Improving Spatial Support for Objects via Multiple Segmentations

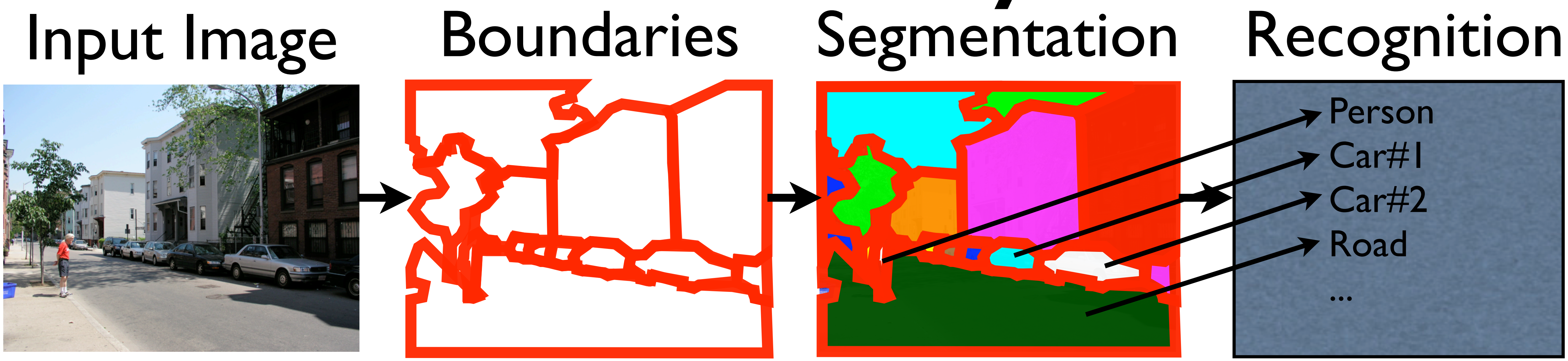
presented at
British Machine Vision Conference 2007
September 12, 2007

Tomasz Malisiewicz and Alexei A. Efros
The Robotics Institute
Carnegie Mellon University

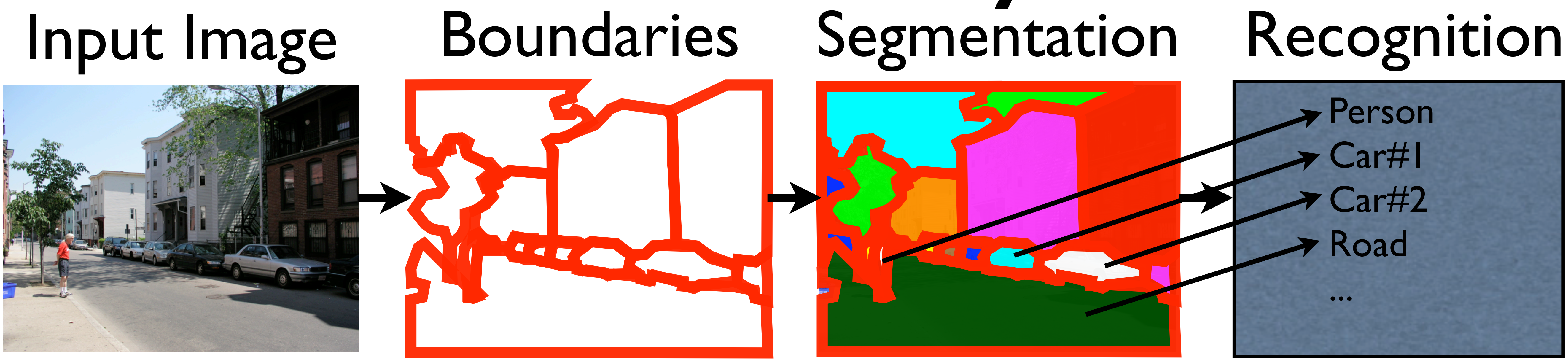
THE ROBOTICS INSTITUTE

Carnegie Mellon

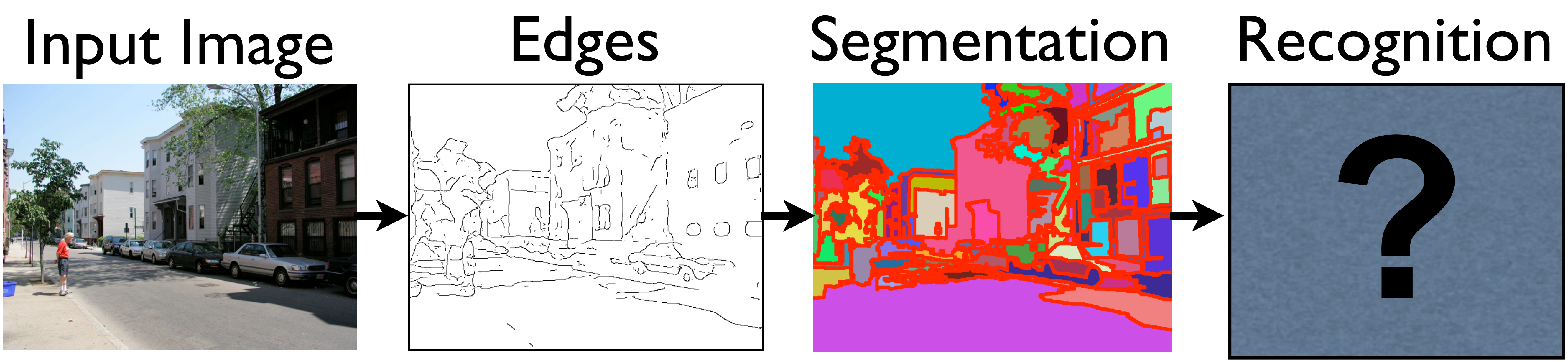
in Theory


Input Image
Boundaries
Segmentation
Recognition
Person
Car#1
Car#2
Road
...

