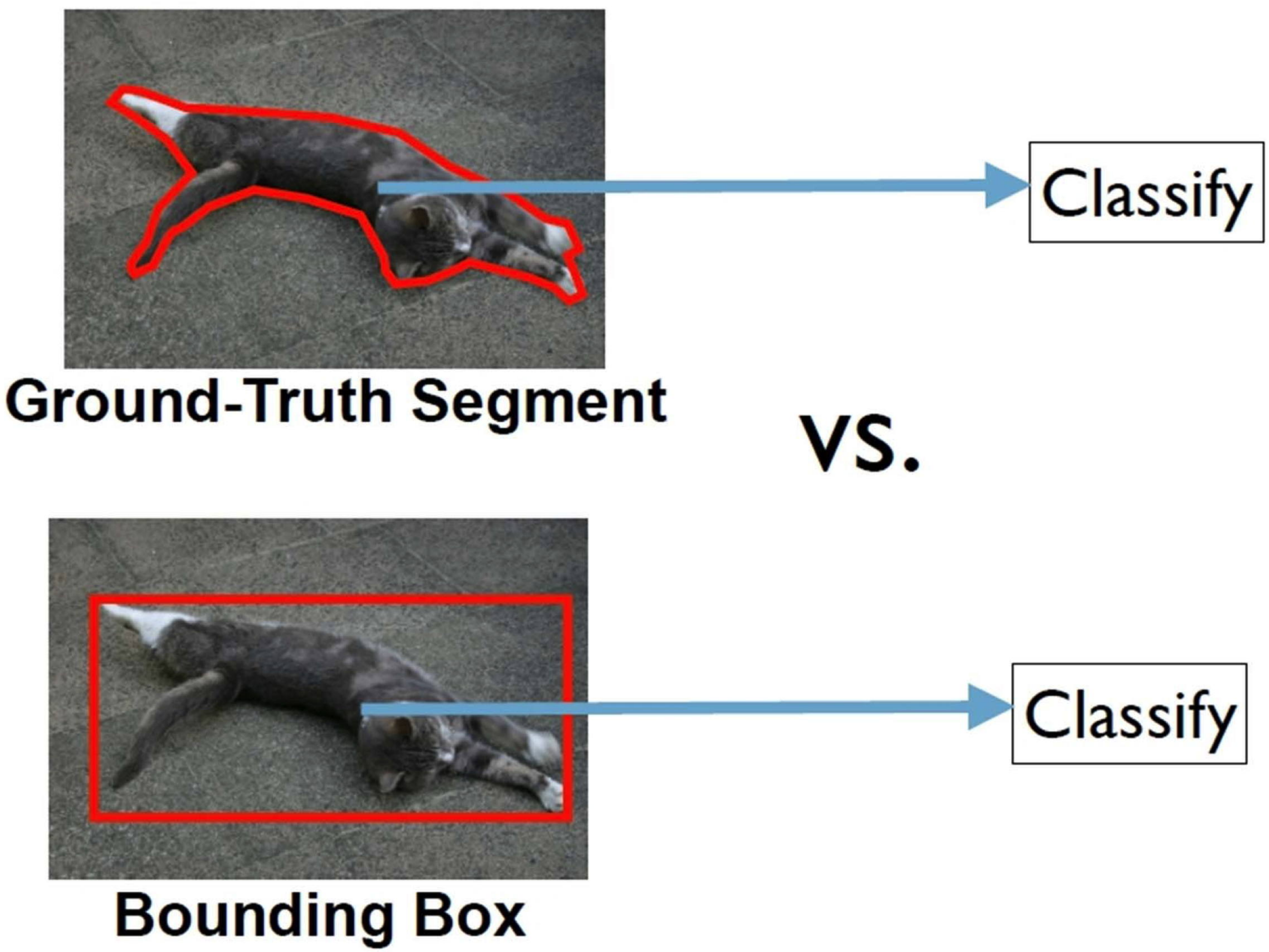
Philosophy: split the image into homogeneous regions



Problems with the multi-scale window recognition-by-detection approach

- Computational overhead
  - 10^4 categories **x** 10^6 windows **x** 10 scales
- Segmentation not considered
  - Improper handling of irregular shapes
  - Window descriptor dominated by background
- Context not considered
  - No criterion for global consistency

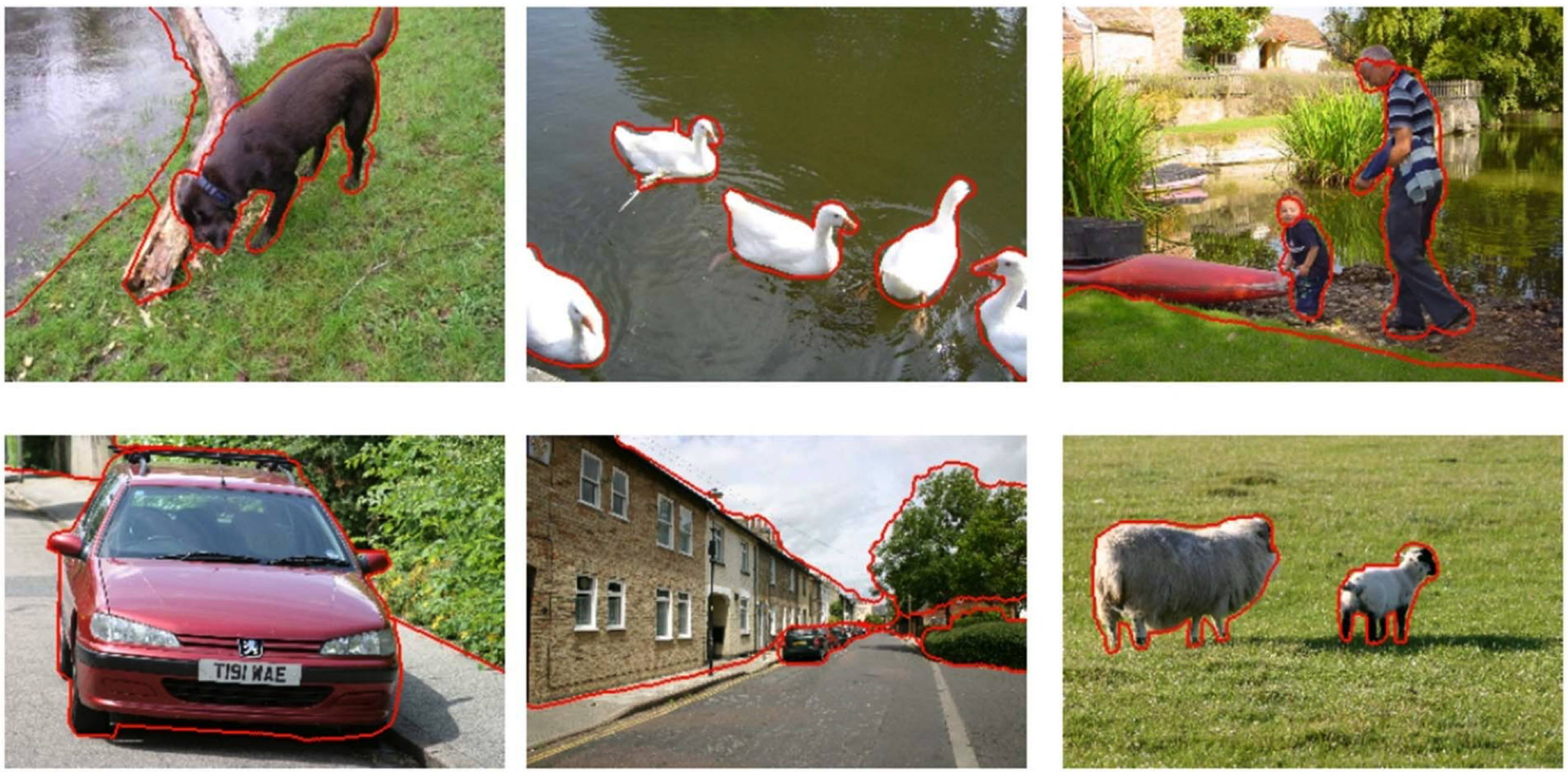
Does spatial support matter?



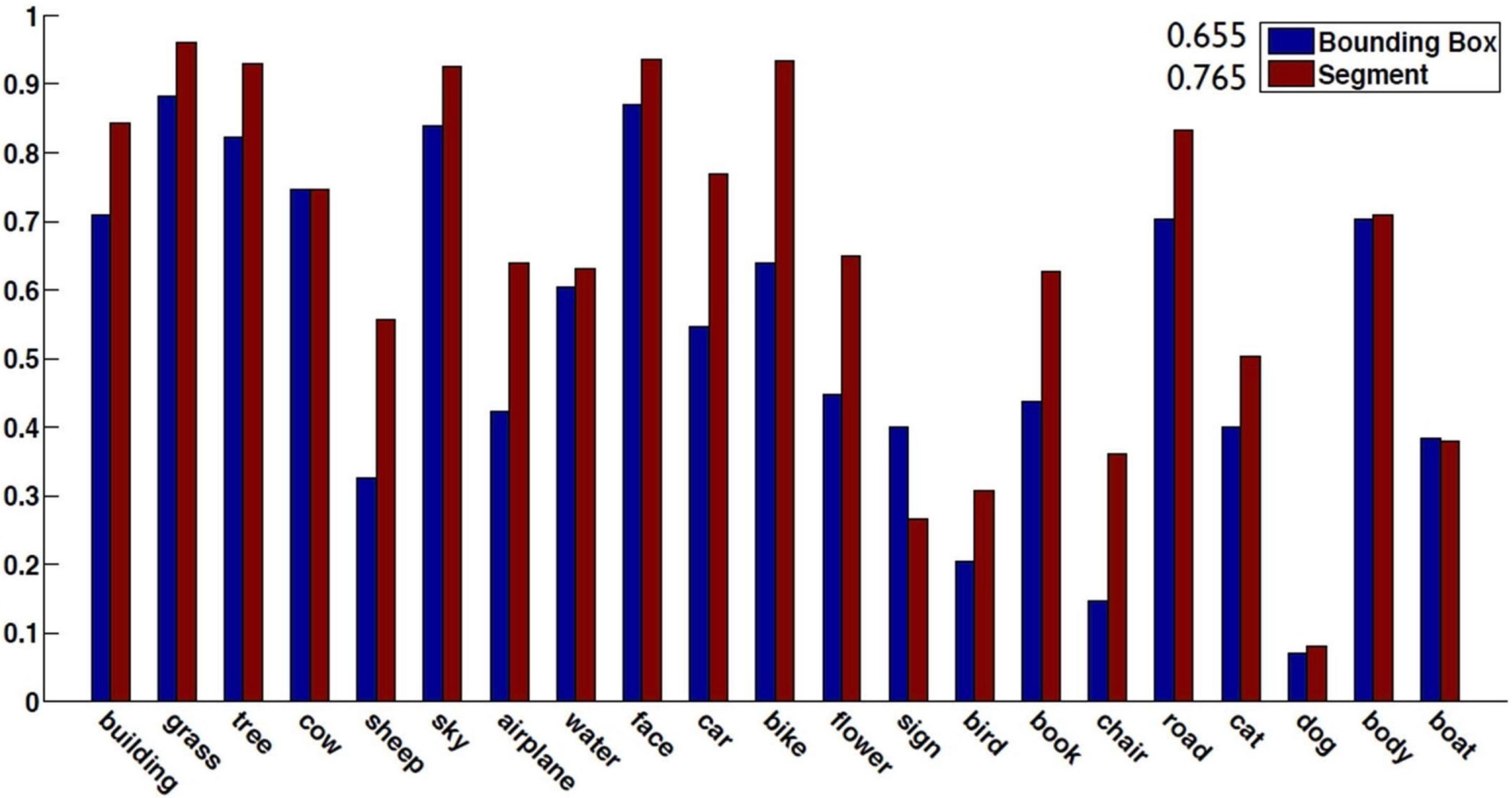
Malisiewicz & Efros, BMVC07

MSRC segmentation dataset

MSRC data-set: 591 images of 23 object classes +
pixel-wise segmentation masks



Spatial support matters!



