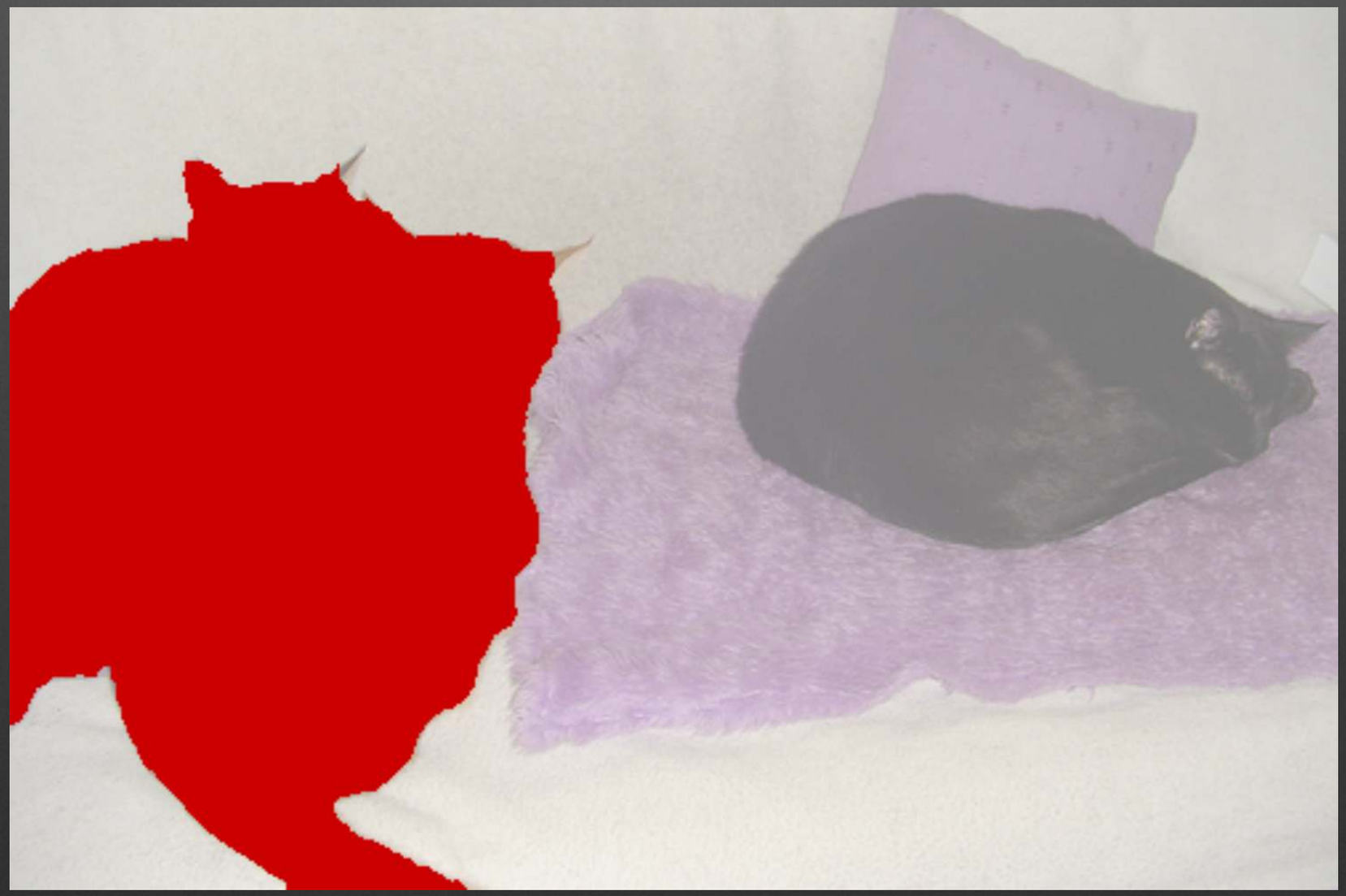
Geodesic Object Proposals

Philipp Krähenbühl
Stanford University

Vladlen Koltun
Adobe Research

Object Proposals

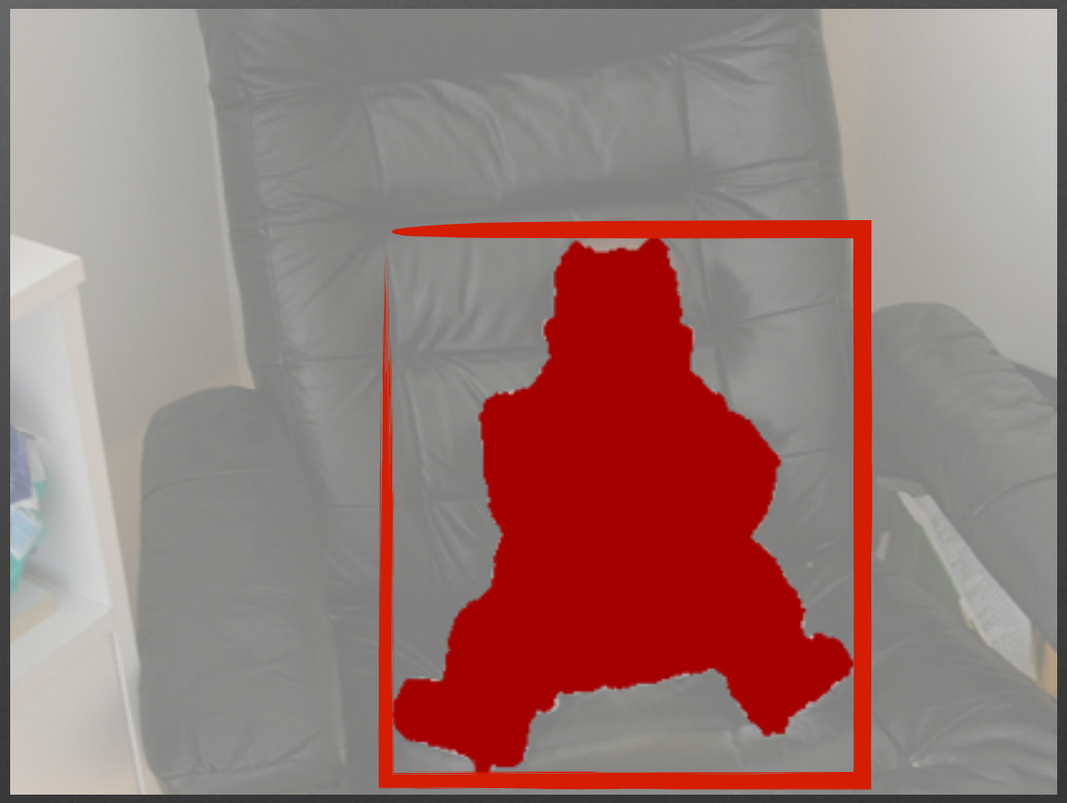
Find small set of proposals
that includes all objects in a scene