in Theory



in Practice



Sliding Windows

Sliding Windows

Person

Sliding Windows

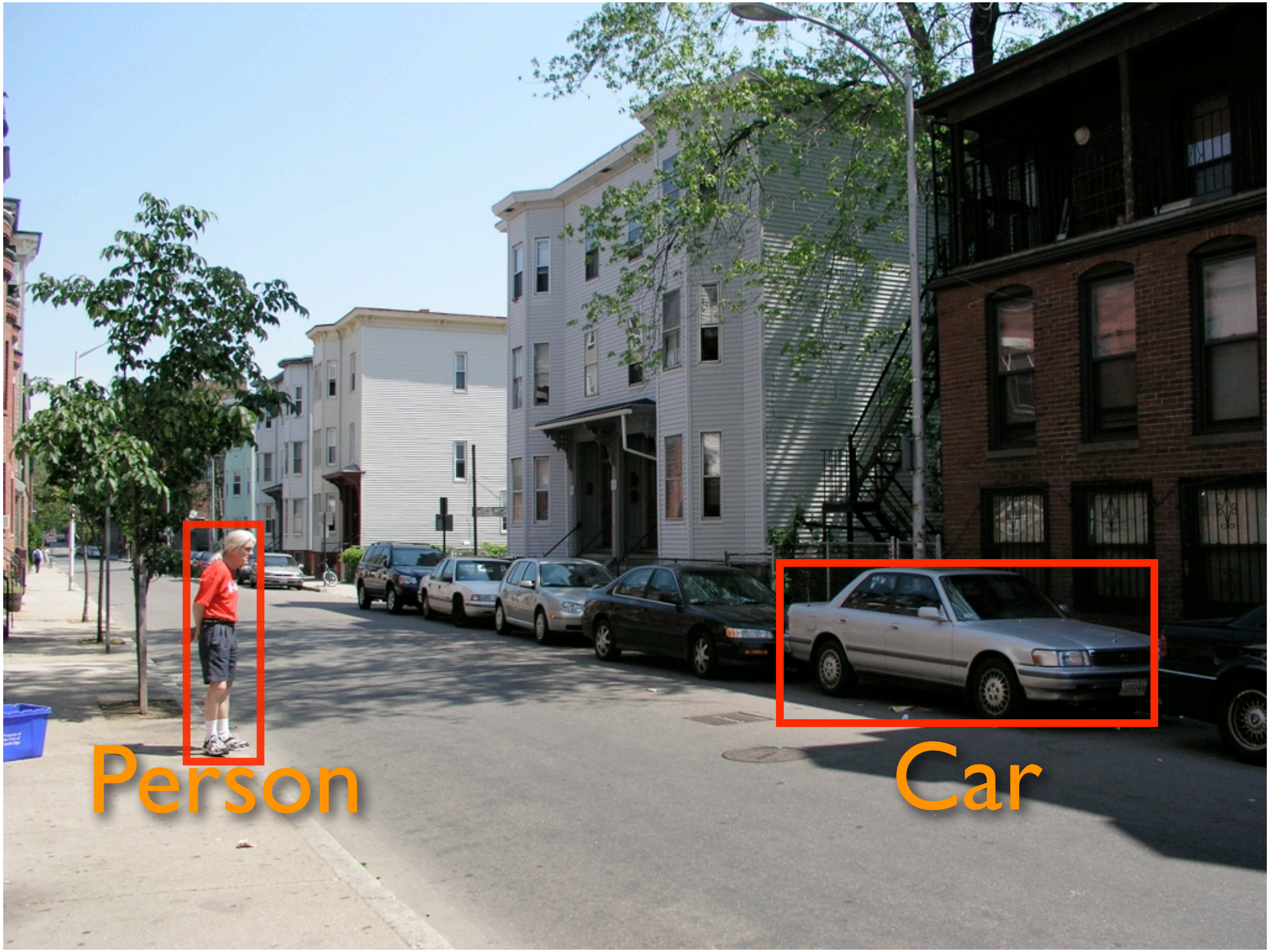
Person

Sliding Windows

Sliding Windows

Sliding Windows

Sliding Windows



Successes of Sliding Windows

cars

Schneiderman & Kanade '00

faces

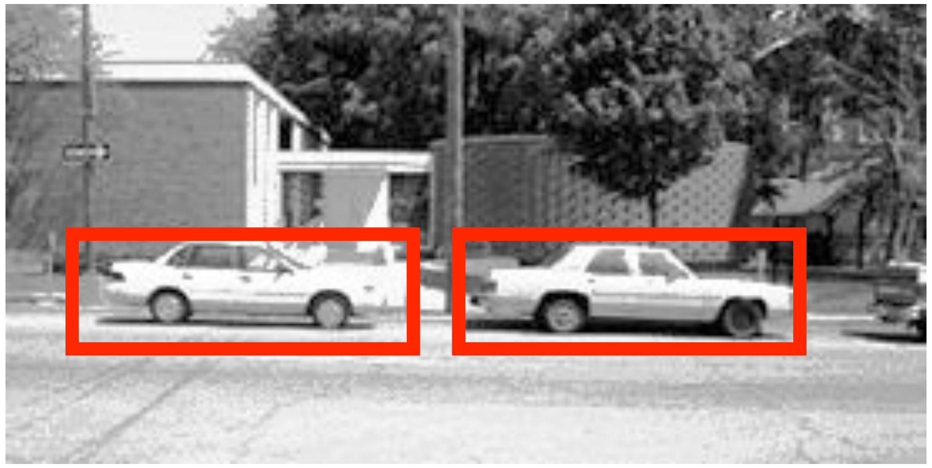
Viola & Jones '04
Schneiderman & Kanade '00

pedestrians

Dalal & Triggs '05
Ferrari et al '07

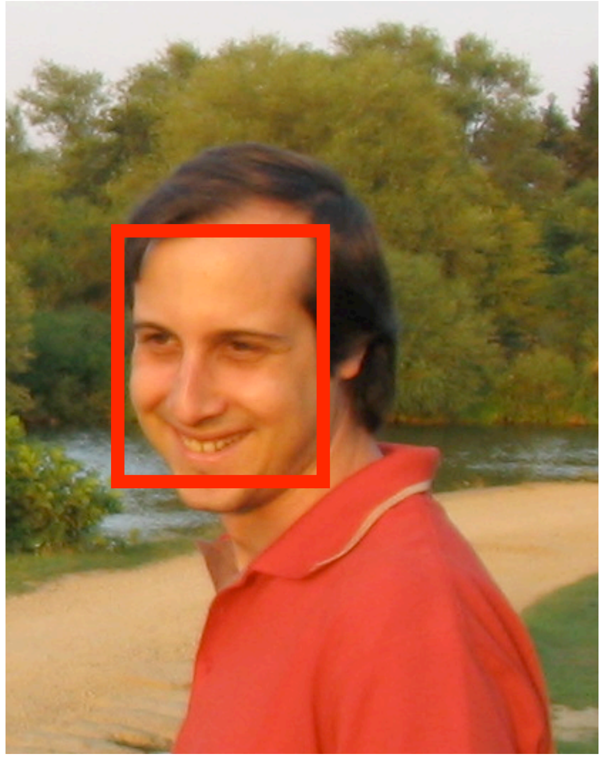
Successes of Sliding Windows

cars



Schneiderman & Kanade '00

faces



Viola & Jones '04
Schneiderman & Kanade '00

pedestrians

Dalal & Triggs '05
Ferrari et al '07

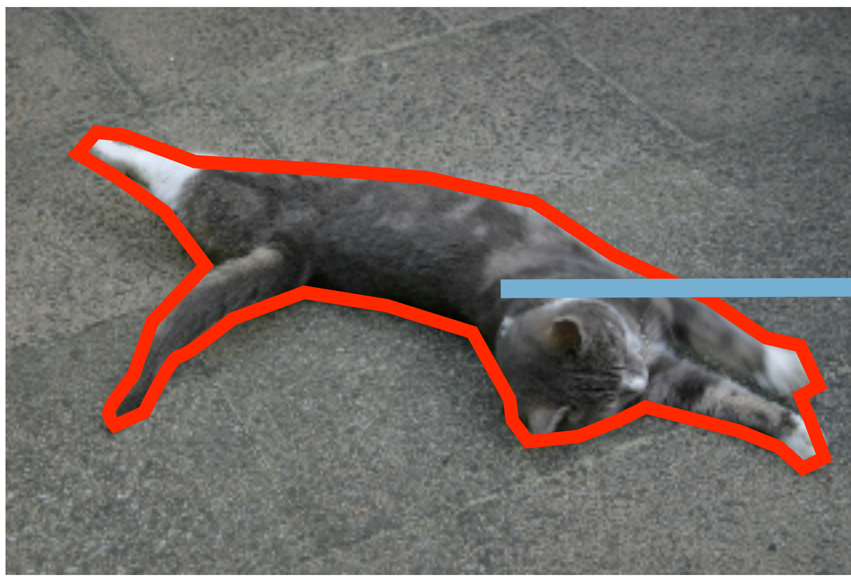




Overview

- Does spatial support matter?
- How to get good spatial support?

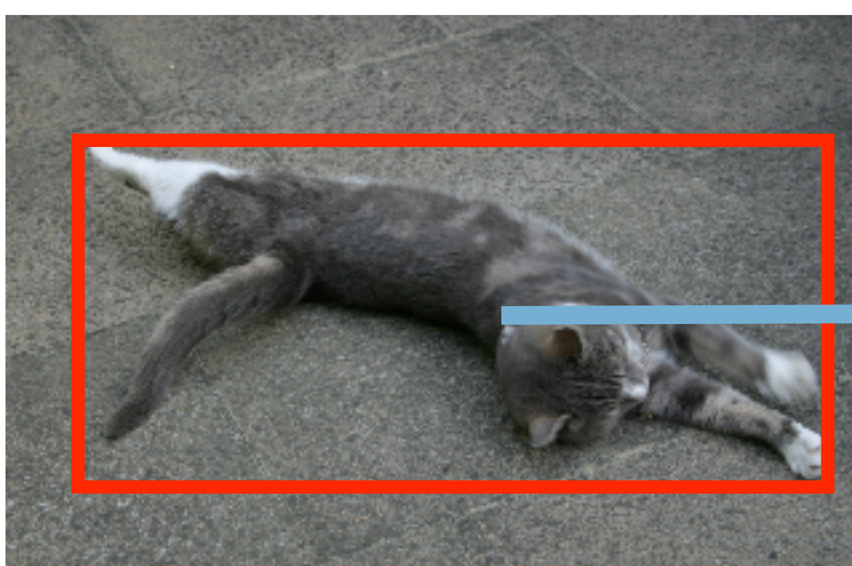
1. Does Spatial Support Matter?