Malisiewicz & Efros, BMVC07

Segmentation

- Important
- Many methods
- Many systems are built along the lines
- Image -> Segments -> Features -> Result
- In this lecture:
  - Discussion on system building and testing
  - More on a few segmentation methods

Overview – Systems & Segmentation

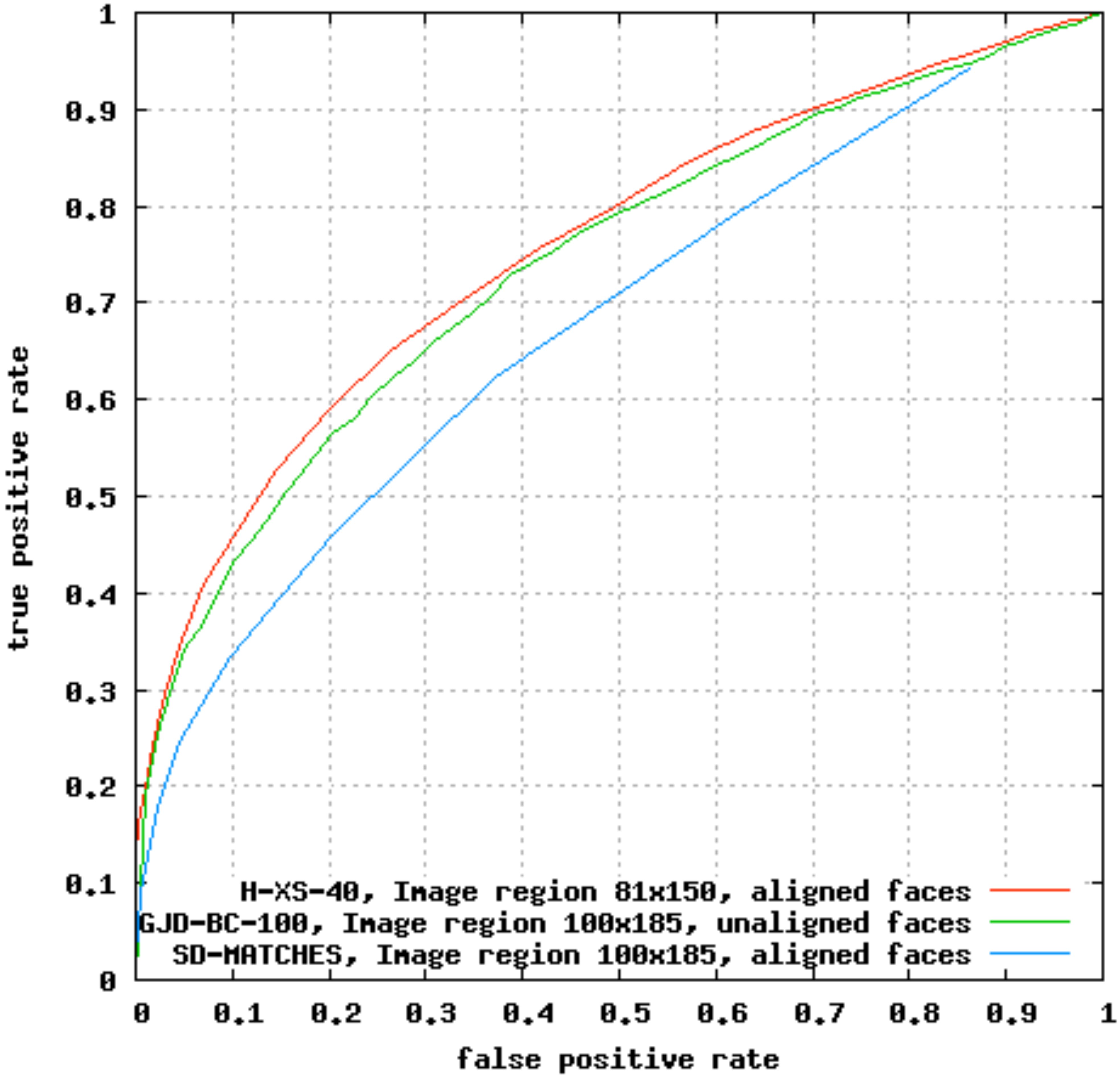
- Recap and outlook
- Computer and Segmentation. Does segmentation matter?
- **System**
  - Build
  - Test
- Segmentation principles
- Segmentation using tools from lectures 1-7
- Mathematical Morphology

Testing your system

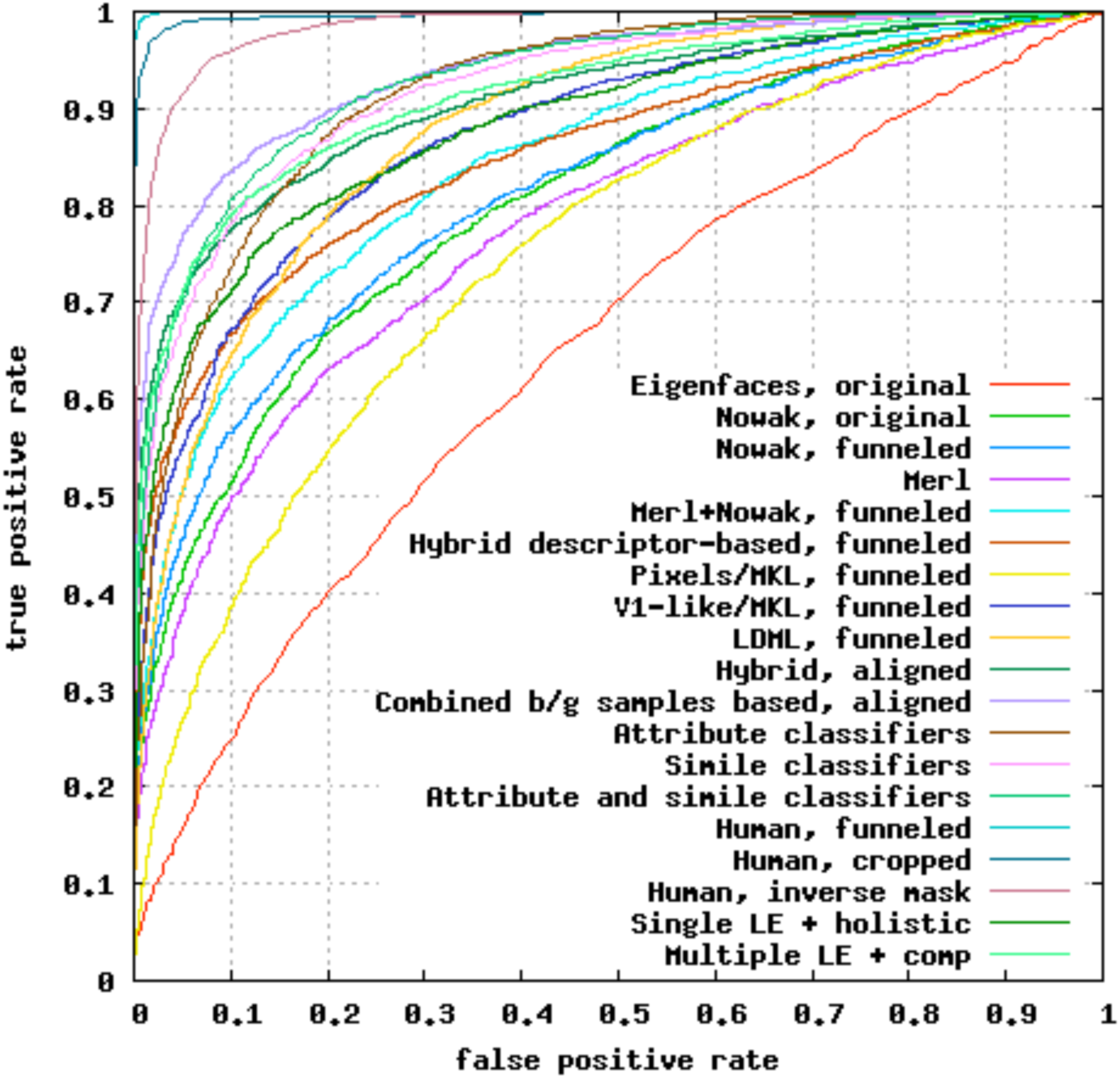
- Image analysis systems
  - Often complex and varying data
  - Often a system of systems
- Important to test your system
- Questions
  - Obtain **data**
  - Obtain **’ground truth**’ (’Gold Standard’)
  - Construct **benchmark scripts**
  - **Visualize** the results
- Adress these questions **early** in a project