Object Proposals

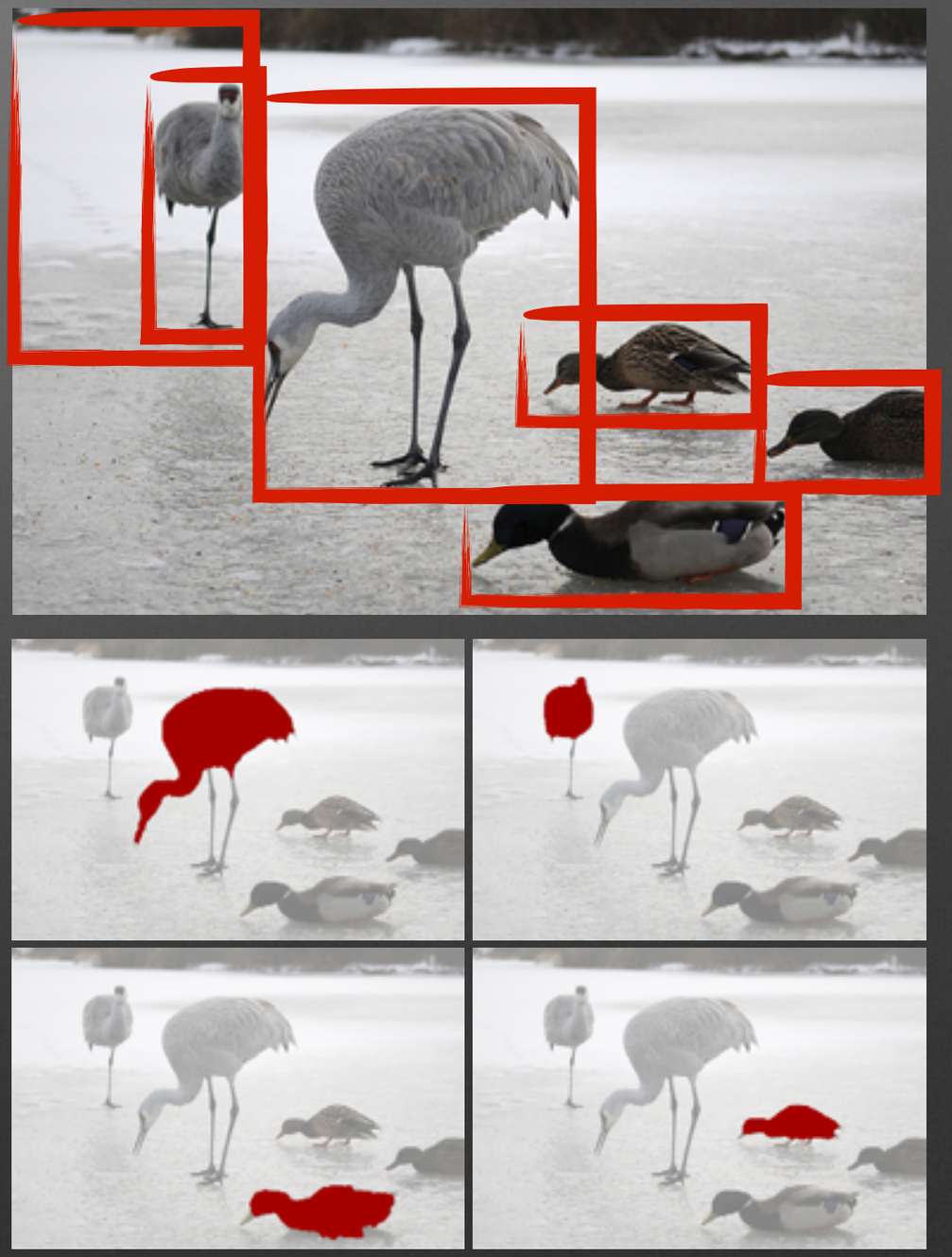
Bounding Box Proposals

Segment Proposals



Uses of object proposals

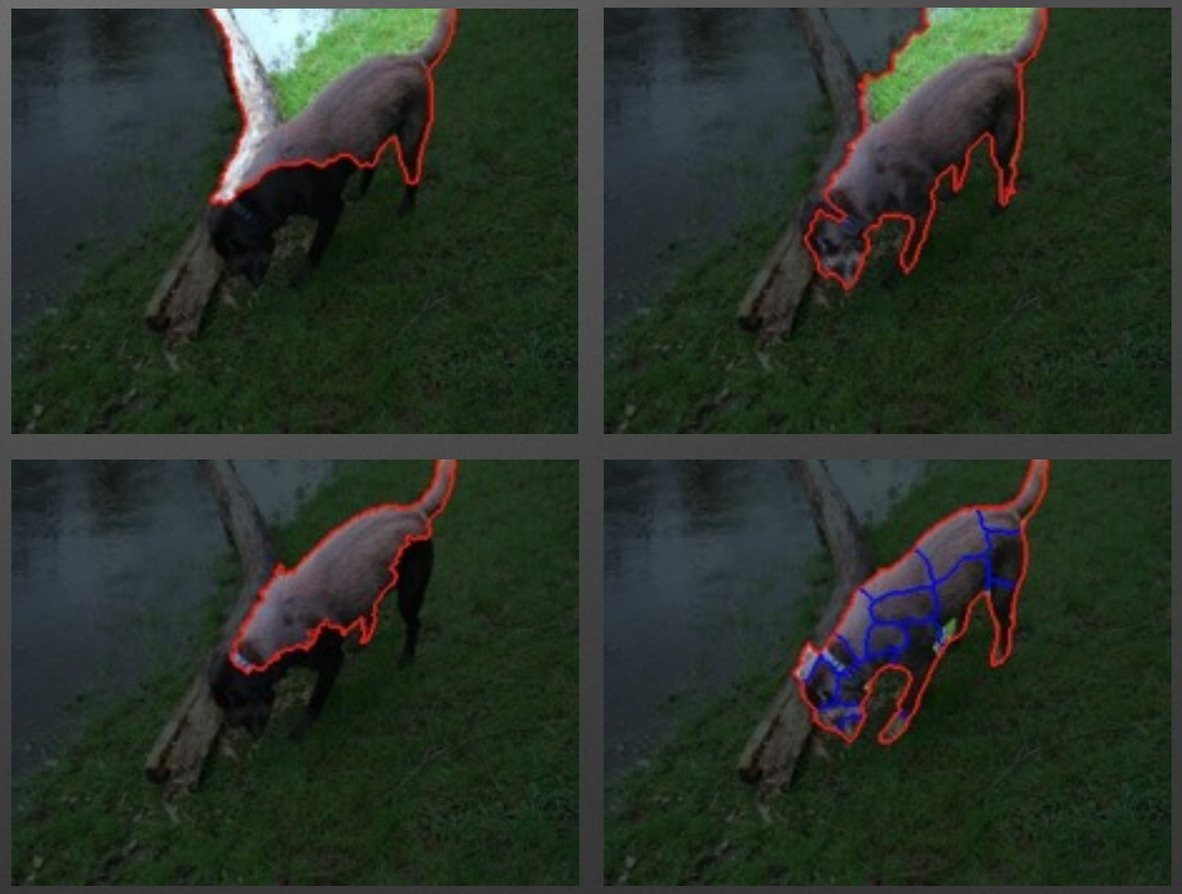
- Object detection
  - faster
  - more expensive features
  - used in almost all state-of-the art methods
- Multi-class image segmentation
  - used in state-of-the art
  - more expressive features



Improving Spatial Support for Objects via Multiple Segmentations

[Malisiewicz and Efros 2007]