Classify

Ground-Truth Segment

vs.

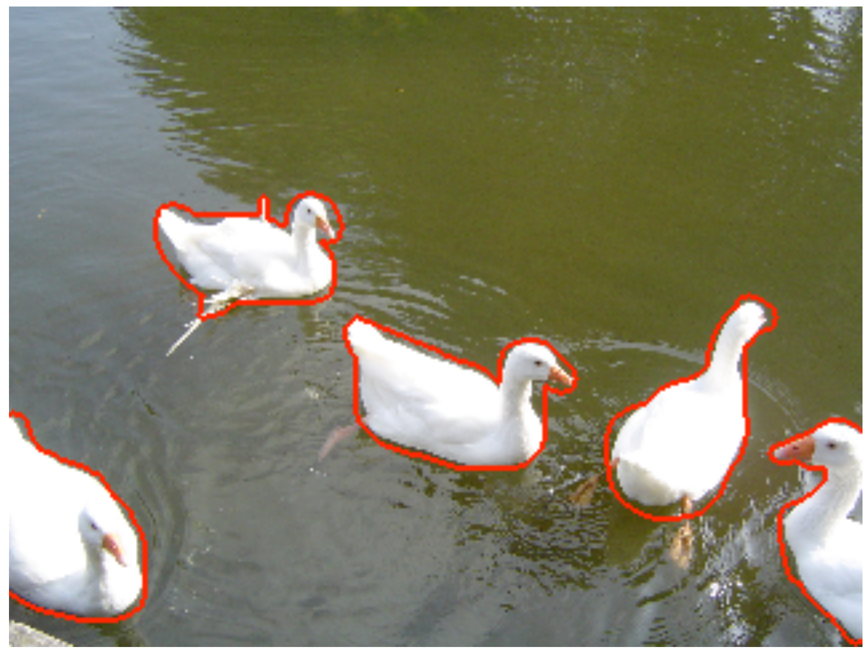
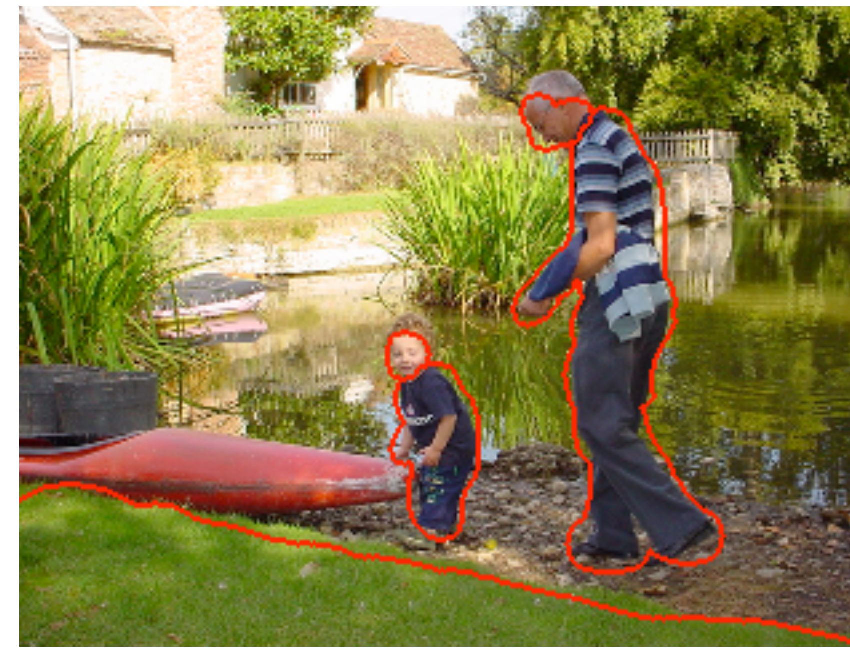


Classify

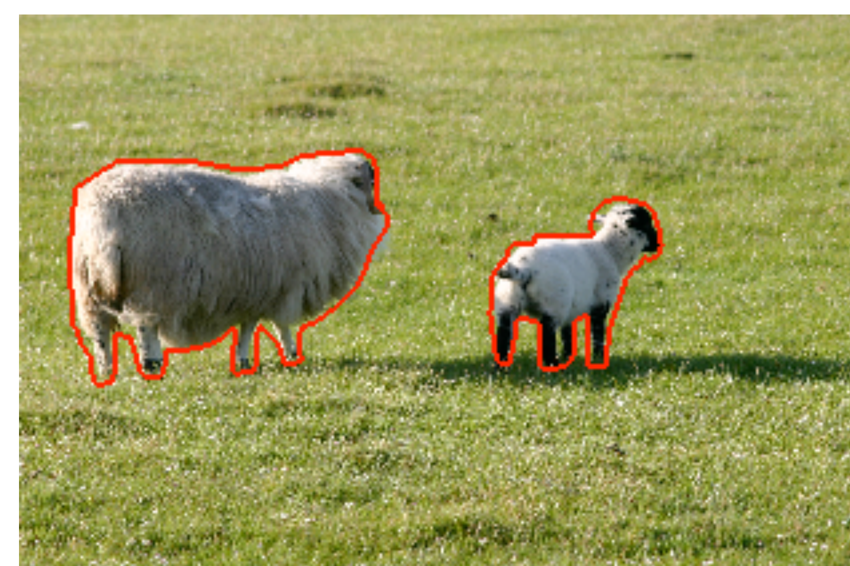
Bounding Box

Does Spatial Support Matter?

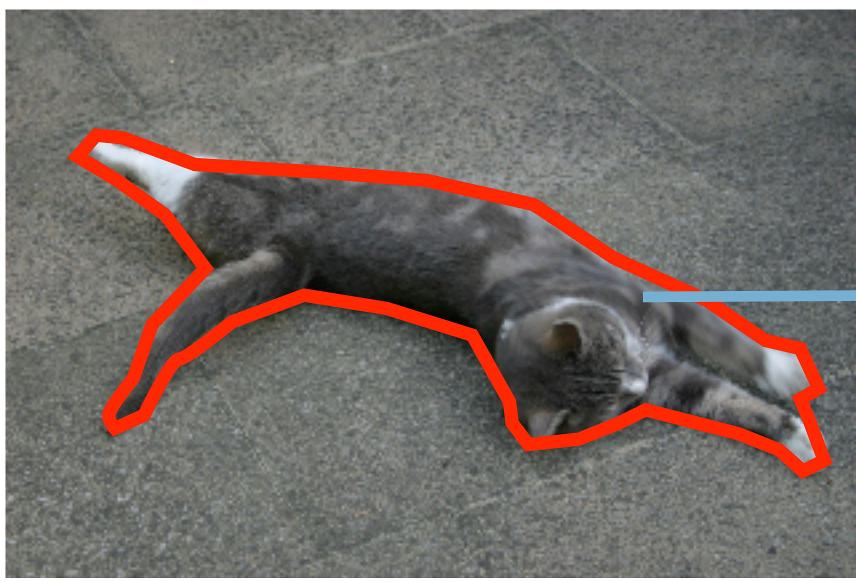
MSRC data-set: 591 images of 23 object classes + pixel-wise segmentation masks





Does Spatial Support Matter?