Testing your system, Example 1 Labelled Faces in the wild

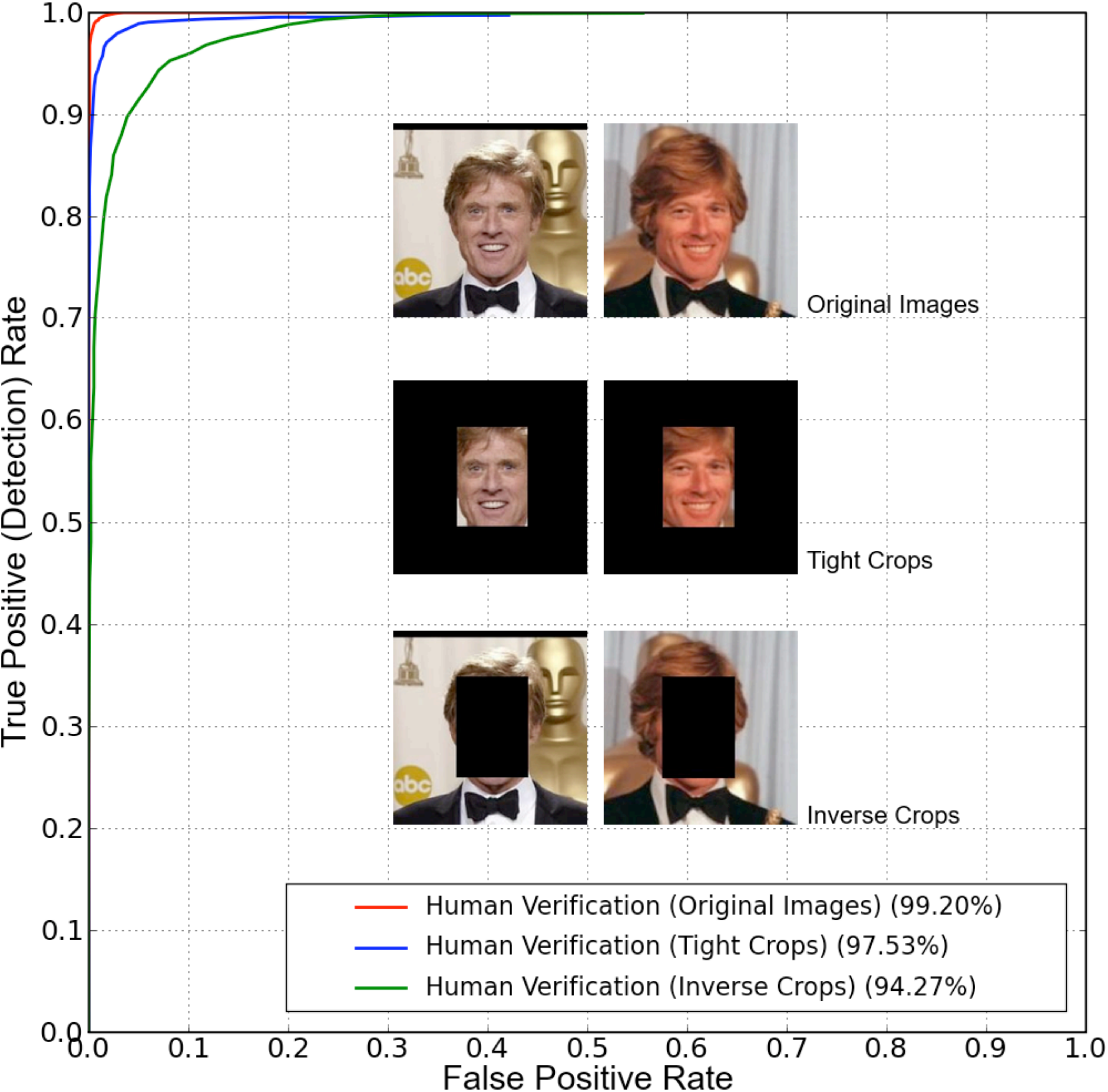
- Collection of images from the web
- Face detection
- Pairs of matching faces
- Pairs of non-matching faces



true positive rate
false positive rate
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
H-XS-40, Image region 81x150, aligned faces
GJD-BC-100, Image region 100x185, unaligned faces
SD-MATCHES, Image region 100x185, aligned faces



true positive rate
false positive rate
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Eigenfaces, original
Nowak, original
Nowak, funneled
Merl
Merl+Nowak, funneled
Hybrid descriptor-based, funneled
Pixels/MKL, funneled
V1-like/MKL, funneled
LDML, funneled
Hybrid, aligned
Combined b/g samples based, aligned
Attribute classifiers
Simile classifiers
Attribute and simile classifiers
Human, funneled
Human, cropped
Human, inverse mask
Single LE + holistic
Multiple LE + comp

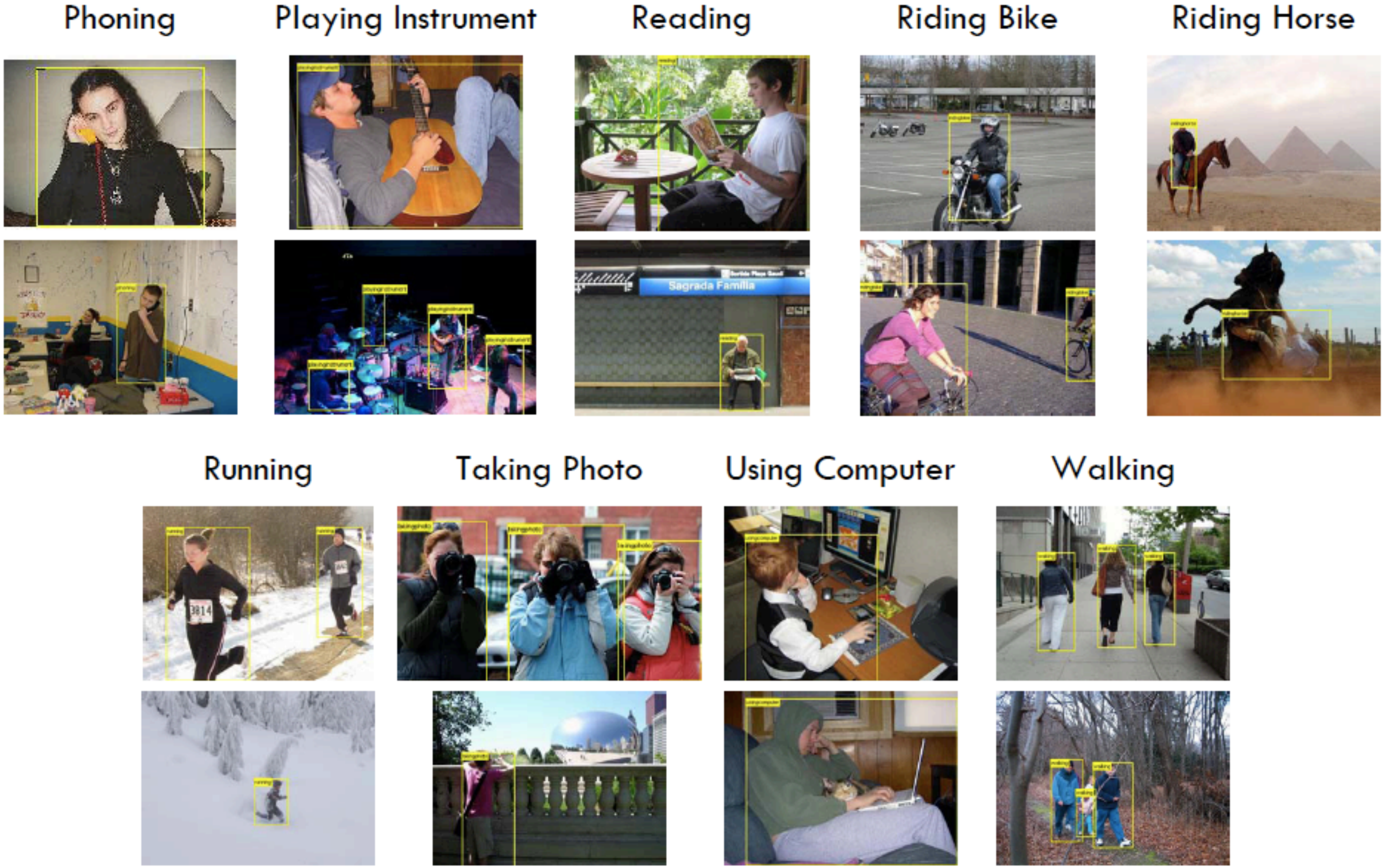


True Positive (Detection) Rate
False Positive Rate
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Original Images
Tight Crops
Inverse Crops
Human Verification (Original Images) (99.20%)
Human Verification (Tight Crops) (97.53%)
Human Verification (Inverse Crops) (94.27%)

Evaluating segmentation

- The Jaccard score
- A – pixels of system segmentation
- B – pixels of ground truth segmentation

$$J = \frac{|A \cap B|}{|A \cup B|}$$

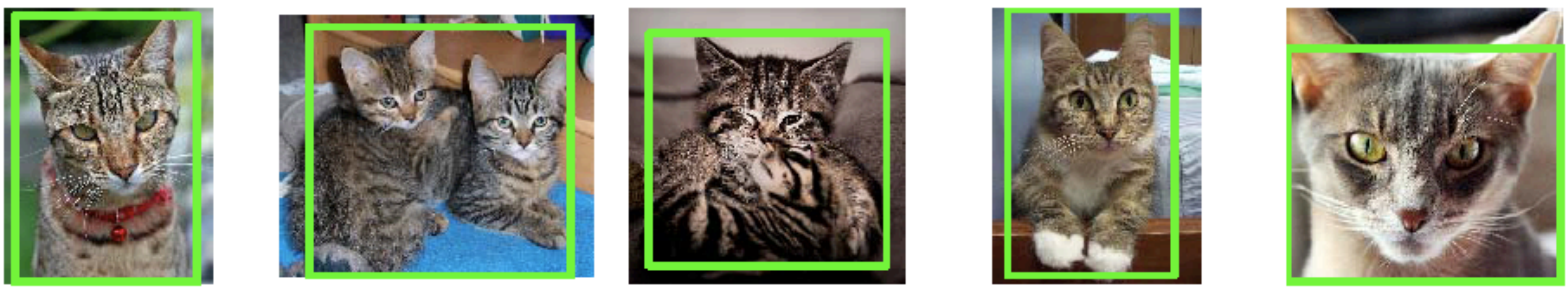


Phoning
Playing Instrument
Reading
Riding Bike
Riding Horse
Sagrada Familia
Running
Taking Photo
Using Computer
Walking