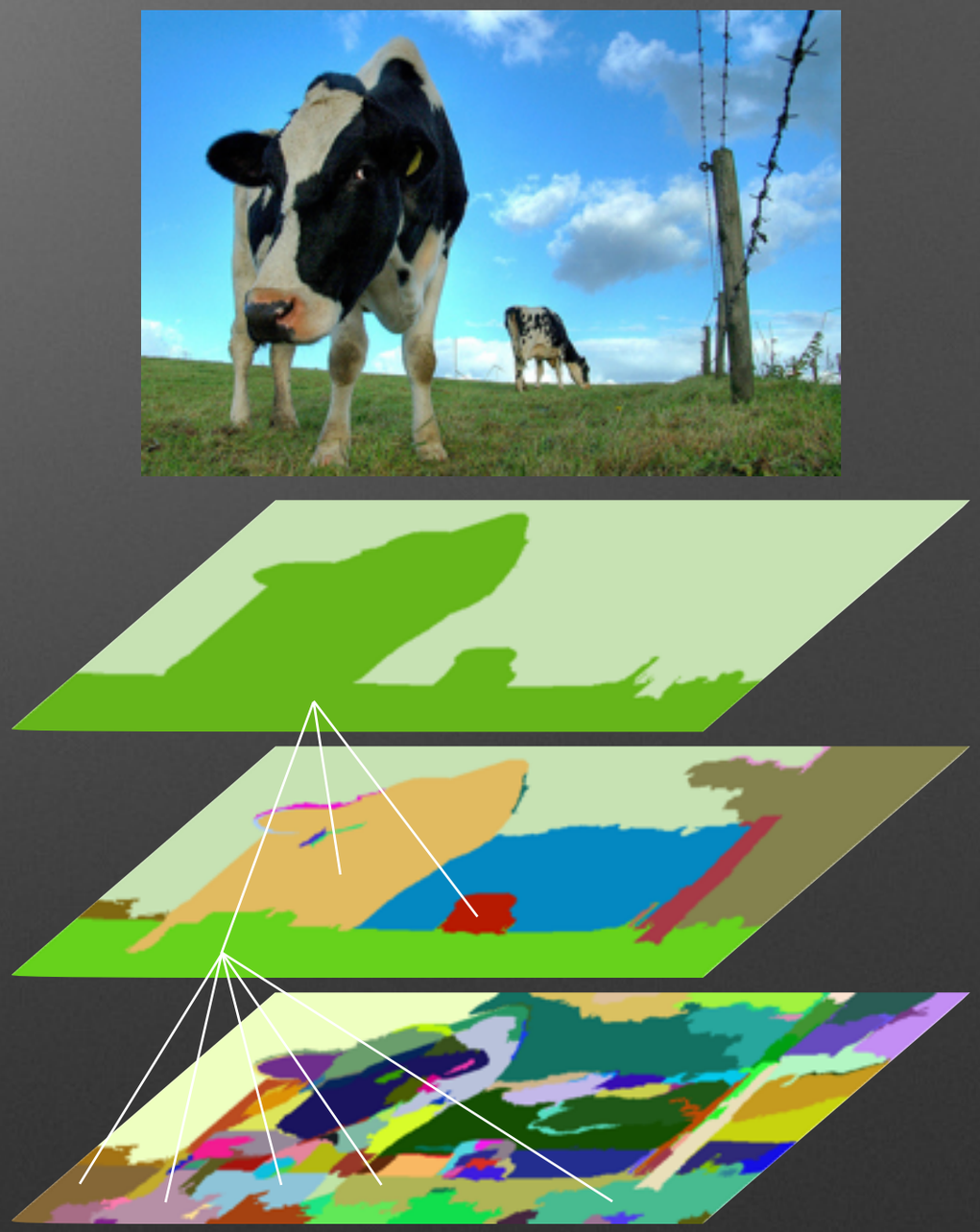
- Soup of segments
  - 3 different unsupervised segmentation algorithms
  - vary parameters
  - combination of up to 3 adjacent segments
- Slow (few minutes / image)
- Hard to segment larger objects well



Segmentation As Selective Search for Object Recognition

[van de Sande et al. 2011]

- Hierarchical segmentations
  - same algorithm
  - different parameters and color spaces
- Fast (few seconds / image)
- Good bounding box proposals
- OK segment proposals



Category Independent Object Proposals
[Endres and Hoiem 2010]

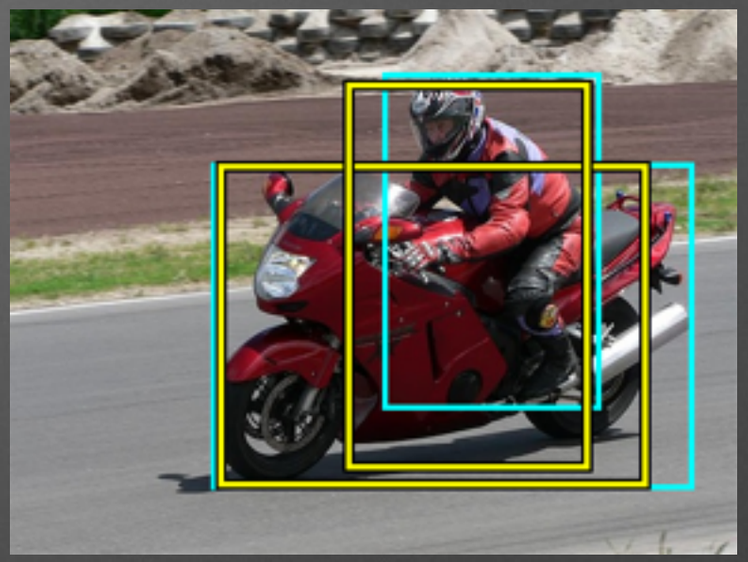
- Series of binary segmentations
  - Superpixel CRF model
  - Select seeds
  - Boosted decision tree as unary
  - Pairwise: boundary detector
  - Optimize with GraphCuts
- Good proposals
- Slow (100+ sec / image)

Constrained Parametric Min-Cuts for Automatic Object Segmentation

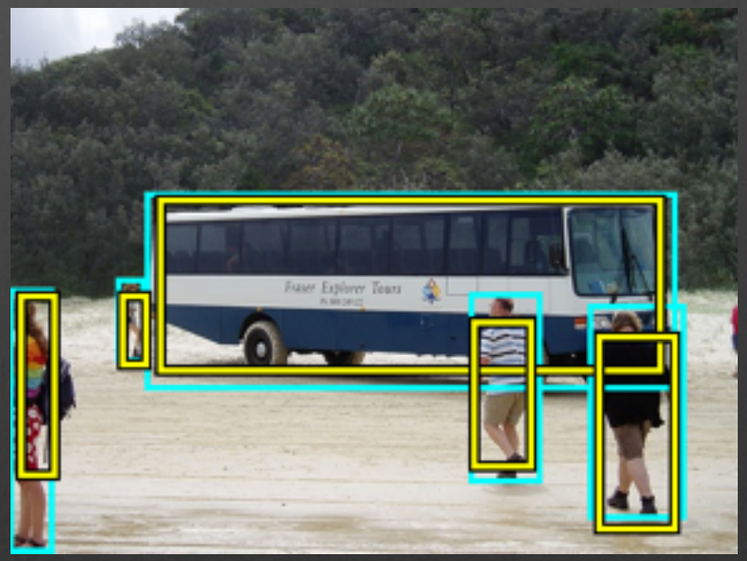
[Carreira and Sminchisescu 2010]

- Series of binary segmentations
  - Pixel CRF model
  - Regularly sampled seeds
  - Color based unary
  - GraphCuts with pairwise term
  - Parametric cut
    - slowly inflate segmentation
- Best proposals
- Slow (200+ sec / image)

What is an object?
[Alexe et al. 2010]



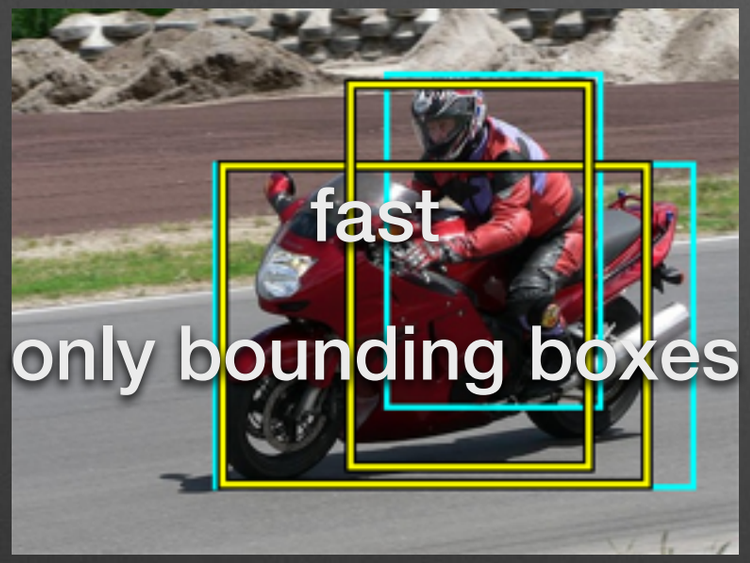
- Bounding box proposals
  - Train a general classifier for "objectness"
- fast (few seconds, BING[1]: few ms)
- Only works for bounding boxes