Features

*

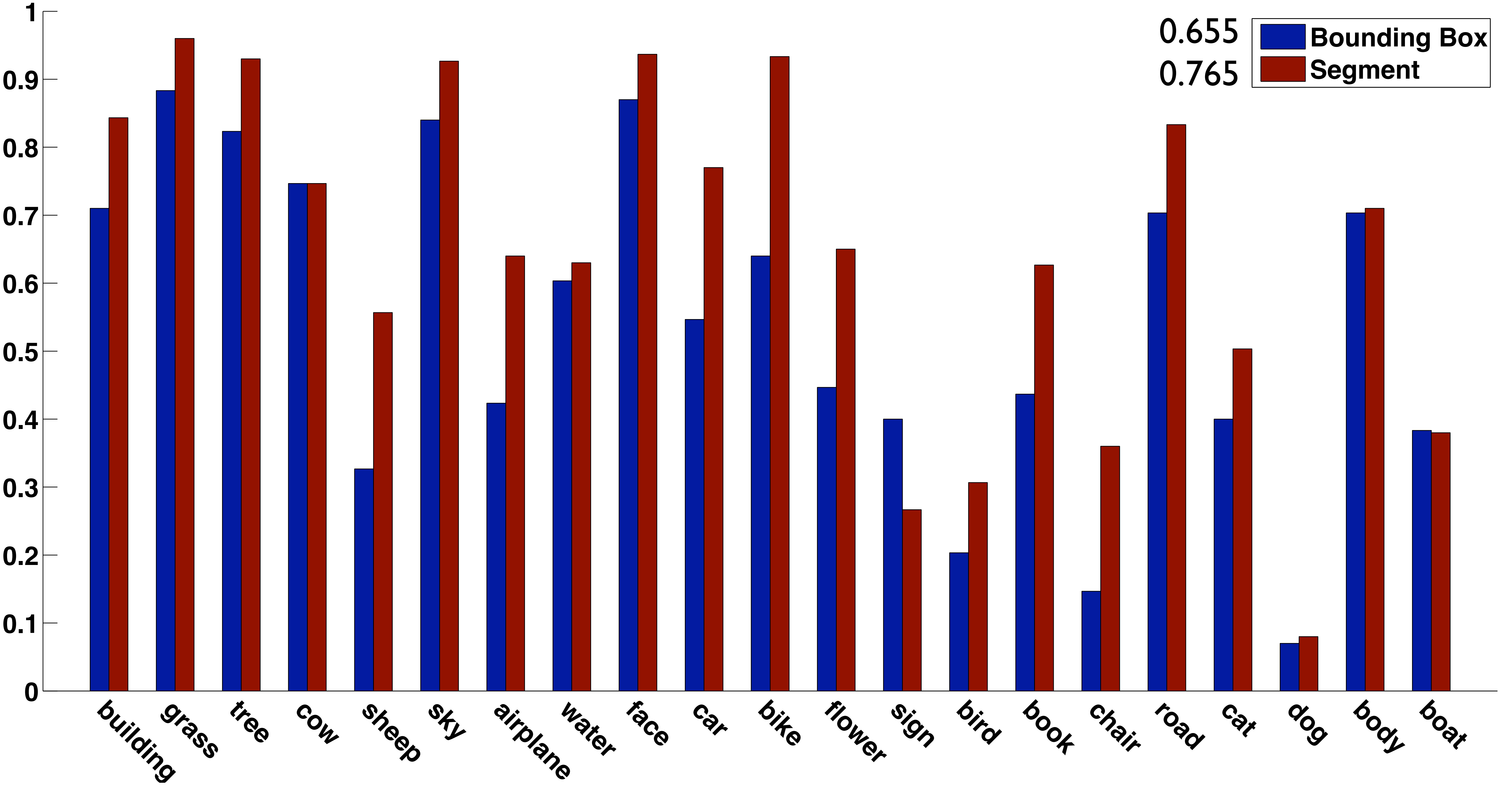
Feature Descriptions	Num
Color	**16**
C1. RGB values: mean	3
C2. HSV values: C1 in HSV space	3
C3. Hue: histogram (5 bins) and entropy	6
C4. Saturation: histogram (3 bins) and entropy	4
Texture	**15**
T1. DOOG filters: mean abs response of 12 filters	12
T2. DOOG stats: mean of variables in T1	1
T3. DOOG stats: argmax of variables in T1	1
T4. DOOG stats: (max - median) of variables in T1	1
Location and Shape	**12**
L1. Location: normalized x and y, mean	2
L2. Location: norm. x and y, 10^{th} and 90^{th} pctl	4
L3. Location: norm. y wrt horizon, 10^{th}, 90^{th} pctl	2
L4. Shape: number of superpixels in region	1
L5. Shape: number of sides of convex hull	1
L6. Shape: $num\ pixels/area(convex\ hull)$	1
L7. Shape: whether the region is contiguous $\in \{0, 1\}$	1
3D Geometry	**35**
G1. Long Lines: total number in region	1
G2. Long Lines: % of nearly parallel pairs of lines	1
G3. Line Intsctn: hist. over 12 orientations, entropy	13
G4. Line Intsctn: % right of center	1
G5. Line Intsctn: % above center	1
G6. Line Intsctn: % far from center at 8 orientations	8
G7. Line Intsctn: % very far from center at 8 orient.	8
G8. Texture gradient: x and y "edginess" (T2) center	2

Classifier

Boosted Decision Tree*

*Hoiem *et al* '05

Does Spatial Support Matter?



2. How to get good spatial support?

2. How to get good spatial support?

- Segmentation is a natural way to obtain spatial support

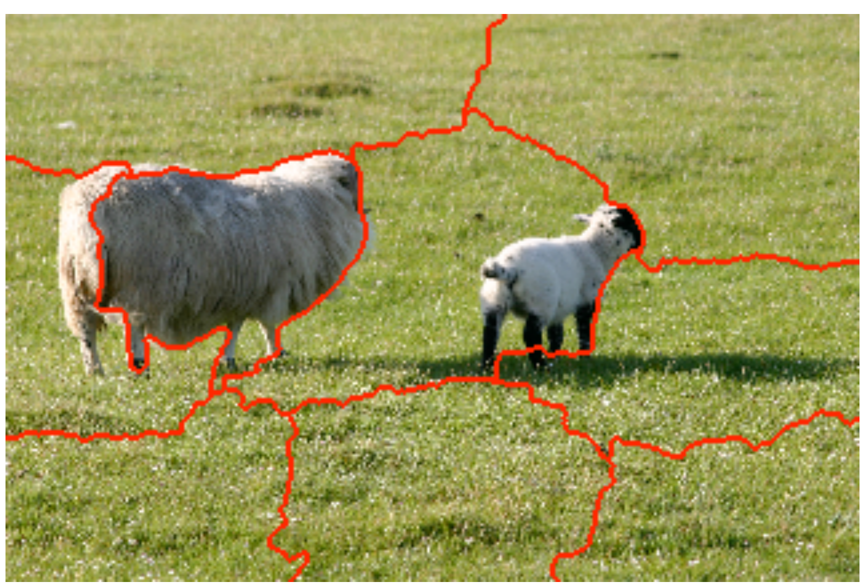
2. How to get good spatial support?

- Segmentation is a natural way to obtain spatial support
- Can an off-the-shelf segmentation algorithm provide good spatial support?

2. How to get good spatial support?

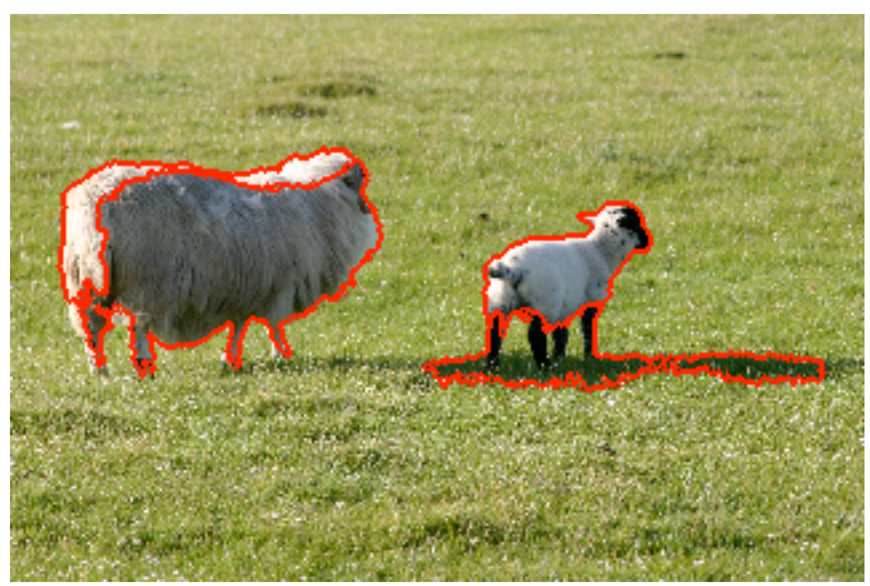
- Segmentation is a natural way to obtain spatial support
- Can an off-the-shelf segmentation algorithm provide good spatial support?