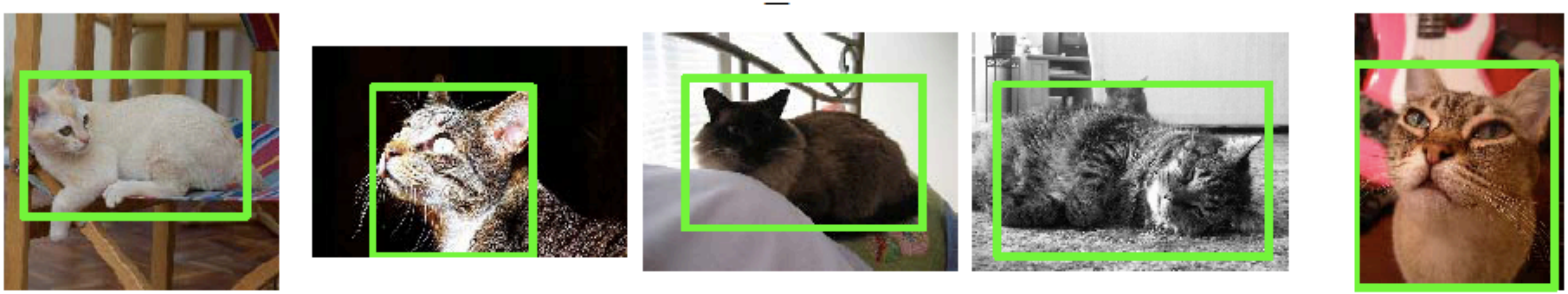
True Positives - Cat

UVA_DETMONKEY

UVA_GROUPLOC

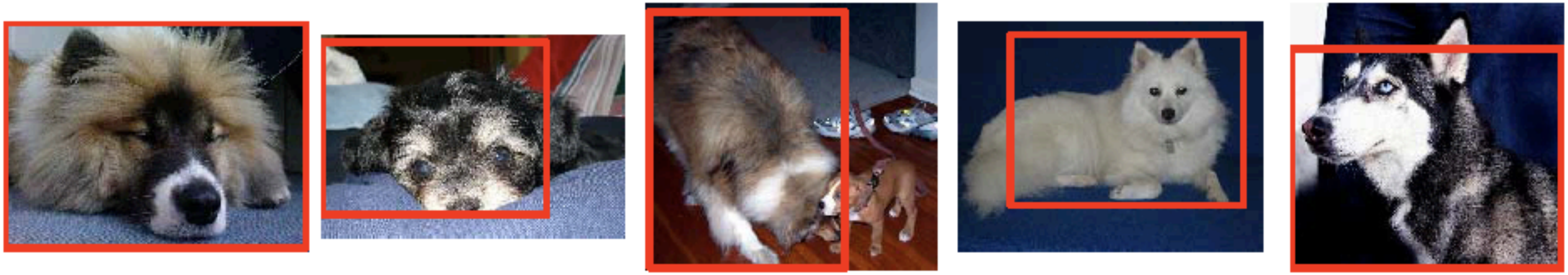


MITUCLA_HIERARCHY

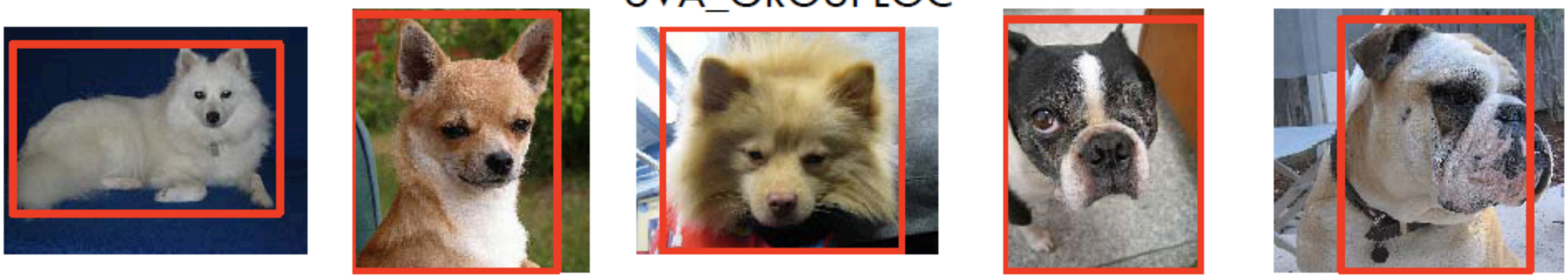


False Positives - Cat

UVA_DETMONKEY

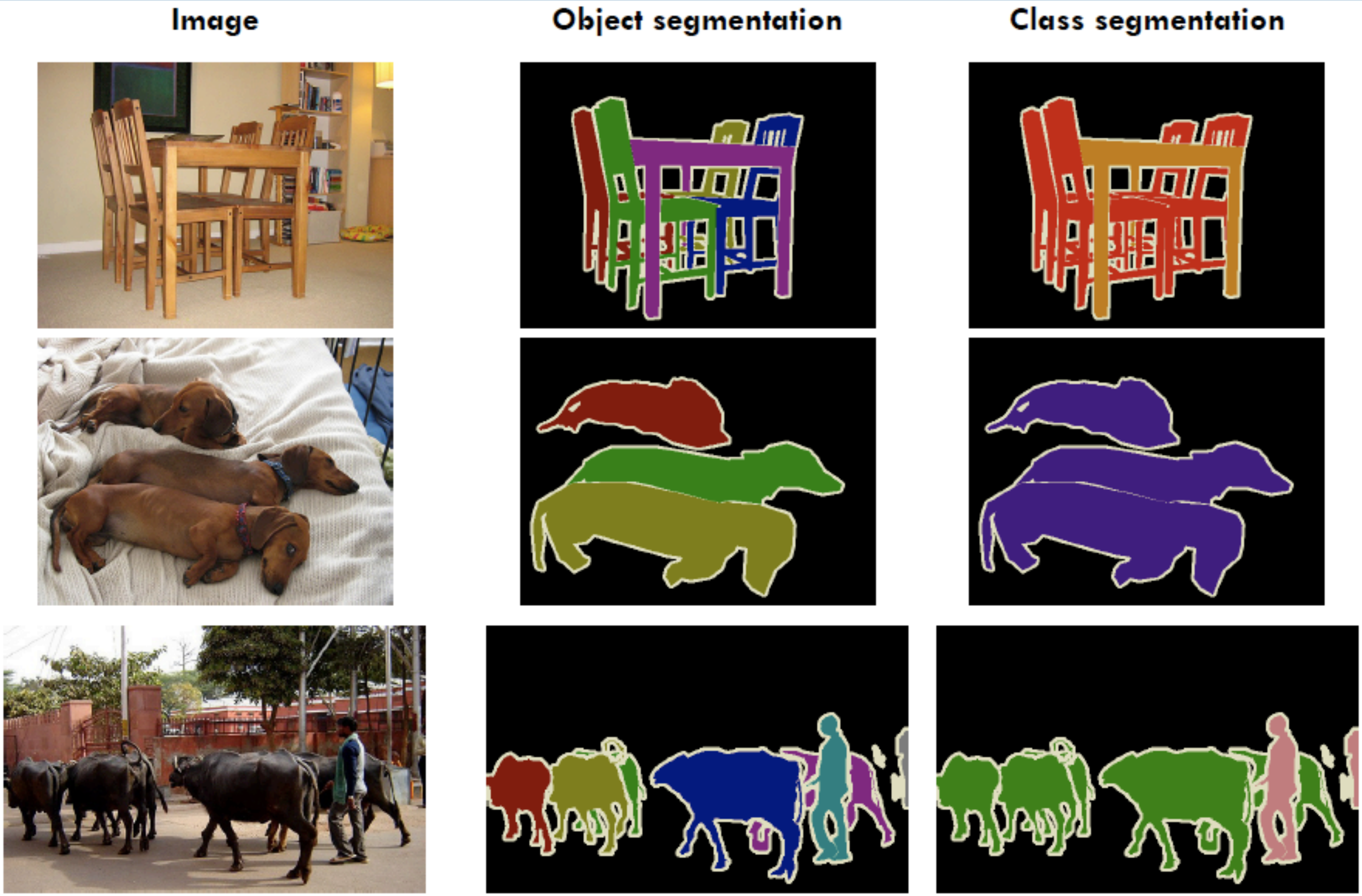


UVA_GROUPLOC



MITUCLA HIERARCHY

Example Annotations



Example Segmentations

Image

Ground truth

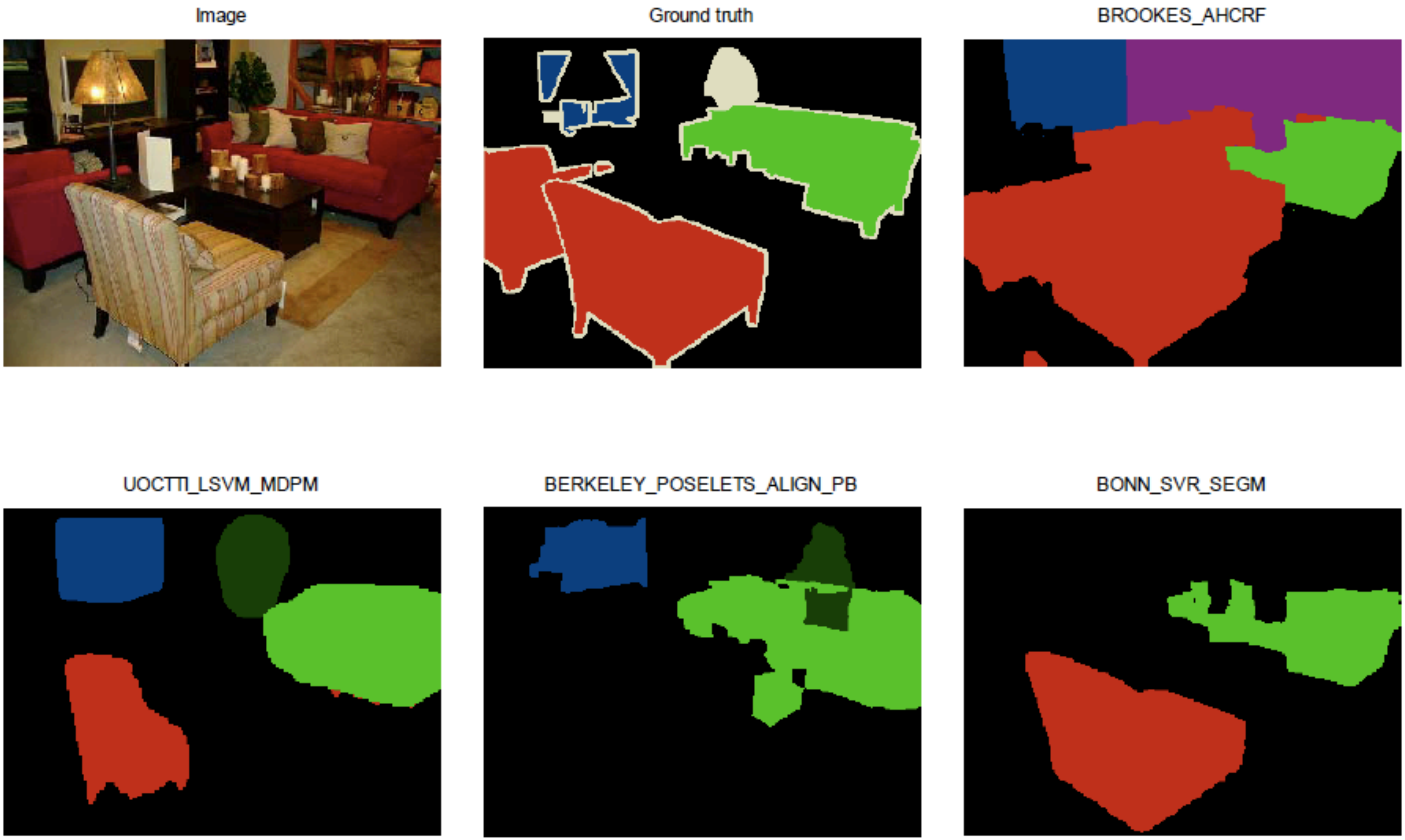
BERKELEY_POSELETS_ALIGN_PB

BROOKES_AHCRF

BONN_FGT_SEGM

CVC_HARMONY_DET

Example Segmentations



Accuracy by Class/Method

Trained on VOC2010 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
BONN_FGT_SEGM	36.5	82.5	54.6	22.5	25.1	27.6	40.0	60.2	48.3	39.4	7.3	30.8	21.3	25.3	34.9	54.1	36.6	22.5	45	17.6	33.5	37.0
BONN_SVR_SEGM	39.7	84.2	52.5	27.4	32.3	34.5	47.4	60.6	54.8	42.6	9.0	32.9	25.2	27.1	32.4	47.1	38.3	36.8	50.3	21.9	35.2	40.9
BROOKES_AHCRF	30.3	70.1	31.0	18.8	19.5	23.9	31.3	53.5	45.3	24.4	8.2	31.0	16.4	15.8	27.3	48.1	31.1	31.0	27.5	19.8	34.8	26.4
CVC_HARMONY	35.4	80.8	56.7	20.6	31.0	33.9	20.8	57.6	51.4	35.8	7.1	28.1	22.6	24.3	29.3	49.4	37.8	23.3	37.6	18.1	45.6	30.7
CVC_HARMONY_DET	40.1	81.1	58.3	23.1	39.0	37.8	36.4	63.2	62.4	31.9	9.1	36.8	24.6	29.4	37.5	60.6	44.9	30.1	36.8	19.4	44.1	35.9
STANFORD_REGLABEL	29.1	80.0	38.8	21.5	13.6	9.2	31.1	51.8	44.4	25.7	6.7	26.0	12.5	12.8	31.0	41.9	44.4	5.7	37.5	10.0	33.2	32.3
UC3M_GENDISC	27.8	73.4	45.9	12.3	14.5	22.3	9.3	46.8	38.3	41.7	0.0	35.9	20.7	34.1	34.8	33.5	24.6	4.7	25.6	13.0	26.8	26.1
UOCTTI_LSVM_MDPM	31.8	80.0	36.7	23.9	20.9	18.8	41.0	62.7	49.0	21.5	8.3	21.1	7.0	16.4	28.2	42.5	40.5	19.6	33.6	13.3	34.1	48.5