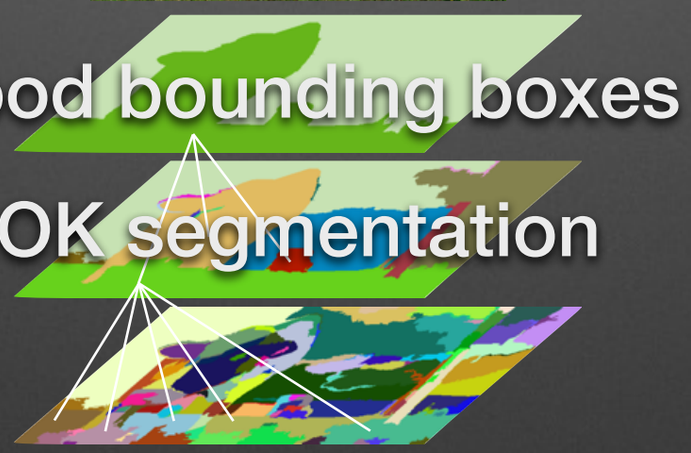
[1] BING: Binarized Normed Gradients for Objectness Estimation at 300fps [Cheng et al. 2014]

Prior work - summary

Objectness



Segmentation based



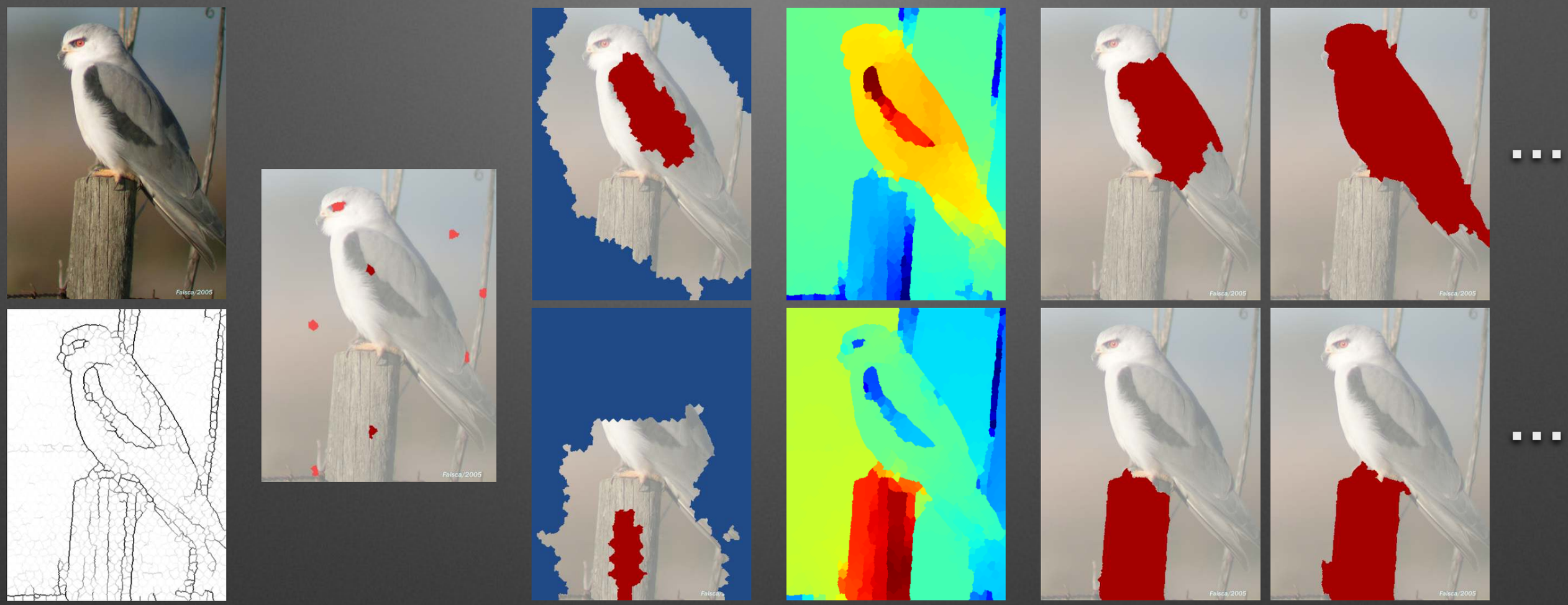
Seed / GraphCuts

Geodesic image segmentation

- (Signed) geodesic distance transform
  - Shortest path to background and foreground scribble
  - small within objects
  - large between objects
  - efficient to compute
- State of the art in interactive image and video segmentation[1,2]

[1] Geodesic Matting: A Framework for Fast Interactive Image and Video Segmentation and Matting [Bai and Sapiro 2008]
[2] Geodesic Image and Video Editing [Criminisi et al. 2011]