Normalized Cuts



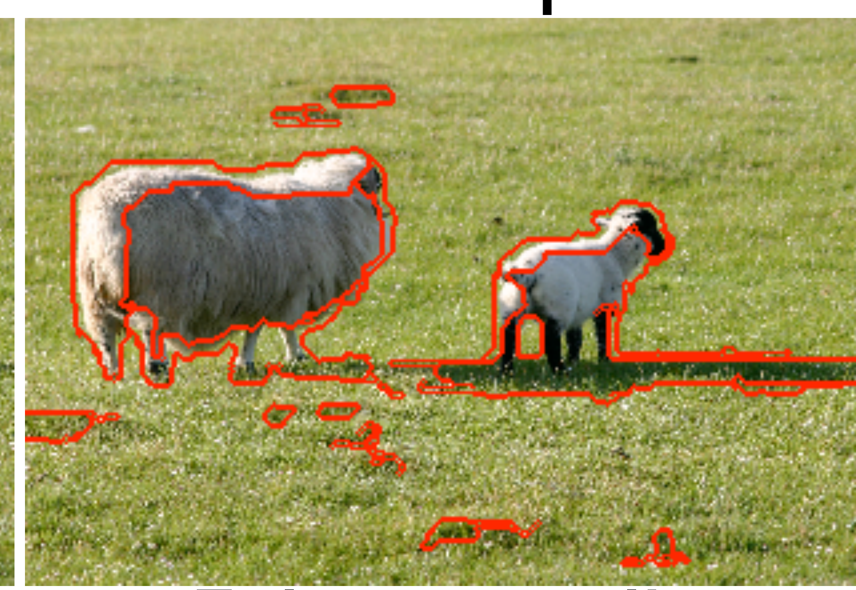
Shi & Malik

Mean Shift



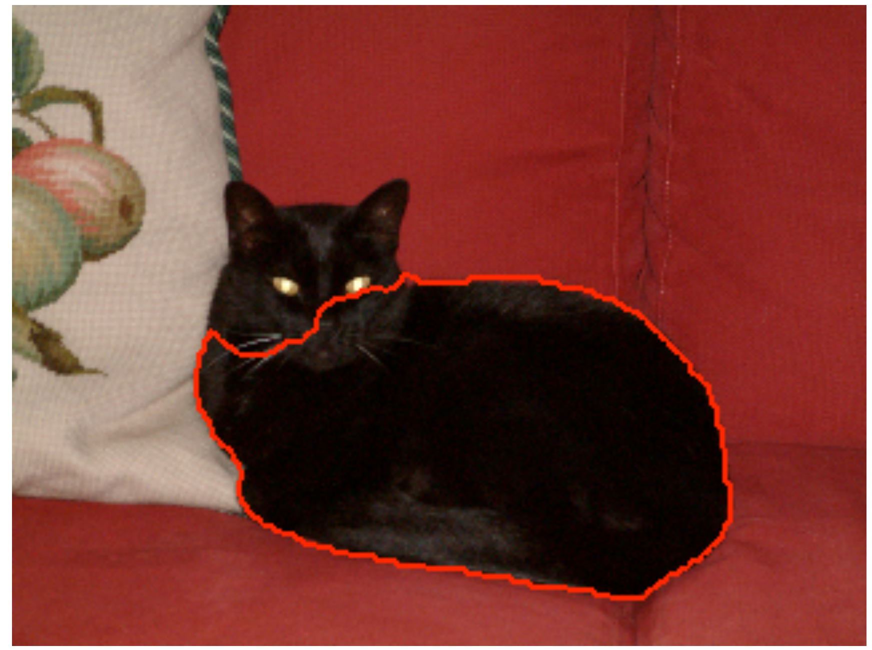
Comaniciu & Meer

Efficient Graph Based



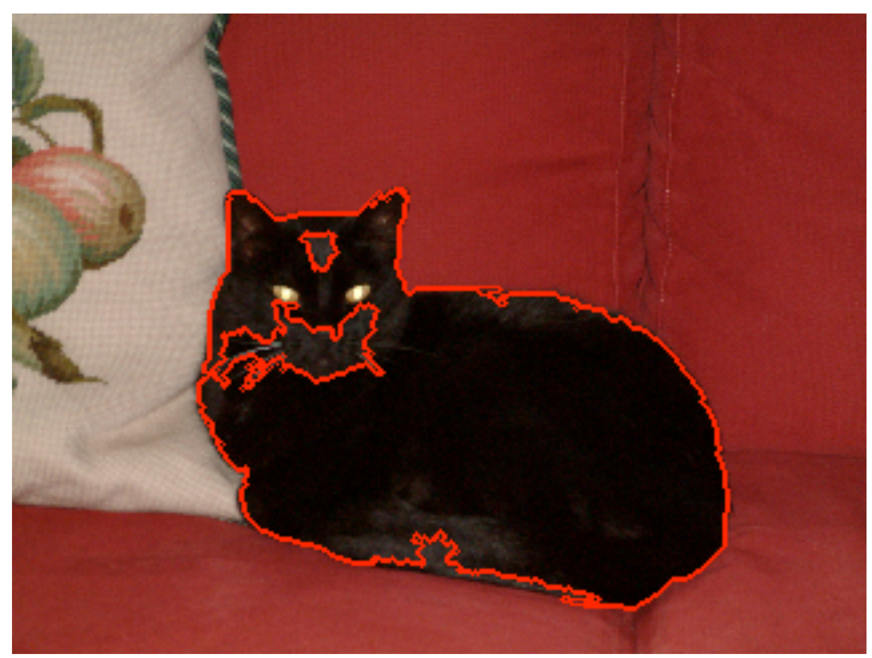
Felzenszwalb & Huttenlocher

Spatial Support



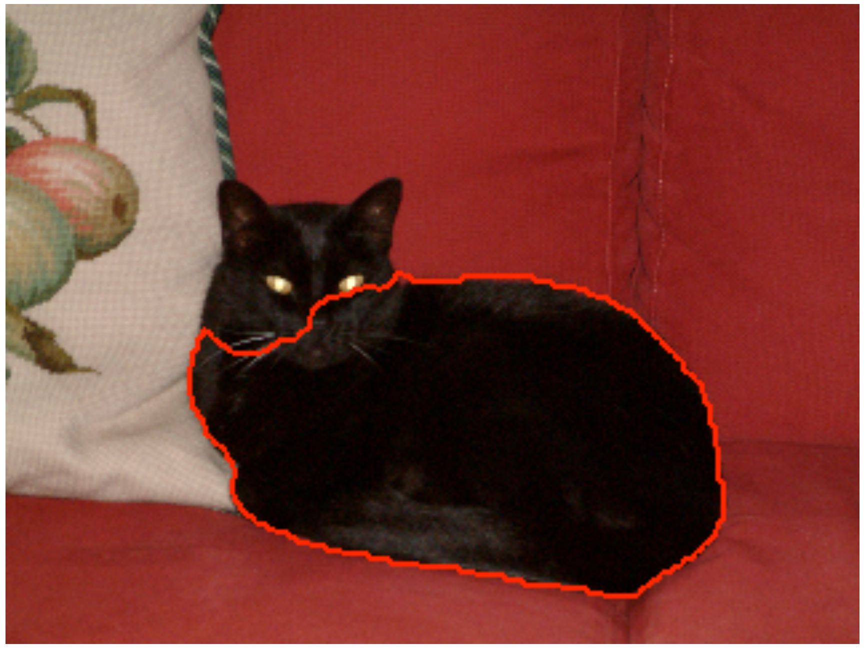
Segment #1

Ground Truth



Segment #2

Spatial Support



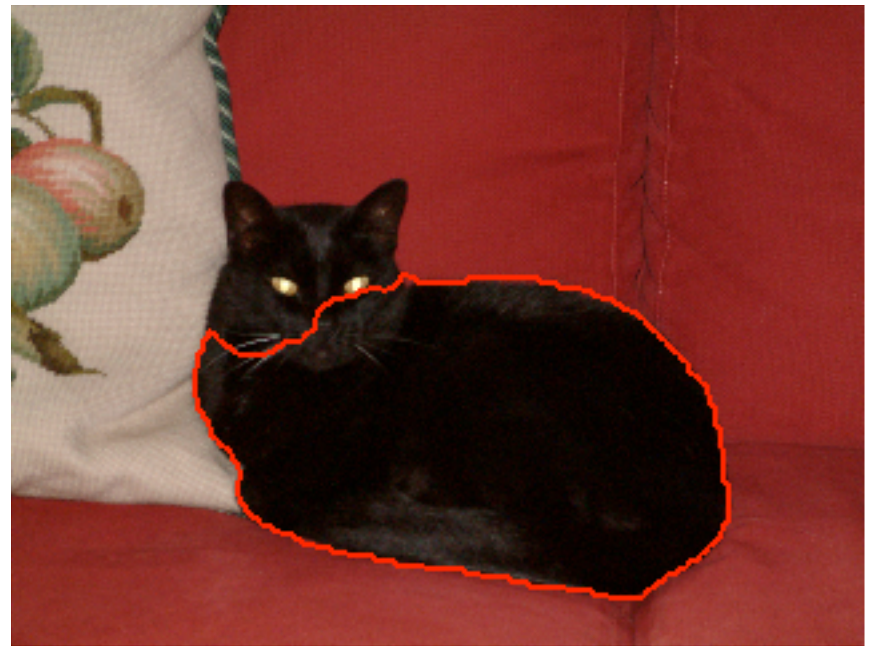
Segment #1

Segment #2