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
BERKELEY_POSELETS	34.7	82.0	49.7	23.3	20.6	19.0	47.1	58.1	53.6	32.5	0.0	31.1	0.0	29.5	42.9	41.9	43.8	16.6	39.0	18.4	38.0	41.5

- Best results exceed best detection-based results for all classes
- BERKELEY_POSELETS method uses additional training annotation for object detection: improves on “horse”

Overview –
Systems & Segmentation

- Recap and outlook
- Computer and Segmentation. Does segmentation matter?
- System
  - Build
  - Test
- **Segmentation principles**
- Segmentation using tools from lectures 1-7
- Mathematical Morphology

Segmentation

Goal: Segment the image into pieces/segments, i.e. regions that belong to the same object or that has the same properties.
Can also be seen as a problem of ’grouping’ of pieces (pixels, regions) together.
Edges, Ridges, Blobs, Interest Points, Textures - already a step towards segmentation.
More generally - segmentation is about cutting out the interestings regions/parts.

Segmentation

Some typical segmentation problems are:

- Cut an image sequence into shots
- Find manufactured parts in an industrial environment
- Find humans in images and video
- Find houses in satellite images
- Find faces in images

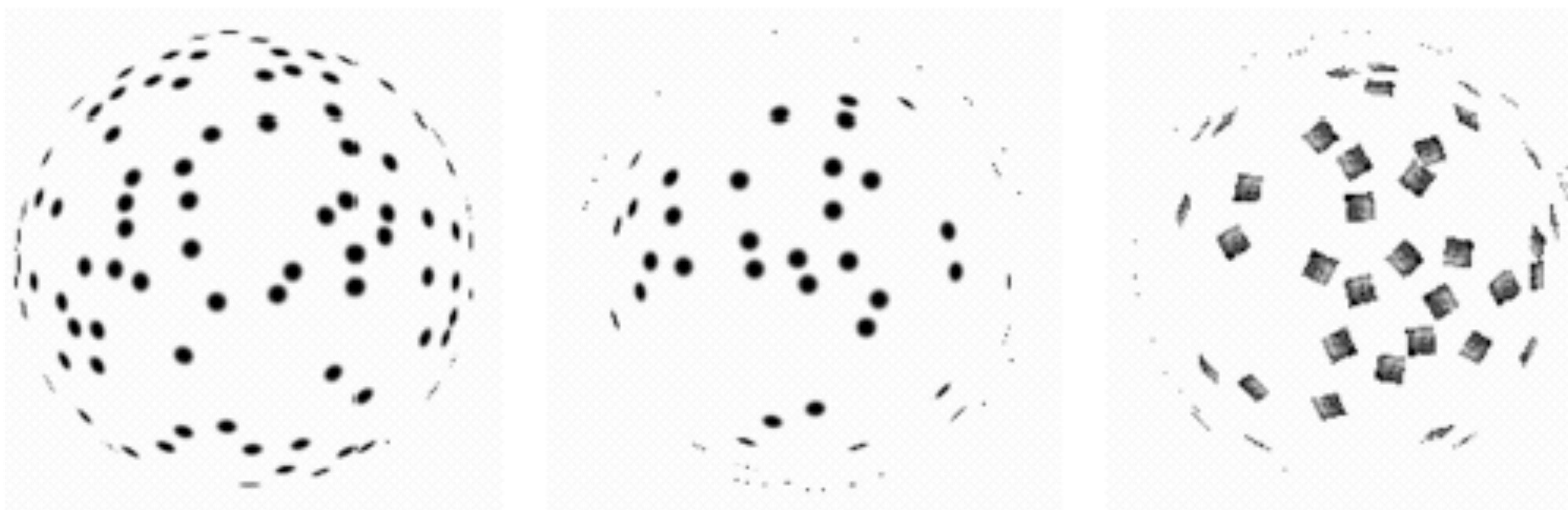
Example: OCR.
Example: Image interpretation
Example: Road user analysis
Example: Medical Image Analysis, e.g. Cell analysis

Segmentation

Using clustering:

- Segment images into pieces
- Fit lines to a set of points
- Fit a fundamental matrix to image pairs

In some cases it is easier to view segmentation as the problem of putting pieces together. This is usually called **grouping** (less precise) or **clustering** (which has a precise meaning in the field of pattern recognition).

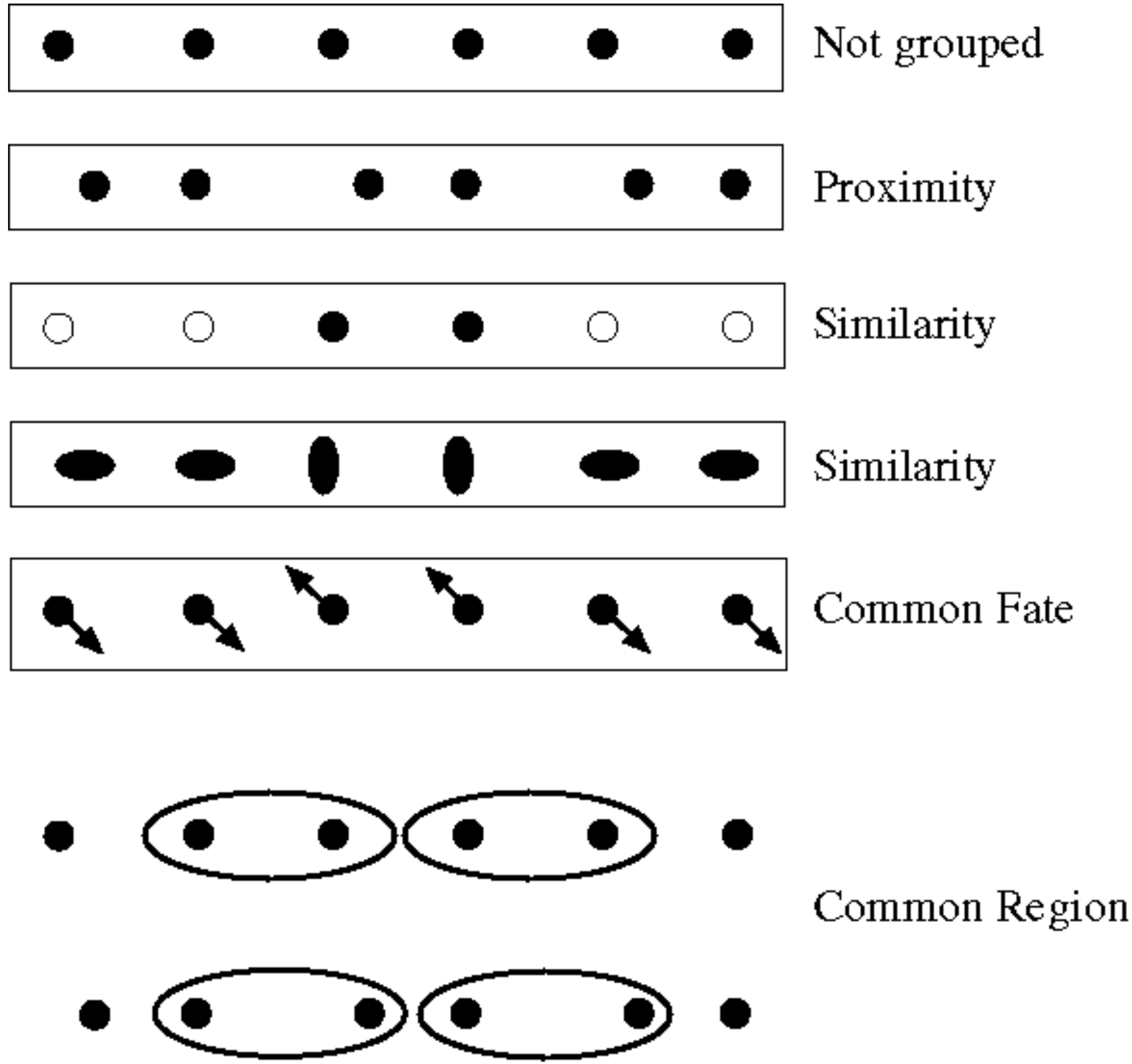


Why do these tokens belong together?

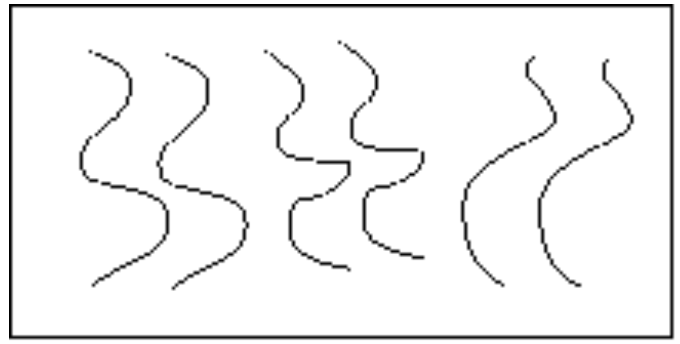
Gestalt theory

Around 1900 the ’Gestalt’ theory was developed by psychologists in Germany, the Berlin school. They developed a descriptive theory of mind and brain. Some principles that they discovered for human grouping of features are:

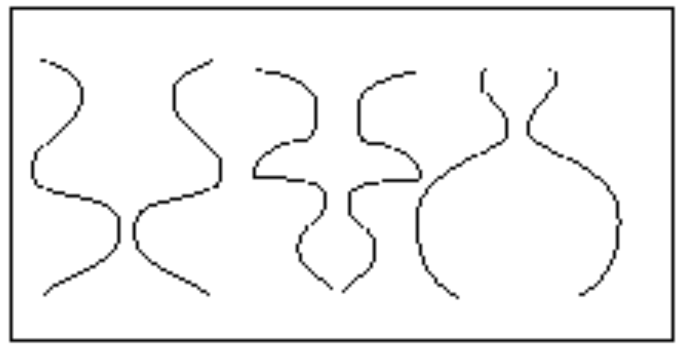
- Proximity
- Similarity
- Same fate
- Same region
- Closedness
- Symmetry
- Parallelism