Geodesic object proposals



image, boundary map and superpixels

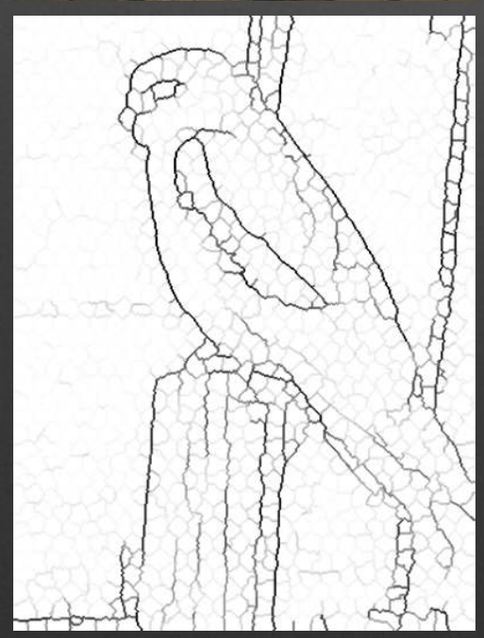
select seeds

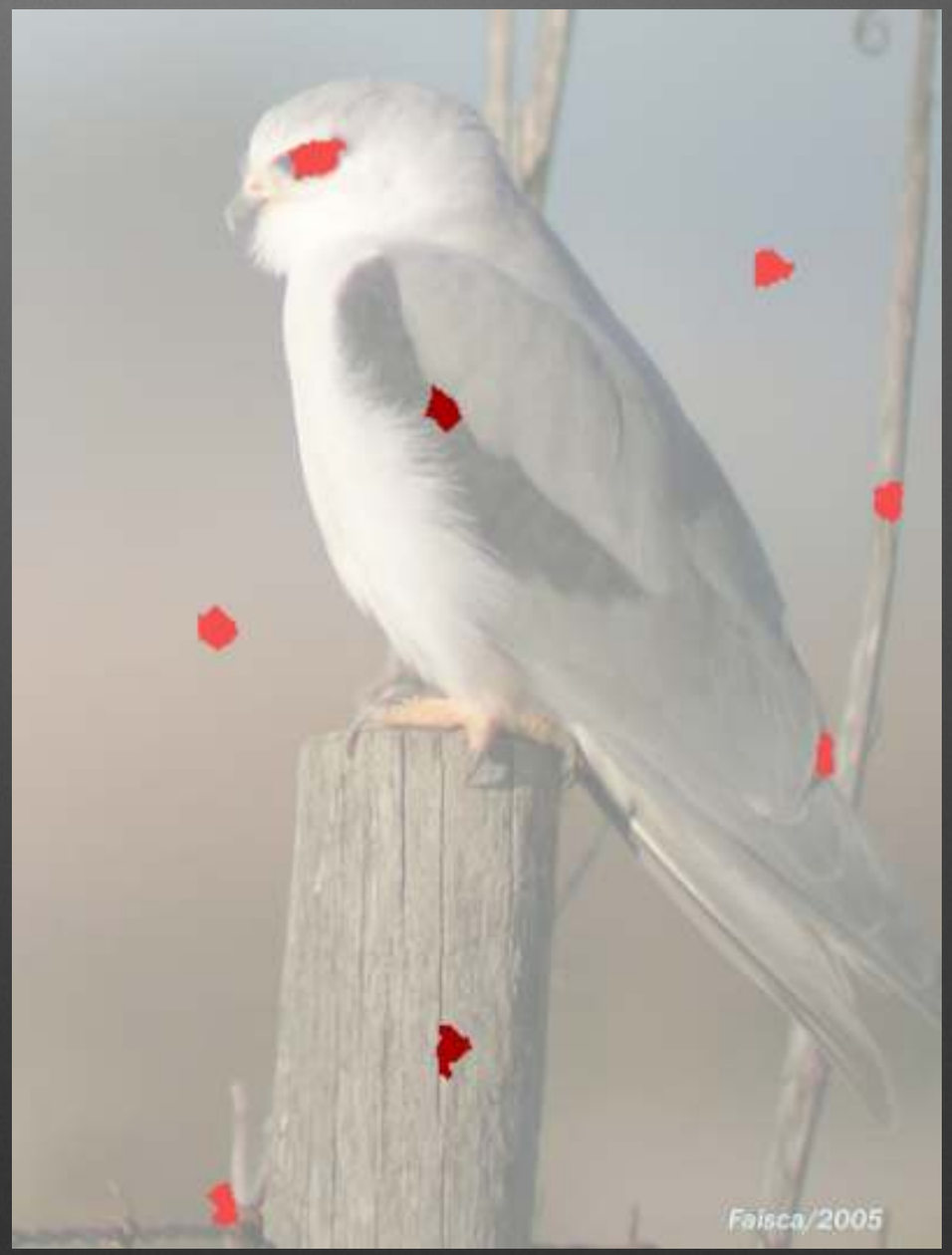
foreground background masks

geodesic distance transform

multiple proposals

Seed placement

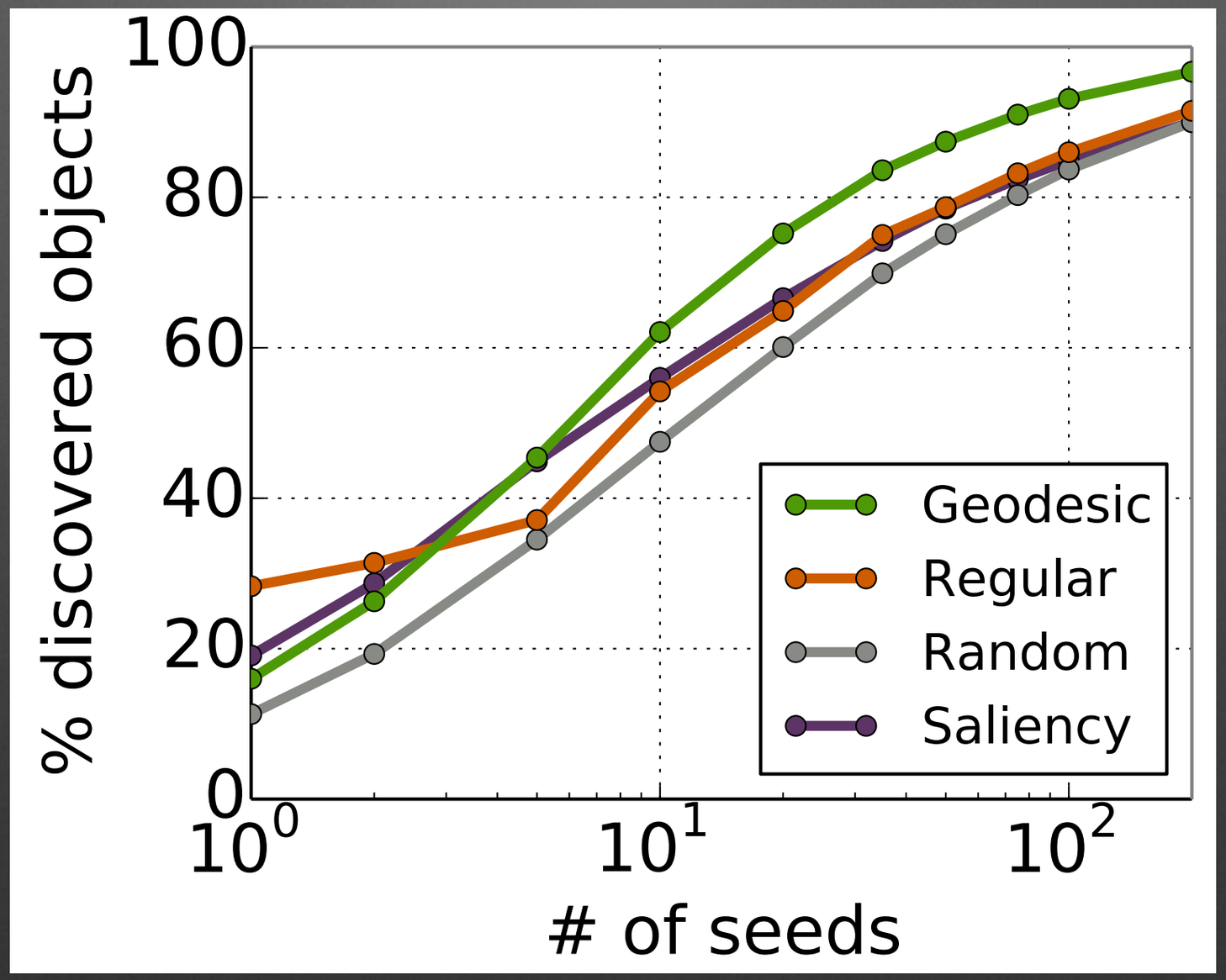




Seed placement

- Place a seed in each object
- Regular or random sampling
  - miss small objects
- Saliency based sampling
  - miss non-salient objects
- Geodesic placement
  - regular sampling in geodesic space
  - greedily place next seed at maximal geodesic distance