Ground Truth

$$OS(S,G) = \frac{|S \cap G|}{|S \cup G|}$$

Spatial Support



.825

Segment #1

$$OS(S,G) = \frac{|S \cap G|}{|S \cup G|}$$

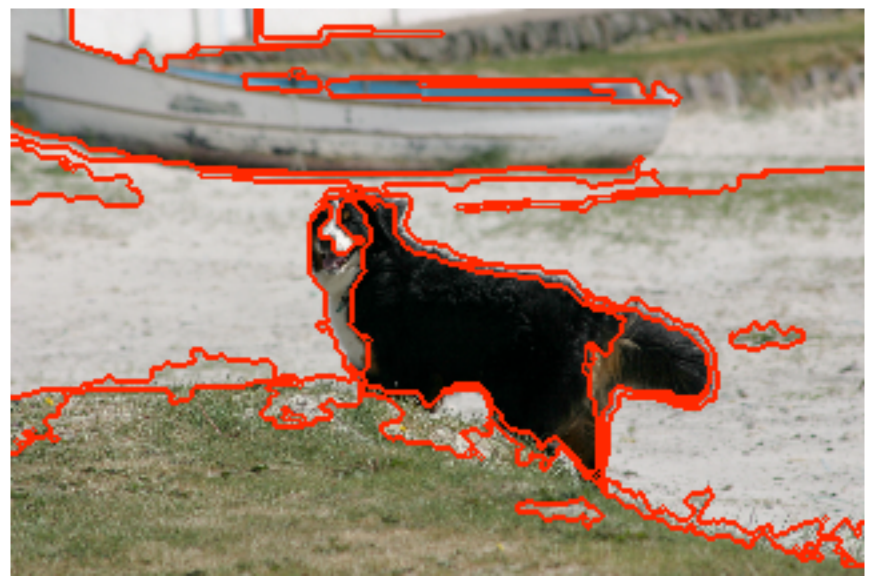
Ground Truth

.892

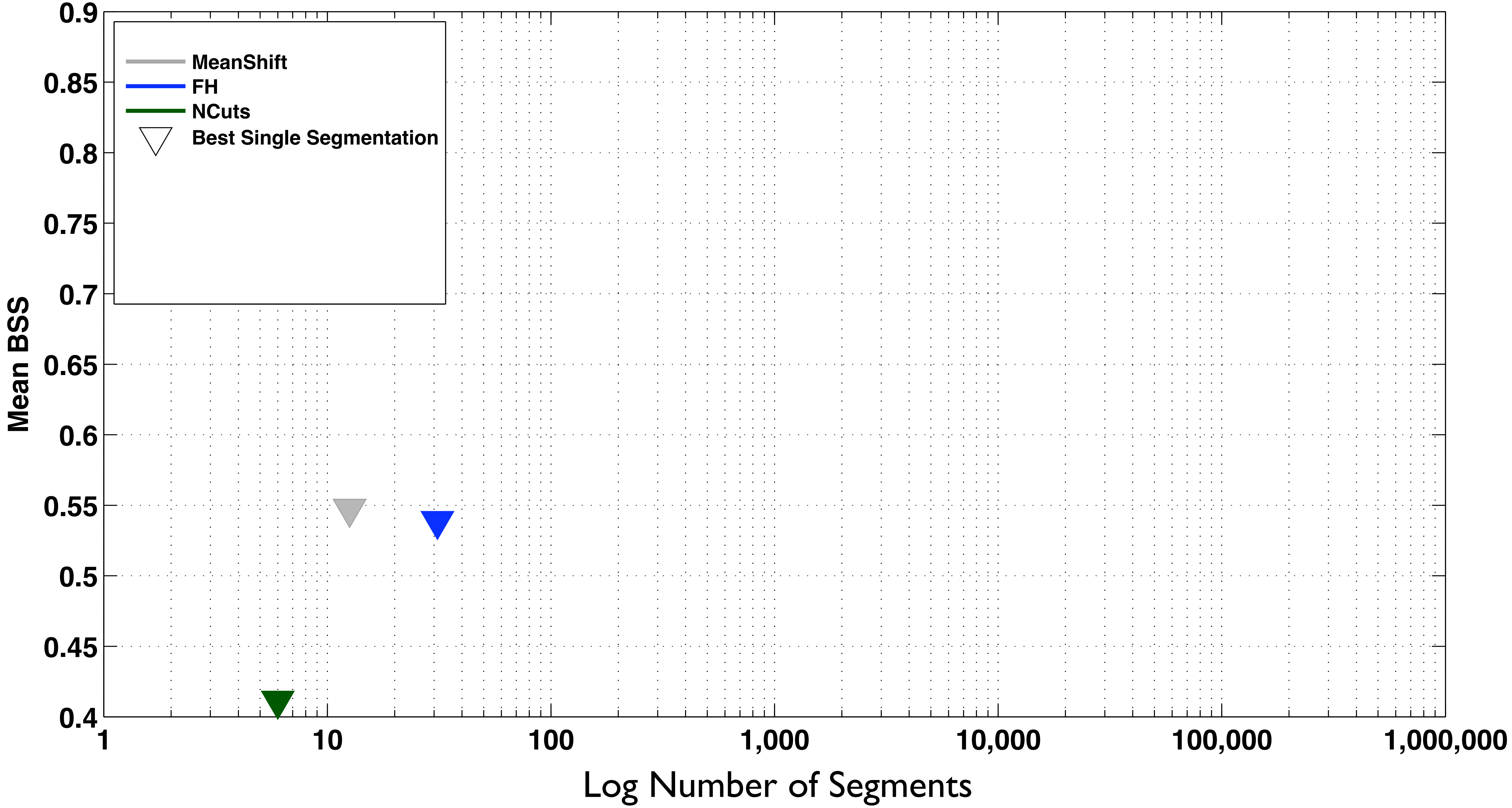
Segment #2

Ground Truth

Mean Shift	FH	NCuts

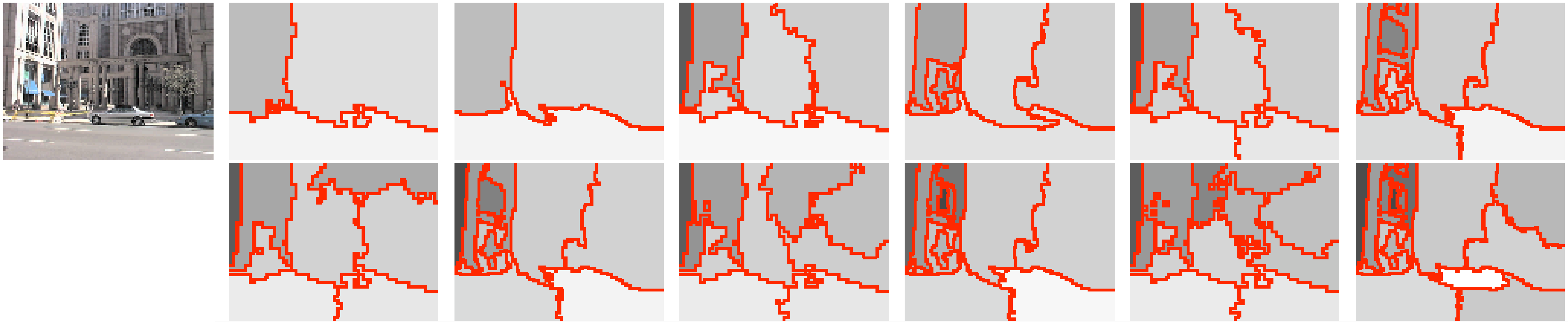


Evaluation*

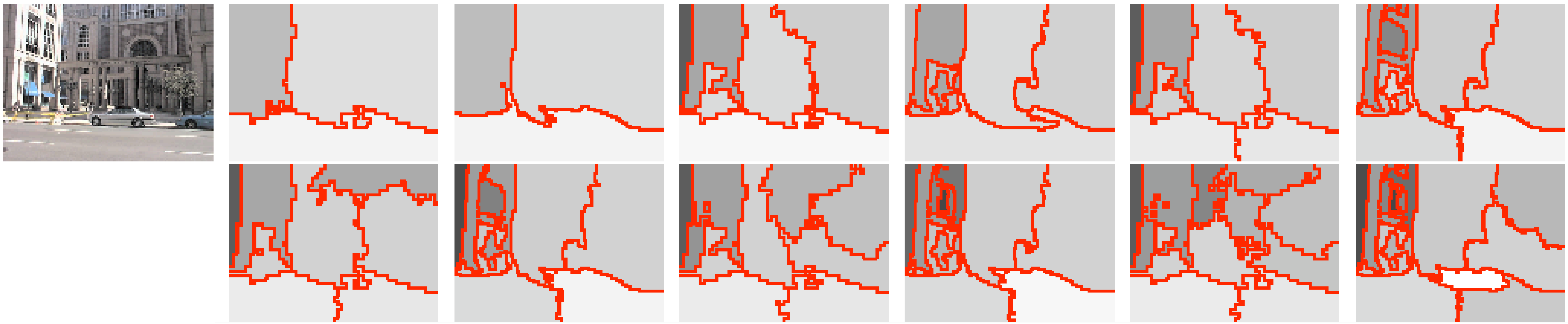


*Unnikrishnan *et al* 2005, Ge *et al* 2006