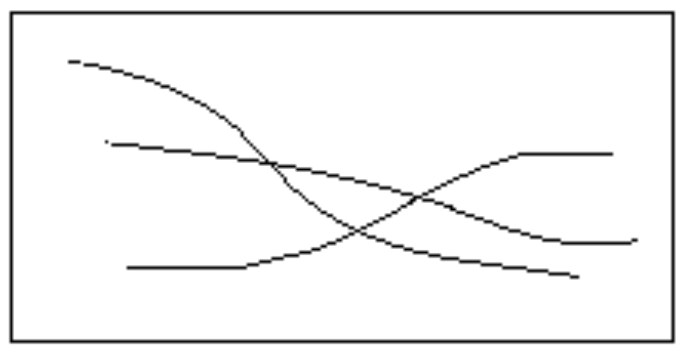
Not grouped
Proximity
Similarity
Similarity
Common Fate
Common Region



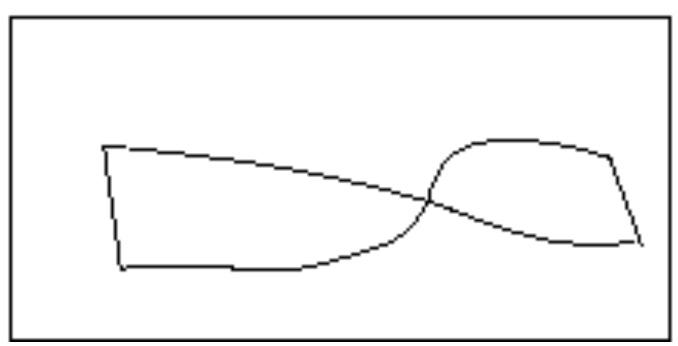
Parallelism



Symmetry



Continuity



Closure

Overview – Systems & Segmentation

- Recap and outlook
- Computer and Segmentation. Does segmentation matter?
- System
  - Build
  - Test
- Segmentation principles
- **Segmentation using tools from lectures 1-7**
- Mathematical Morphology

Segmentation using what we have learnt so far

- Pixelwise classification
- Colour-pixel classification
- Filterbanks and classification
- Deep learning
- Clustering
- Connected Components
- Cleaning up among the segments
  - Classification
  - Morphology

Pixels, clustering, segmentation

- At each pixel one could define a feature vector
  - Intensity f(i,j)
  - RGB colour channel (r,g,b)
  - Multispectral channel
  - Position (i,j)
  - Response from a filter bank
- Use machine learning to define a mapping from pixel feature vector to segment
- Either supervised (using lots of old examples) …
- … unsupervised (k-means, other clustering methods)

Overview – Systems & Segmentation

- Recap and outlook
- Computer and Segmentation. Does segmentation matter?
- System
  - Build
  - Test
- Segmentation principles
- Segmentation using tools from lectures 1-7
- **Mathematical Morphology**

Mathematical Morphology

Operations on binary images. Can be regarded as non-linear filtering.
$A = \{ (x, y) \in \mathbb{Z}^2 \,|\, f(x, y) = 1 \}$ is considered as a subset of the image.

Definition

Let A and $B \subset \mathbb{Z}^2$.
The **translation** of A with $x = (x_1, x_2) \in \mathbb{Z}^2$ is defined as

$$(A)_x = \{ c \,|\, c = a + x,\ a \in A \} \ .$$

The **reflection** of A is defined as

$$\hat{A} = \{ c \,|\, c = -a,\ a \in A \} \ .$$

Mathematical Morphology

Definition

The **complement** of A is defined as

$$A^c = \{\, c \mid c \notin A \,\} \ .$$

The **difference** of A and B is defined as

$$A - B = \{\, c \mid c \in A,\ c \notin B \,\} = A \cap B^c \ .$$

Mathematical Morphology

Let $B \subset \mathbb{Z}^2$ denote a **structure element.** (Usually B="a circle" with centre at the origin is chosen.)

Definition

The **dilatation** of A with B is defined by

$$A \oplus B = \{ x \mid (\hat{B})_x \cap A \neq \emptyset \} \ .$$

∎

This can also be written

$$A \oplus B = \{ x \mid ((\hat{B})_x \cap A) \subseteq A \} \ .$$

The dilation of A with B can be seen as extending A with B.

Mathematical Morphology

Definition

The **erosion** of *A* with *B* is defined by

$$A \ominus B = \{\, x \mid (\hat{B})_x \subseteq A \,\} \quad .$$

■

The erosion of A with B can be seen as diminishing (eroding) A with B.

Mathematical Morphology

Definition

The **opening** of A with B is defined by

$$A \circ B = (A \ominus B) \oplus B \quad .$$

Opening = first erosion, then dilation.

- Gives smoother contours.
- Removes narrow passages.
- Eliminates thin parts.

Mathematical Morphology

Definition

The **Closing** of A with B is defined by

$$A \cdot B = (A \oplus B) \ominus B \quad .$$

Closing = first dilation, then erosion.

- Gives smoother contours.
- Fills up small parts.
- Fills up holes.

Image dilation

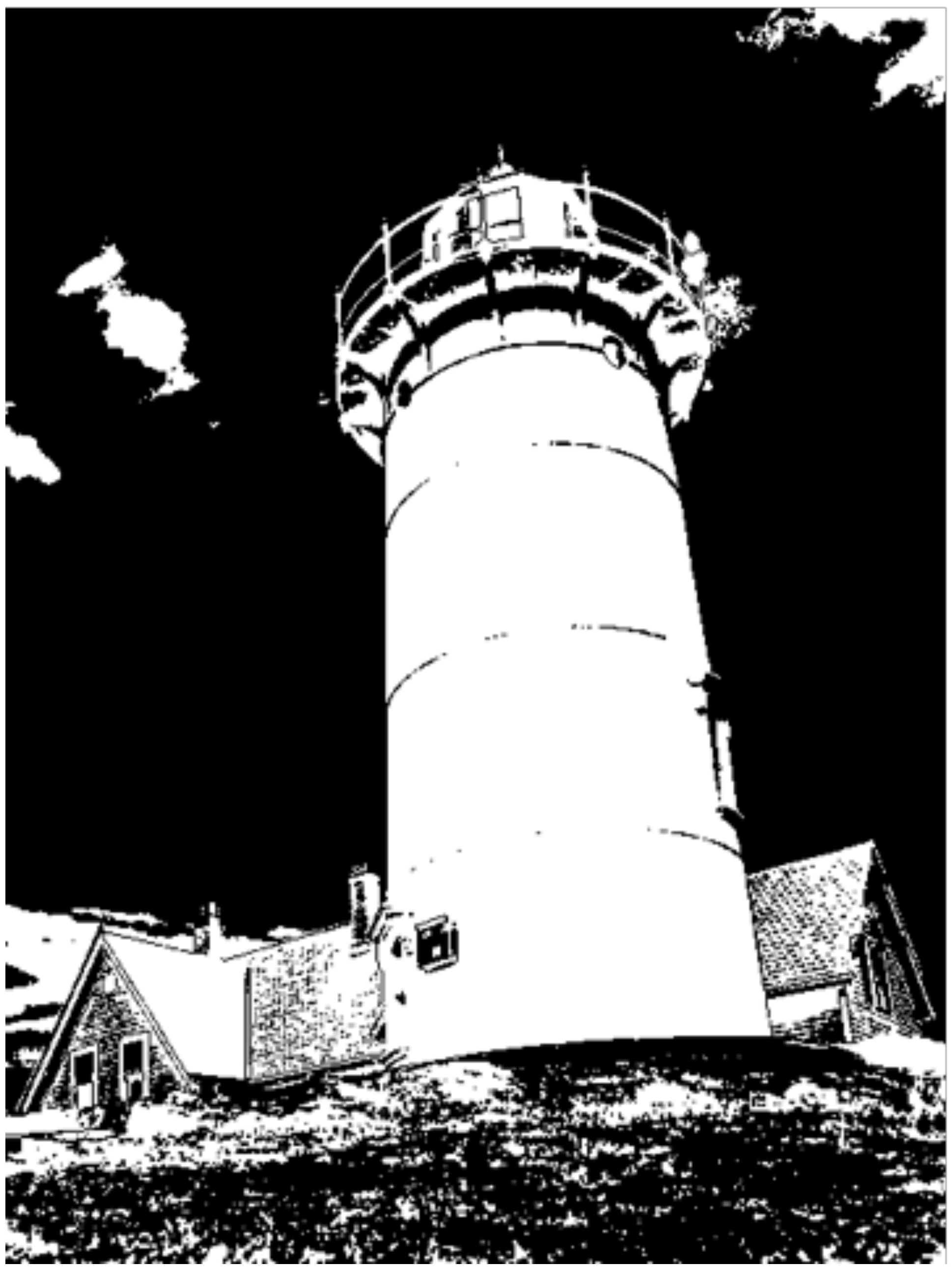
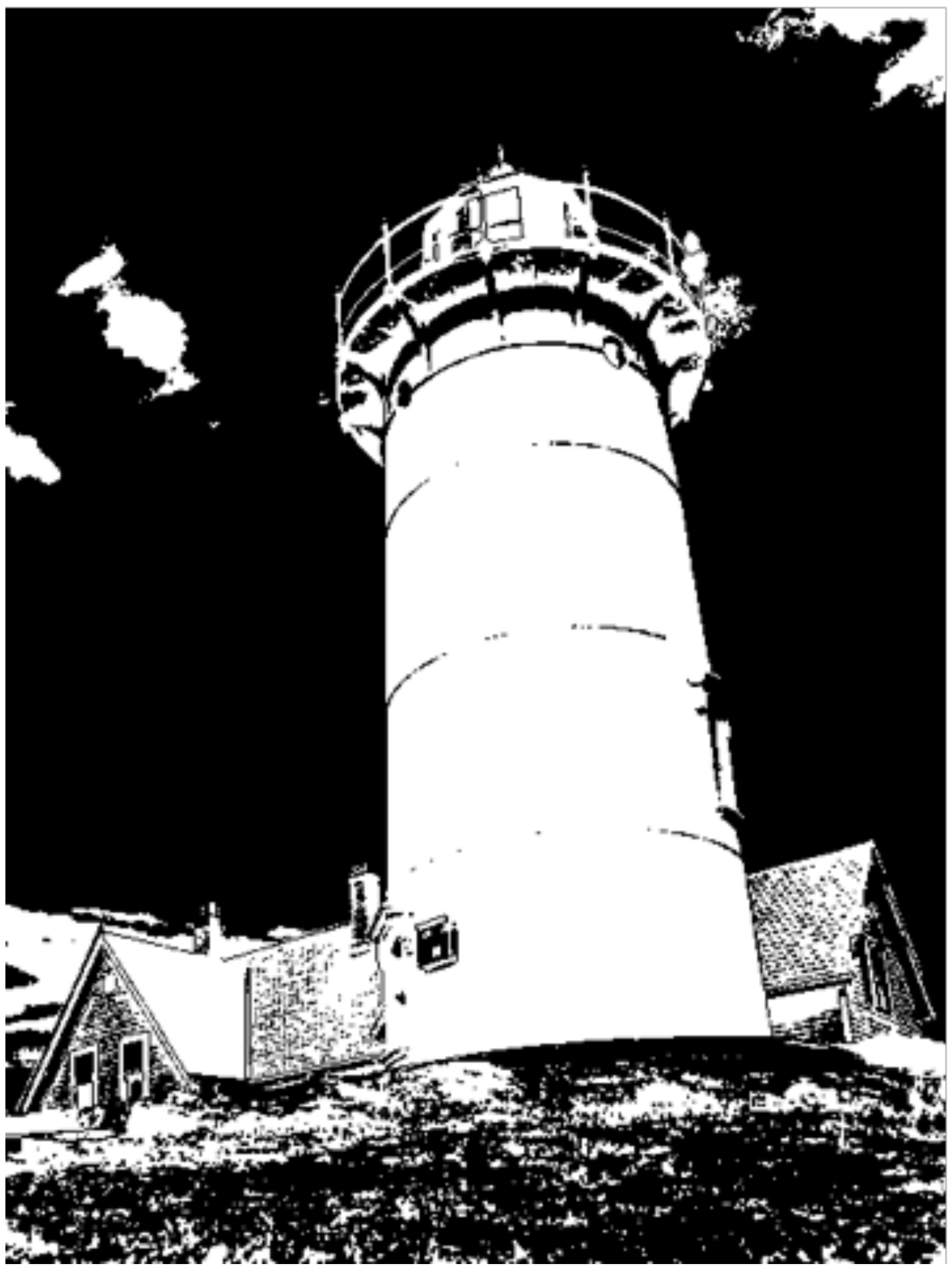