Seed placement



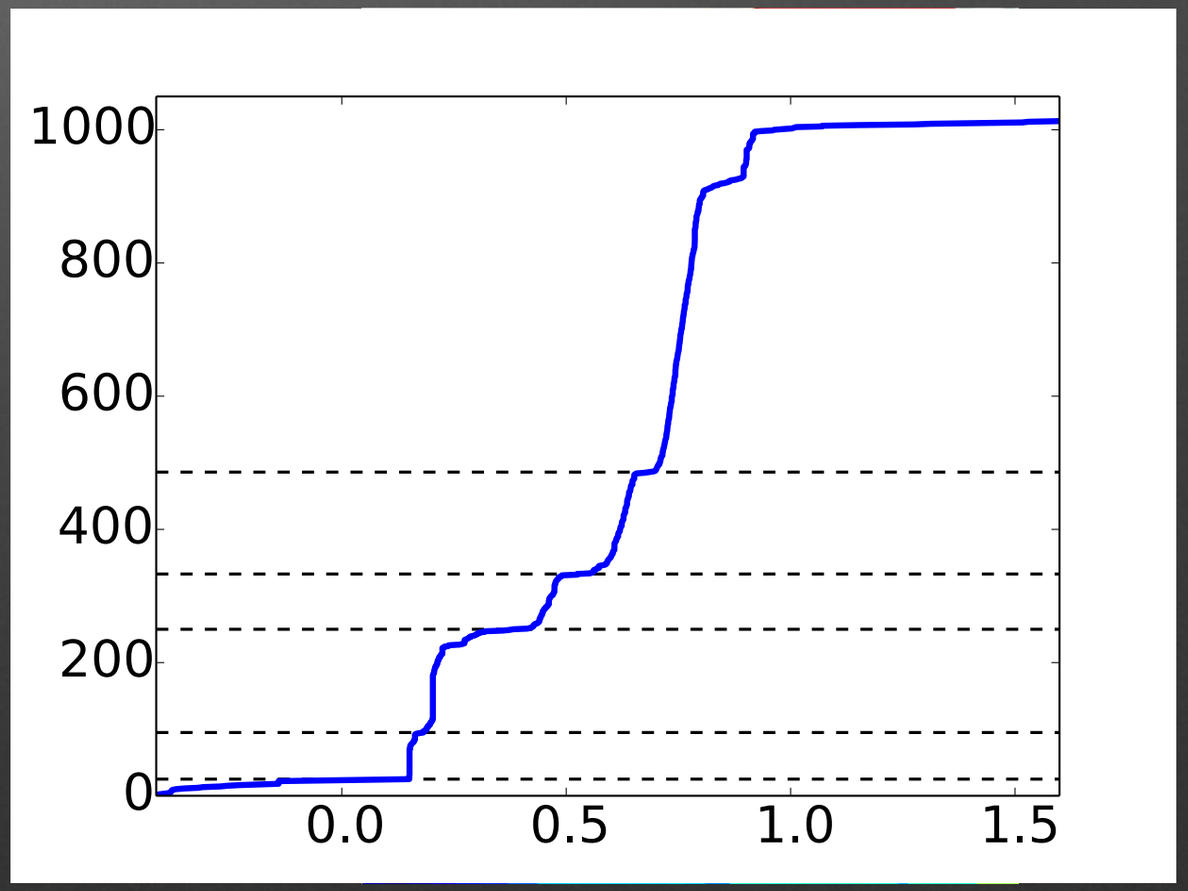
Mask generation

Mask generation

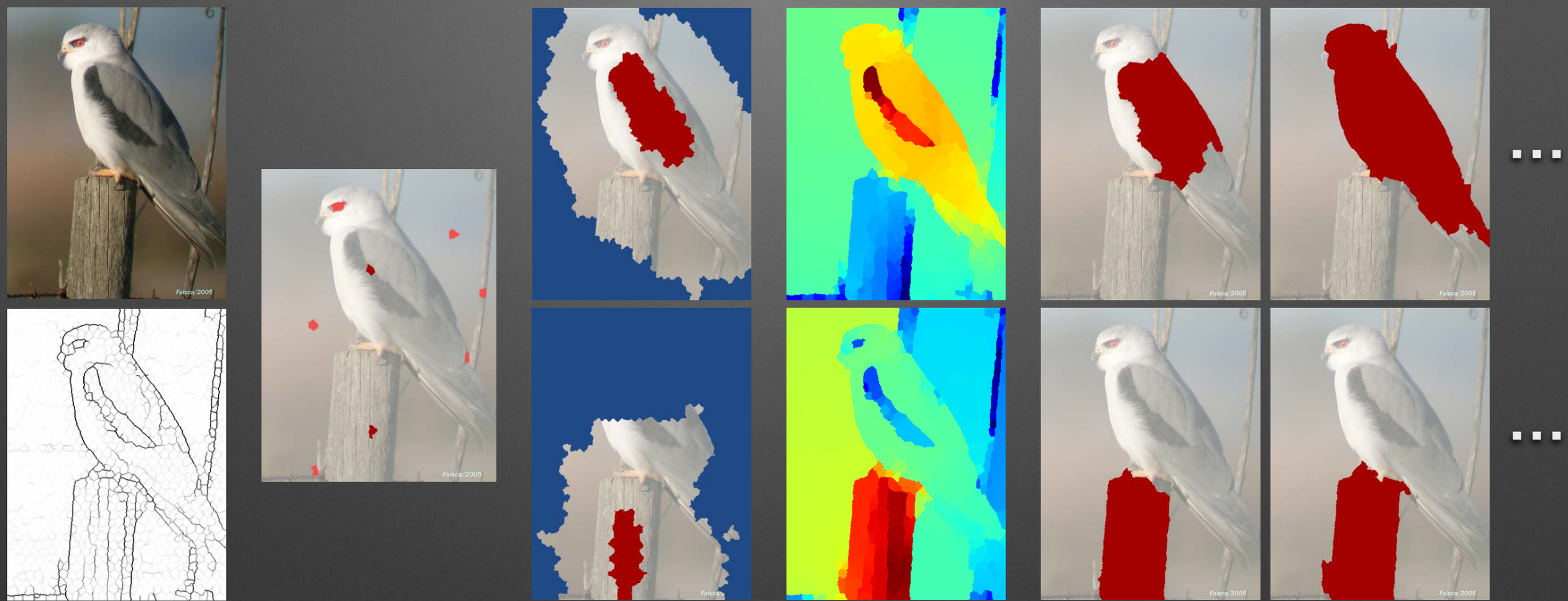
- No errors in masks
  - errors propagate
- Foreground mask
  - seed
- Background mask
  - boundary
  - empty

Geodesic segmentation

- Signed geodesic distance transform
- Each level set is a segmentation
- Find critical level sets
  - stationary points in geodesic function
  - evolution of Eikonal equation