The problem with segmentation

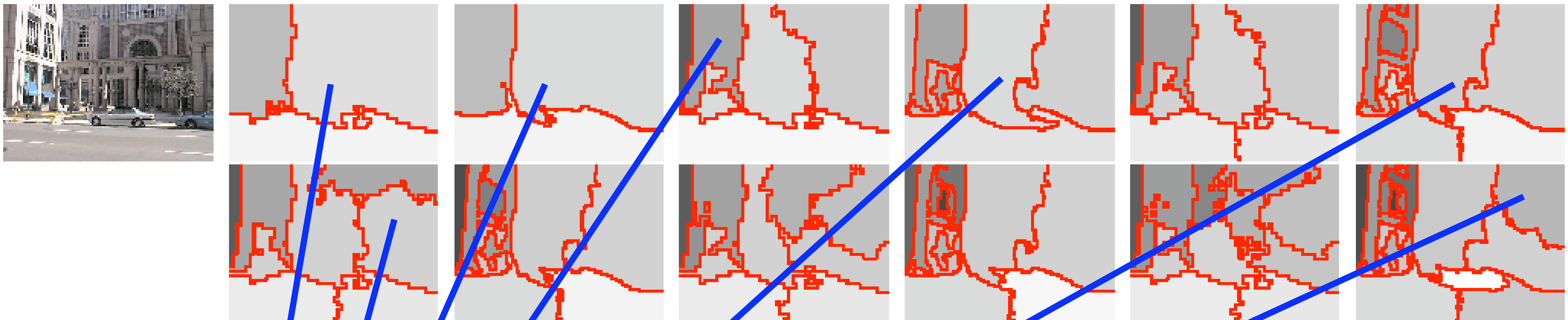


The problem with segmentation

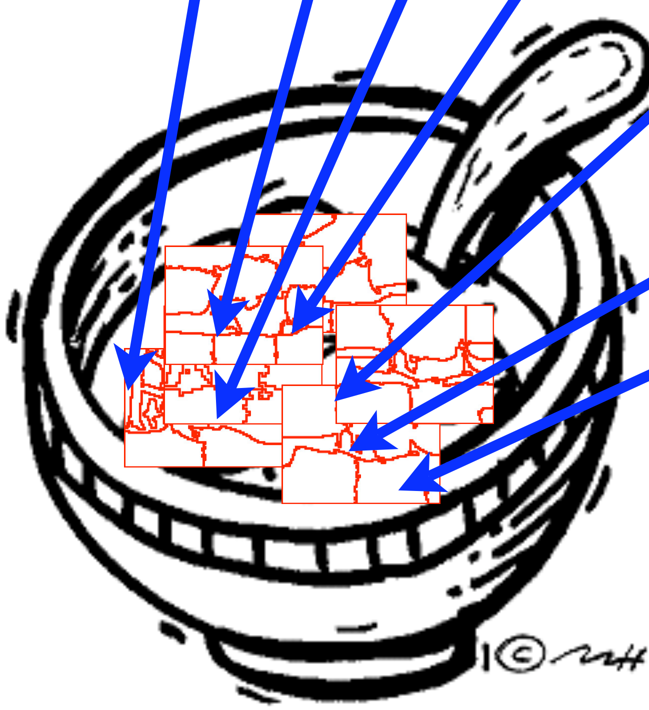


No Single Segmentation provides adequate spatial support

The problem with segmentation



No Single Segmentation provides adequate spatial support



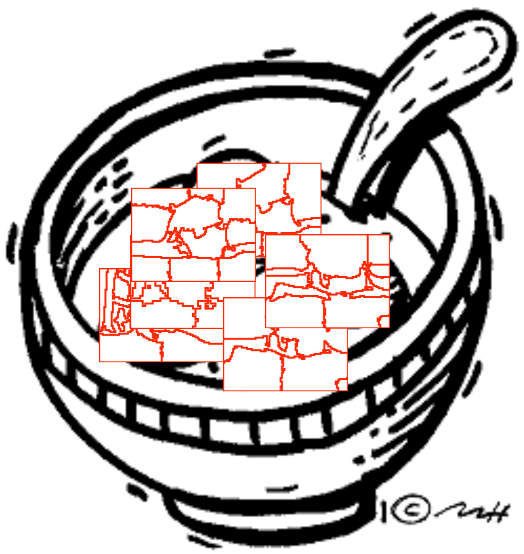
Use a Soup of Segments
(Hoiem *et al* 2005, Russell *et al* 2006)