Image erosion



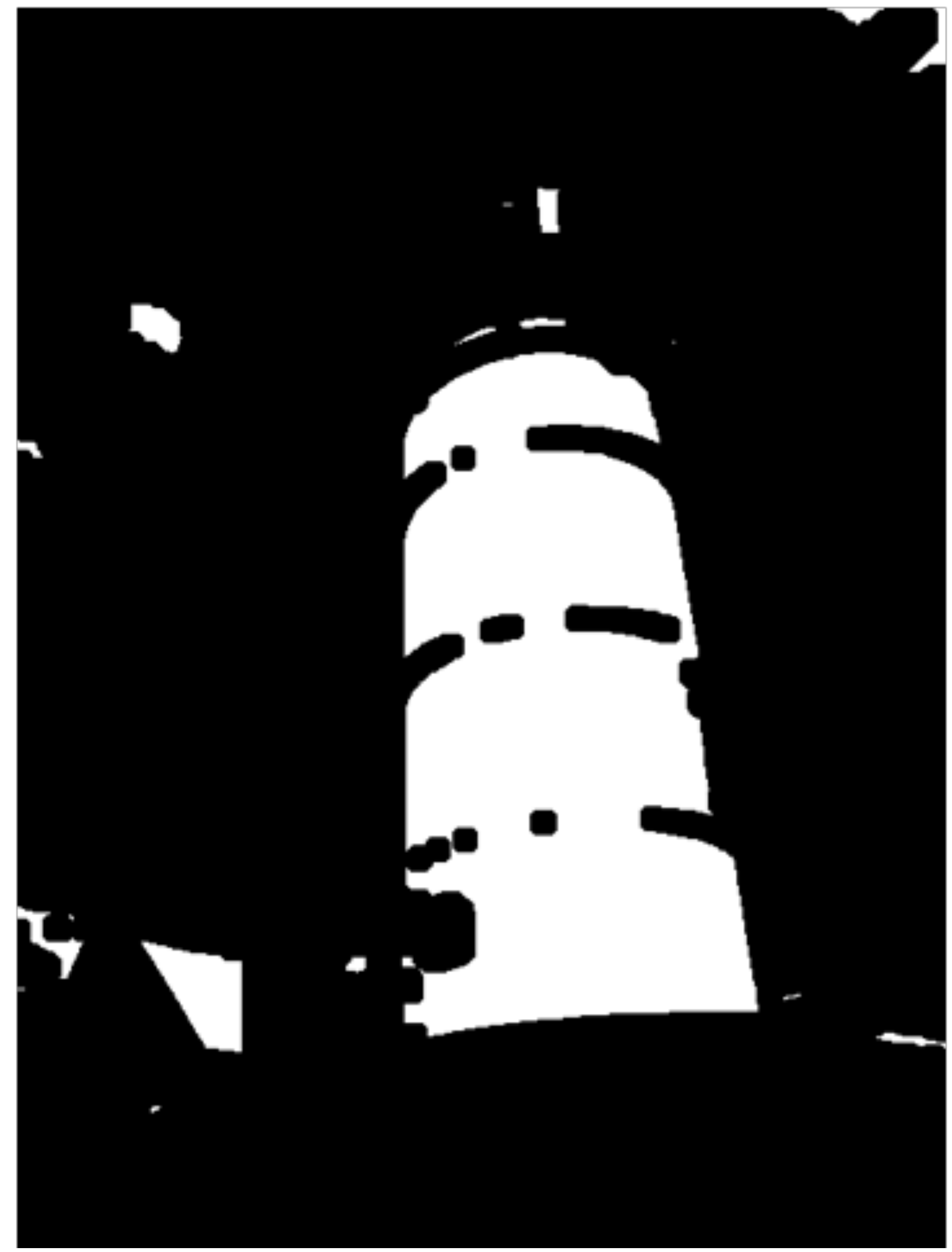
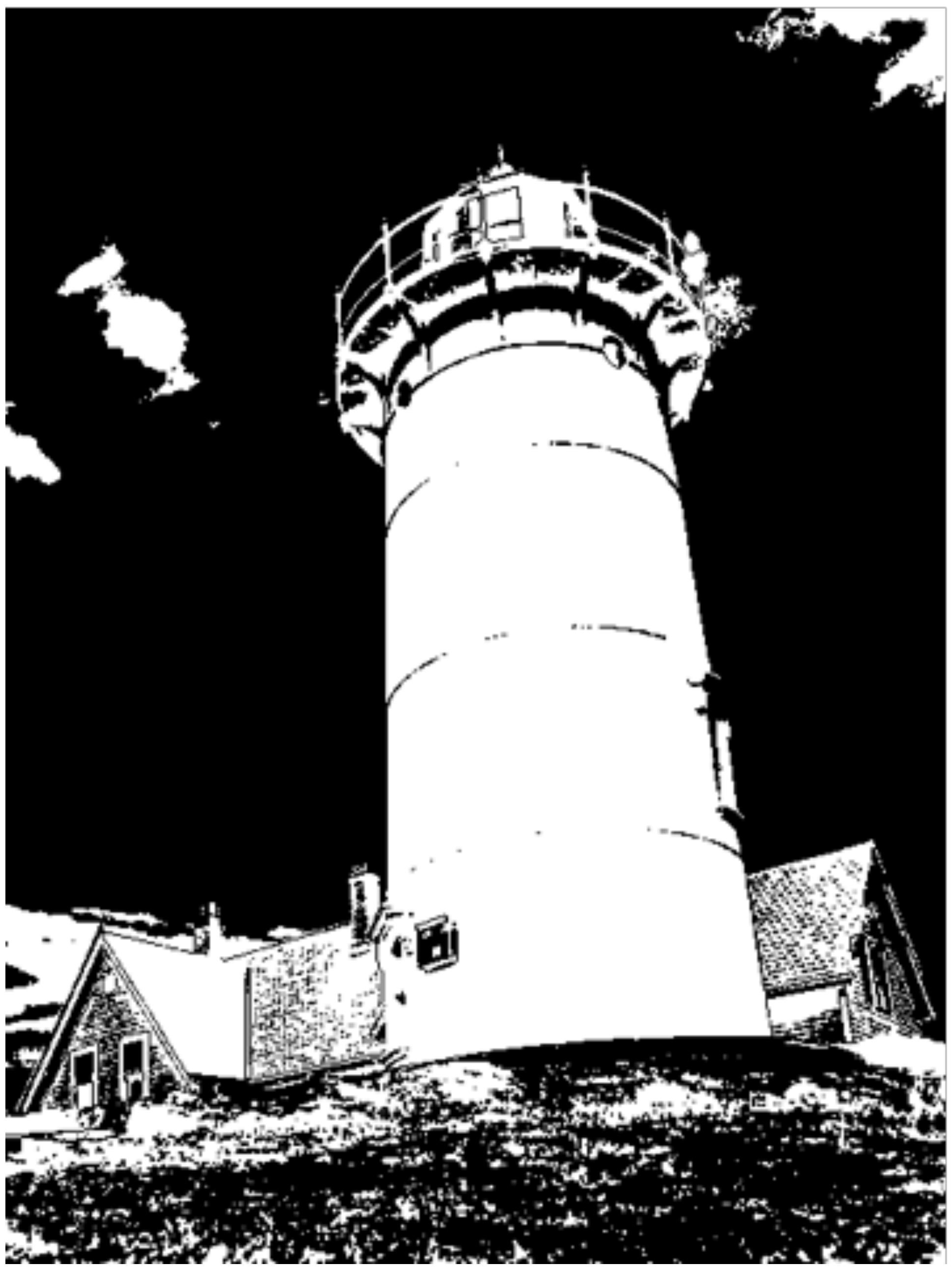
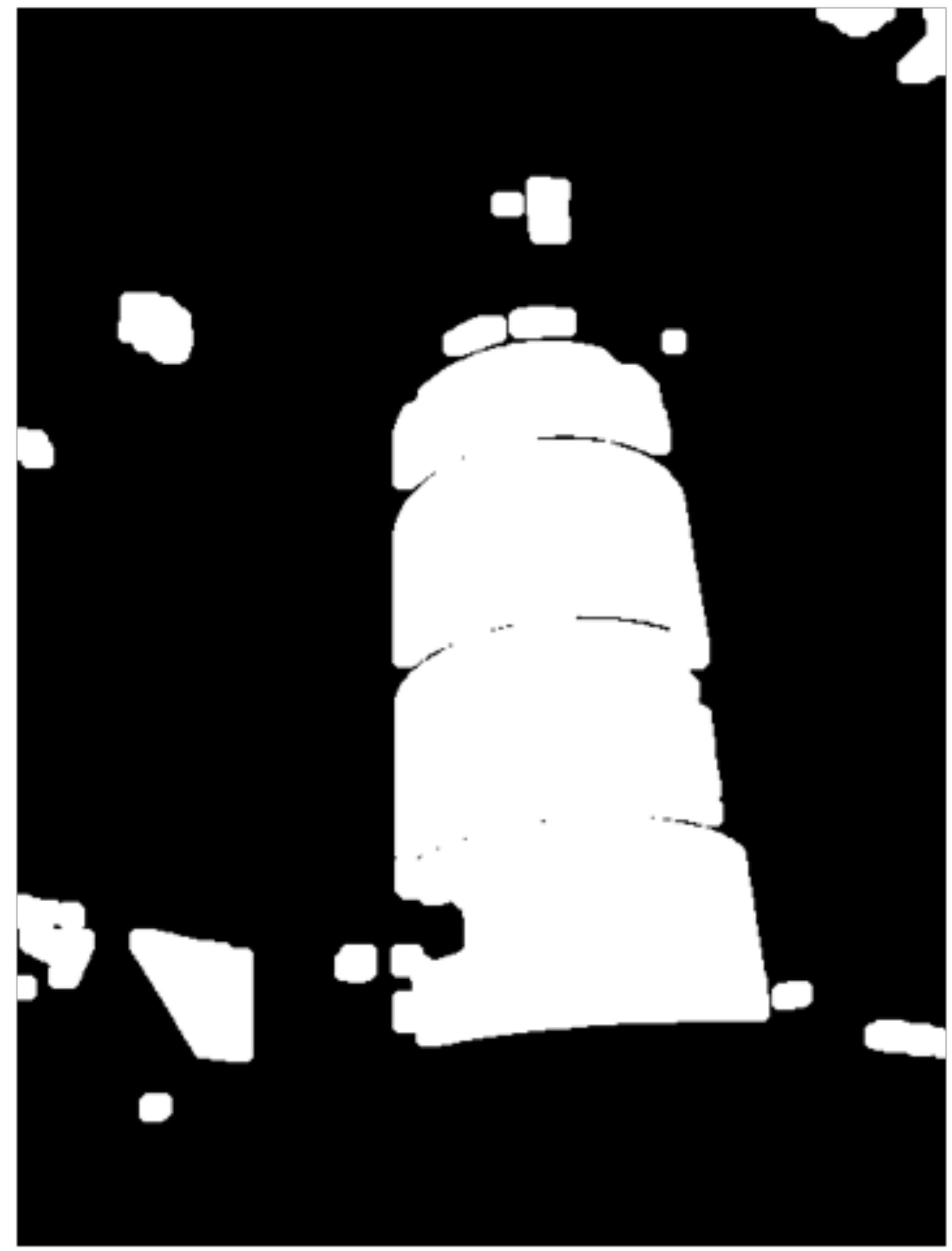