Baseline GOP



image, boundary map and superpixels

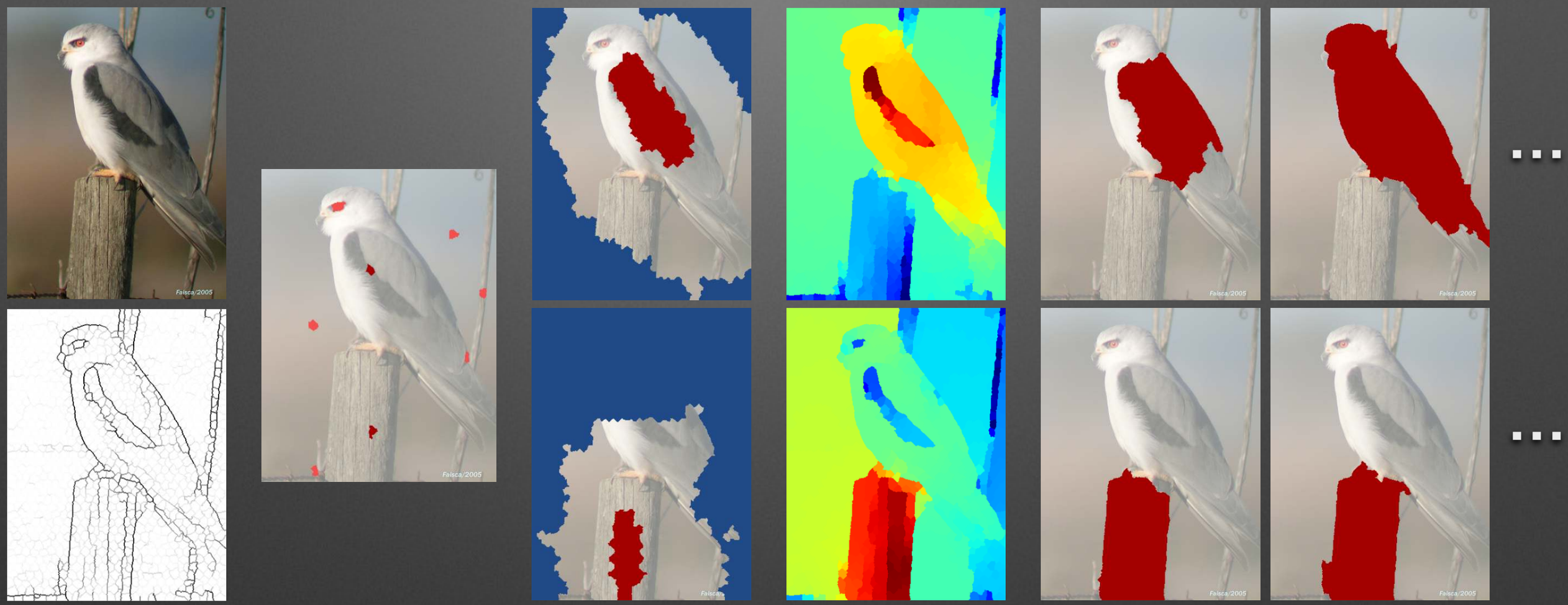
select seeds

foreground background masks

geodesic distance transform

multiple proposals per transform

Learned GOP



image, boundary map and superpixels

learned seeds

learned masks

geodesic distance transform

multiple proposals per transform

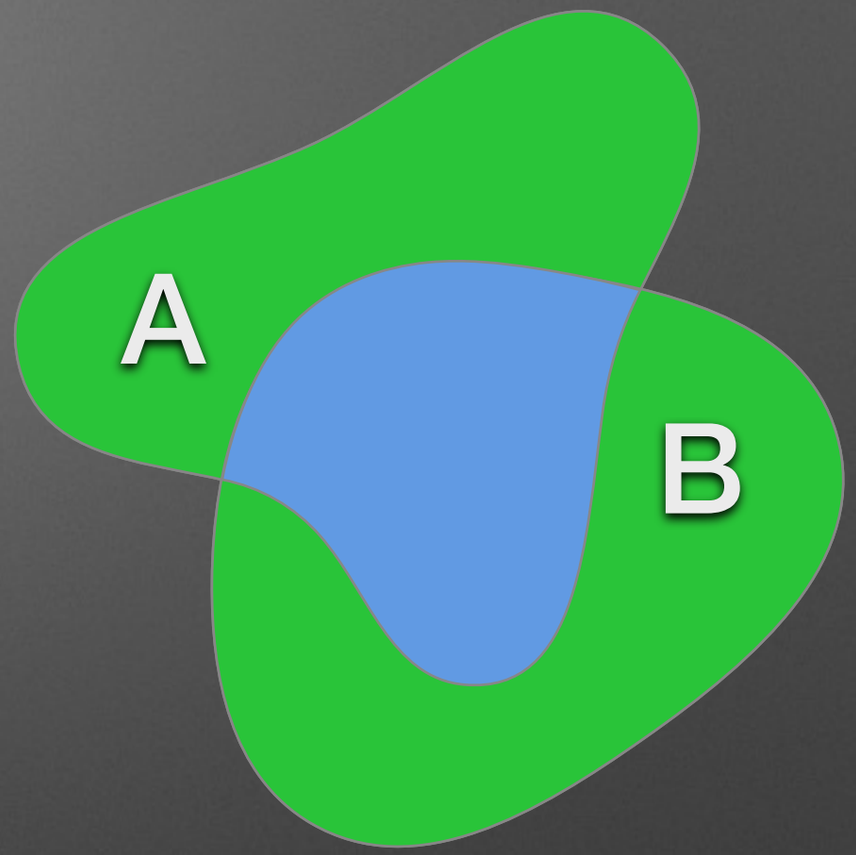
Results

- VOC 2012 dataset
- Evaluation metric
  - overlap $\mathcal{J}(A, B) = \frac{|A \cap B|}{|A \cup B|}$
  - best overlap $b(O_k) = \max_P \mathcal{J}(O_k, P)$
  - Average best overlap (ABO)

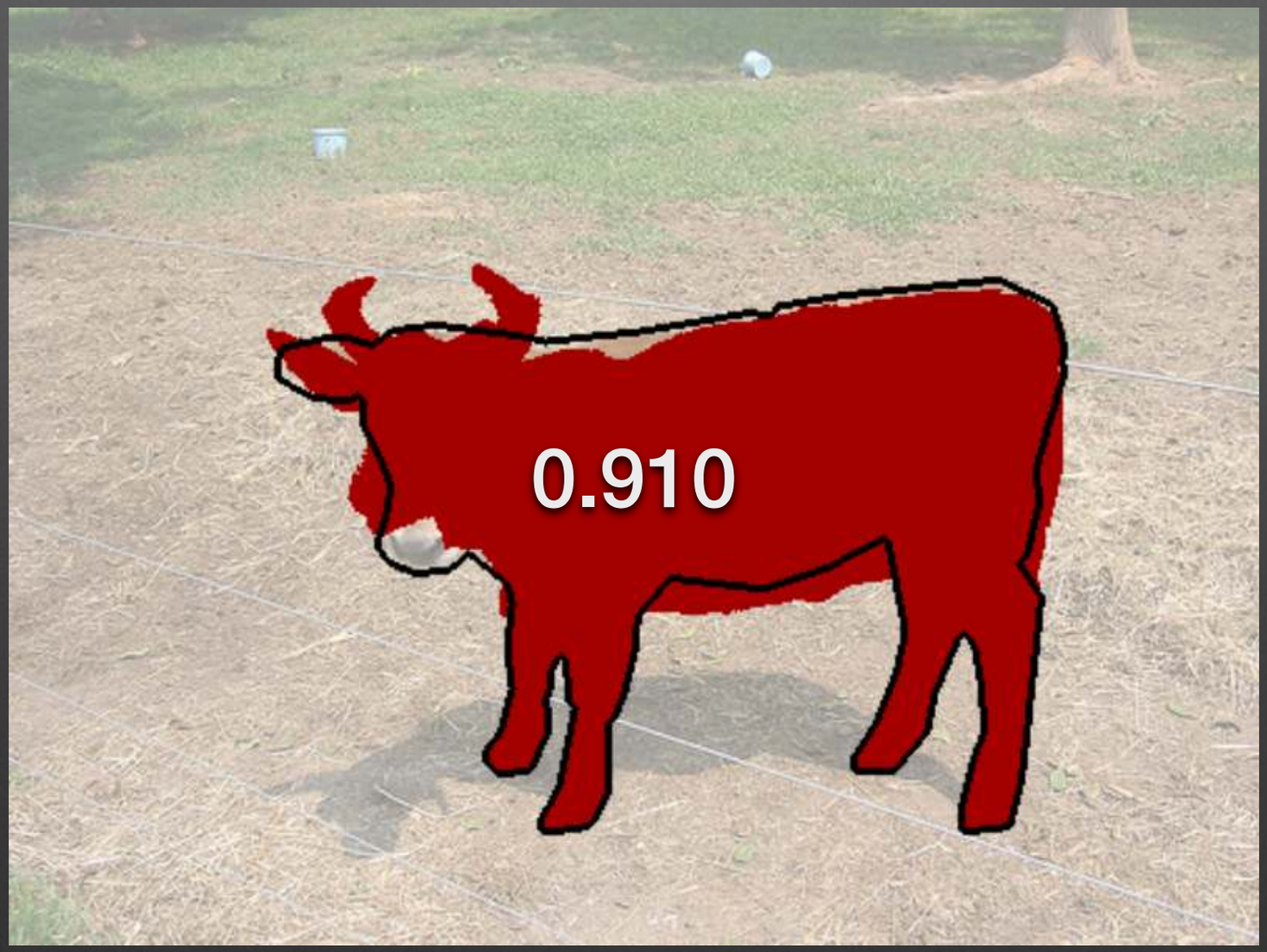
$$\frac{1}{N} \sum_k b(O_k)$$

  - α-recall

$$\frac{1}{N} \sum_k [b(O_k) > \alpha]$$



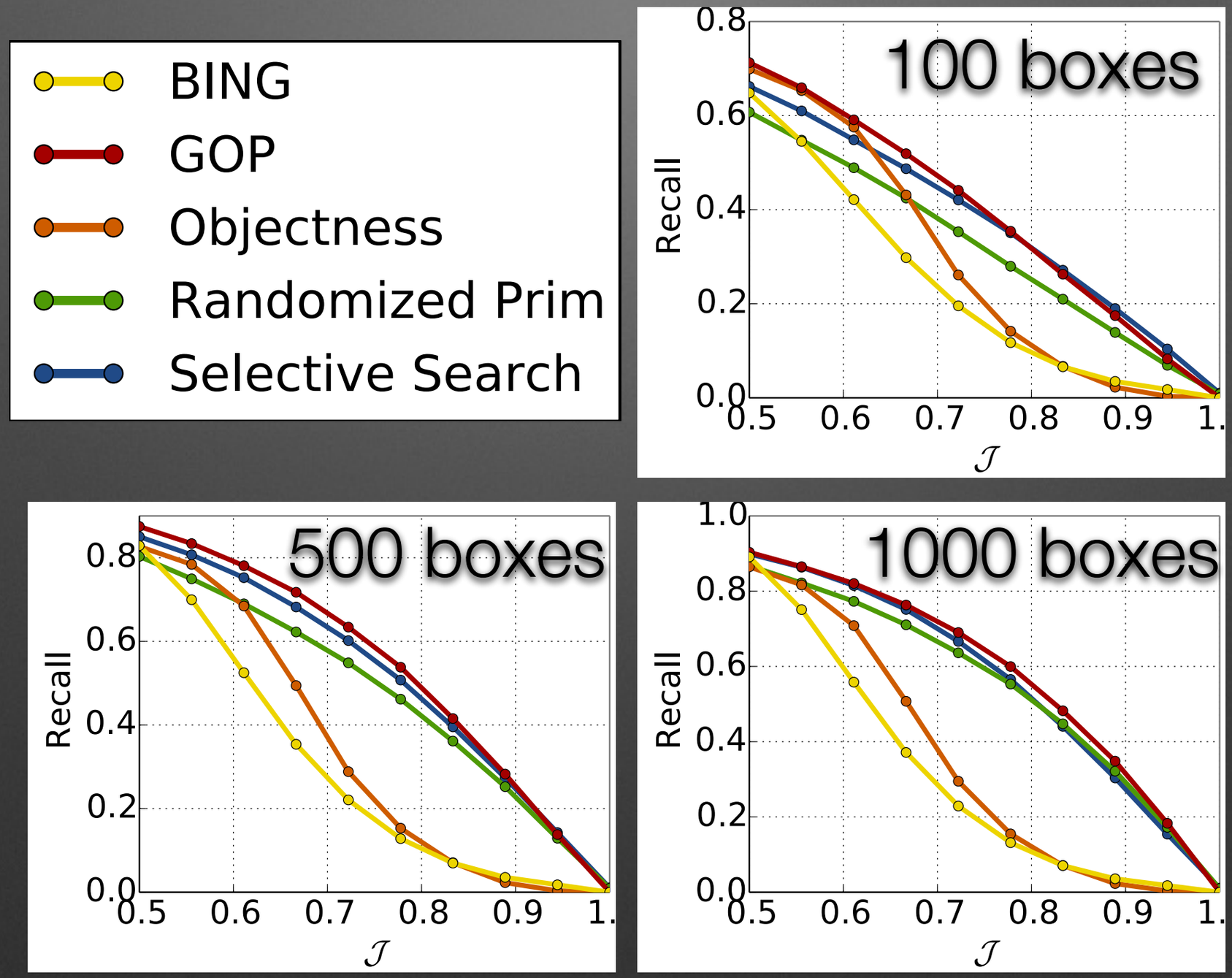
What does overlap mean?



Segmentation results

METHOD	# PROP.	ABO	50%-RECALL	70%-RECALL	TIME
CPMC	646	0.703	0.784	**0.609**	252s
Baseline GOP	653	0.712	0.833	**0.622**	0.6s
Learned GOP	652	0.720	0.844	**0.632**	1.0s
Cat-Ind OP	1536	0.718	0.820	**0.624**	119s
Baseline GOP	1090	0.727	0.847	**0.644**	0.65
Learned GOP	1199	0.741	0.865	**0.673**	1.1s
Sel Search	4374	0.735	0.891	**0.597**	2.6s
Baseline GOP	2089	0.744	0.867	**0.673**	0.9s
Learned GOP	2286	0.756	0.877	**0.699**	1.4s
Baseline GOP	3958	0.756	0.881	**0.699**	1.2s
Learned GOP	4186	0.766	0.889	**0.715**	1.7s

Bounding box results



Bounding box results

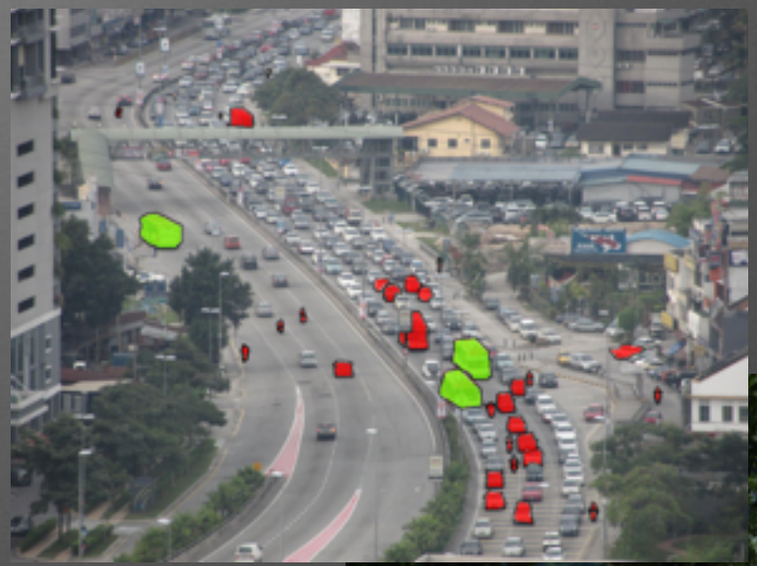
	VOLUME UNDER SURFACE (VUS)
BING	0.278
Objectness	0.324
Randomized Prim	0.511
Selective Search	0.528
GOP	0.546

COCO dataset - segments

METHOD	# PROP.	50%-REC.	70%-REC.
Sel Search	6504	0.717	**0.369**
Baseline GOP	6106	0.704	**0.426**
Learned GOP	6264	0.717	**0.447**

LARGE OBJECTS ≥ 25			
Sel Search	6504	0.810	**0.442**
Baseline GOP	6106	0.882	**0.582**
Learned GOP	6264	0.891	**0.609**

SMALL OBJECTS < 25			
Sel Search	6504	0.525	**0.219**
Baseline GOP	6106	0.337	**0.106**
Learned GOP	6264	0.356	**0.112**



Summary

- Geodesic Object Proposals
  - fast
  - good segment proposals
  - good bounding box proposals
- Future work
  - small objects
  - learn proposals directly from data

Questions

C++, Python and Matlab Code:
http://www.philkr.net/home/gop