Image close

Image open



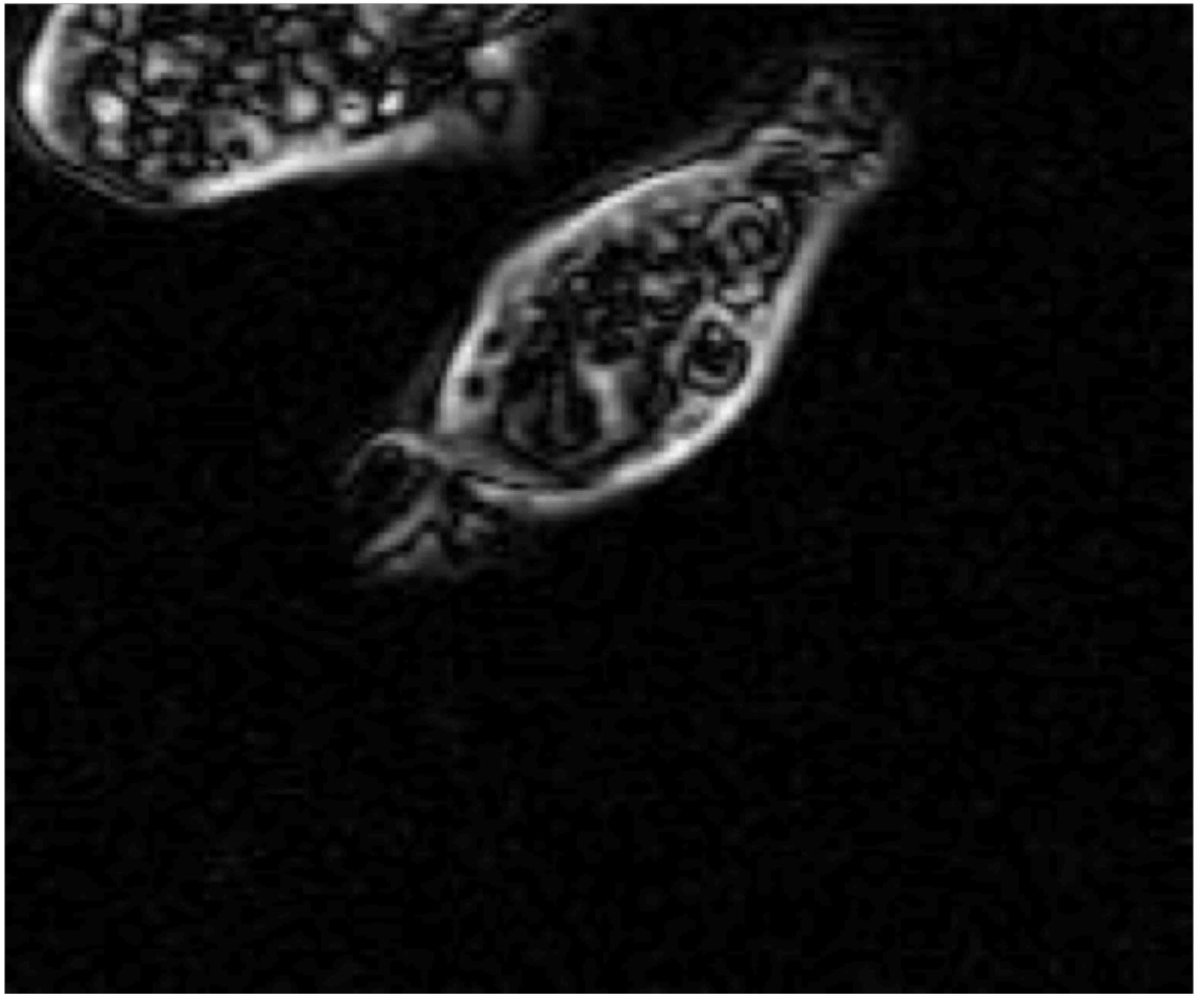


Morphological example

I

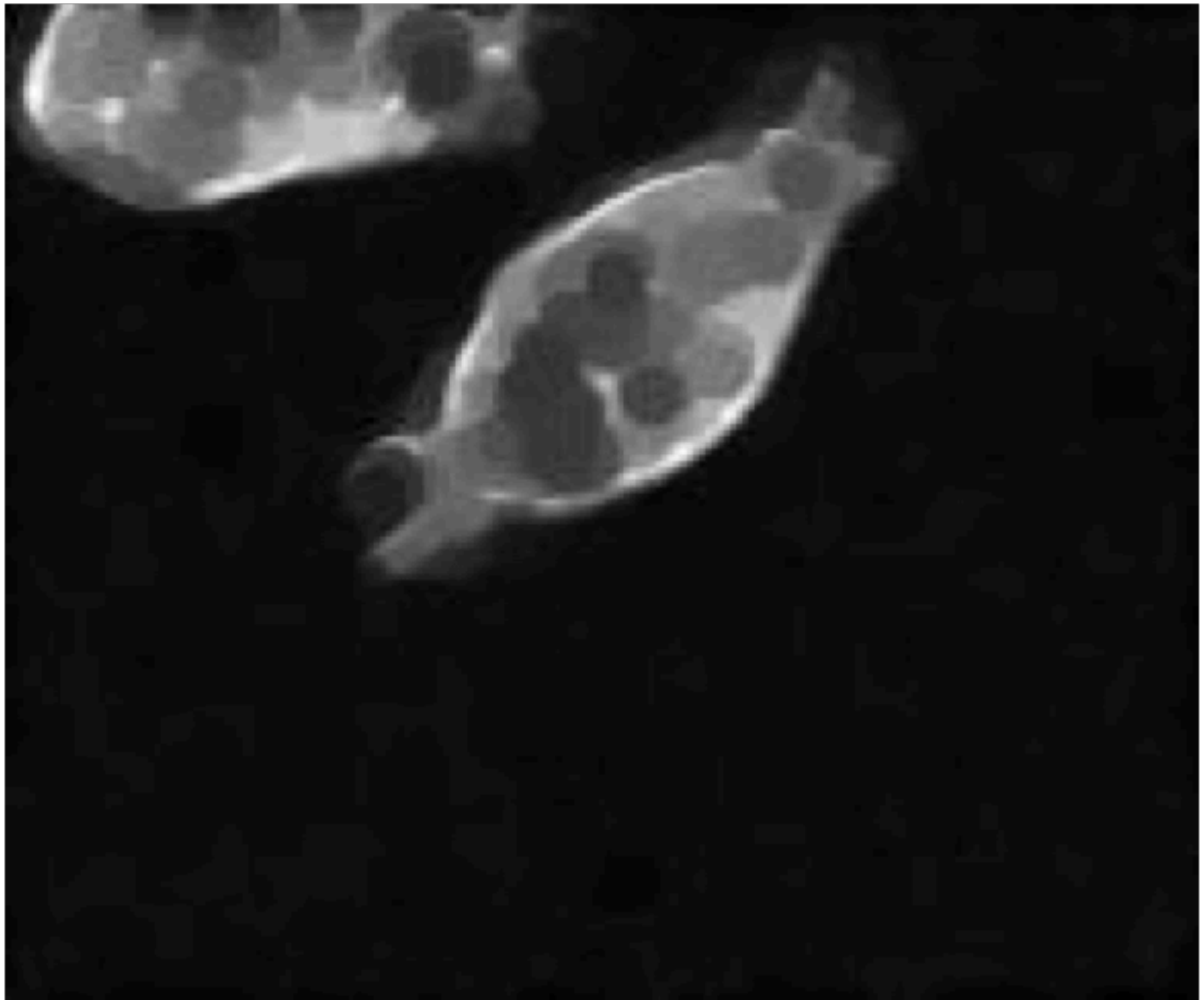
Morphological example

$$I_2 = |I - \text{median}(I)|$$



Morphological example

$$I_3 = \text{imopen}(I_2)$$



Morphological example

$$I_4 = I \cdot (I_3 > 0.2)$$

Distance Transform

Definition

Start with a binary image $A \subset \mathbb{Z}^2$ and a metric $d(x, y)$ that defines the distance between x and y and fulfils

- $d(x, y) \geq 0$ with equality iff $x = y$.
- $d(x, y) = d(y, x)$.
- $d(x, z) \leq d(x, y) + d(y, z)$ (the triangle inequality)

Try to for each pixel calculate the shortest distance to A.

Distance Transform

Different metrics gives different distances!

- $d^E(x, y) = \sqrt{x^2 + y^2}$ (Euclidean metric)
- $d^4(x, y) = |x| + |y|$ (Manhattan)
- $d^8(x, y) = \max(|x|, |y|)$ (Chess-board)
- d^{oct} = compromise between d^4 and d^8 (Octagonal)
- d^{ch} = Chamfer 3-4 given by the mask

$$\begin{bmatrix} 4 & 3 & 4 \\ 3 & 0 & 3 \\ 4 & 3 & 4 \end{bmatrix}$$

Distance Transform

The distance transform can be calculated by

- Forward propagation
- Backward propagation

A "mask" is propagated through the image row-wise from the upper left corner to the lower right corner and another "mask" is propagated in the reverse direction. This procedure is repeated until convergence.

Distance Transform

The **Skeleton** to a binary image, *A*, is defined by

- For each point, *x*, in *A* find the closest boundary point.
- If there are more than one closest boundary point, then *x* belongs to the skeleton of *A*.

The skeleton is dependent on the chosen metric!

Given the skeleton and the actual distance to the boundary for each skeleton point, the binary image *A* can be recovered.

Distance Transform

Calculating the skeleton:

- Using a distance map
- Using morphological operations (thinning).

Overview – Systems & Segmentation

- Recap and outlook
- Computer and Segmentation. Does segmentation matter?
- System
  - Build
  - Test
- Segmentation principles
- Segmentation using tools from lectures 1-7
- Mathematical Morphology

LUND
UNIVERSITY

350