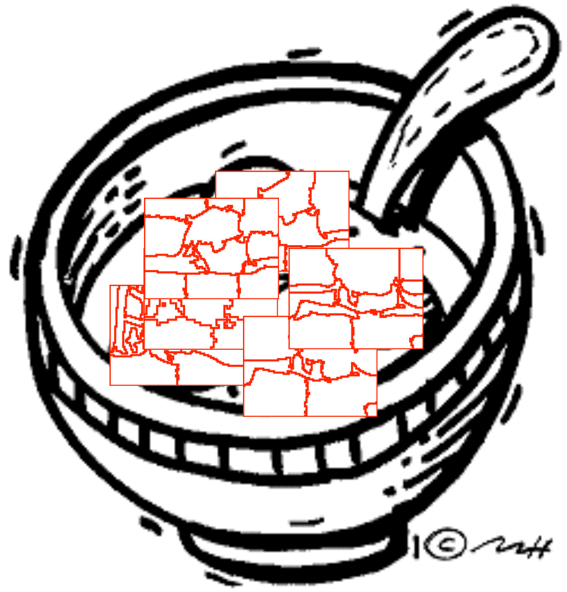
Ground Truth

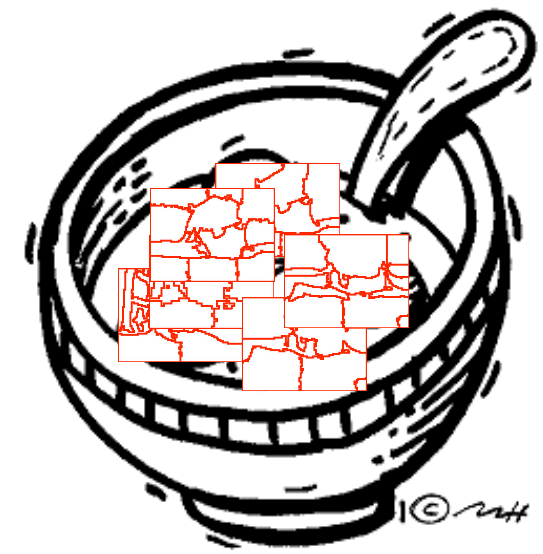
Mean Shift (33)

FH (24)

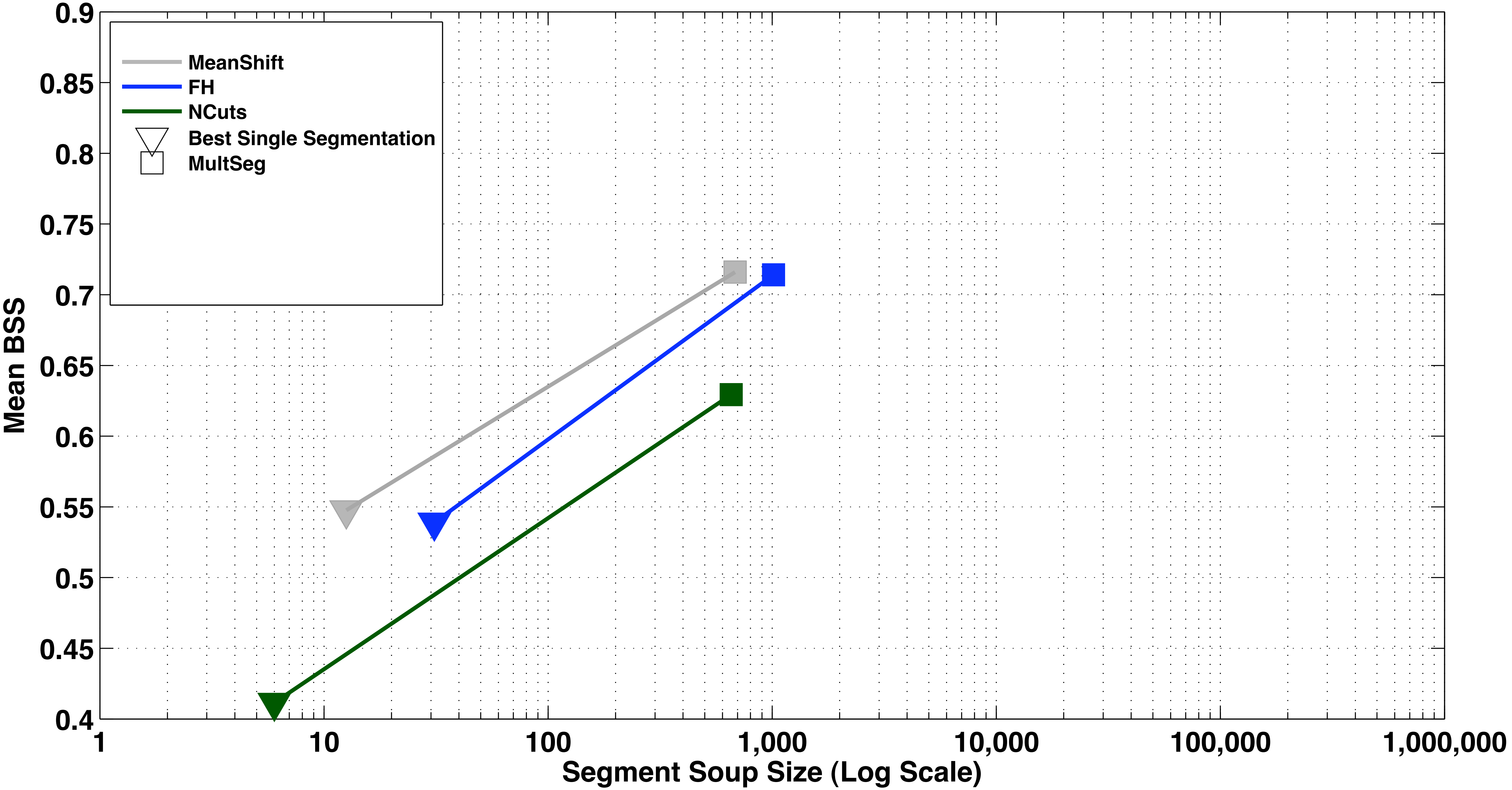
NCuts (33)







Quantitative Results



A closer look

A closer look

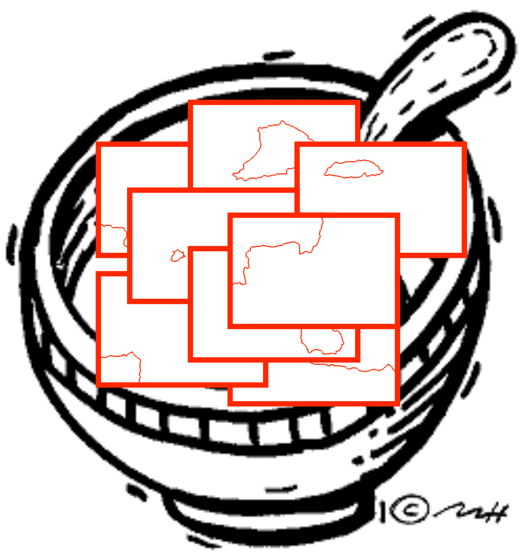
Merging Segments

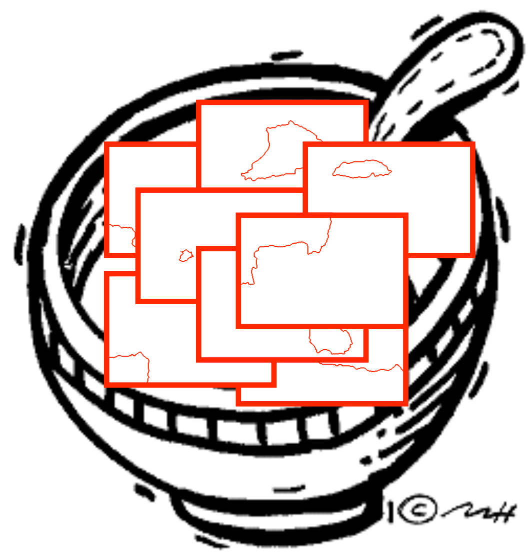
- Enumerate all pairs/triplets of adjacent segments
- Inexpensive and fast given an adjacency graph

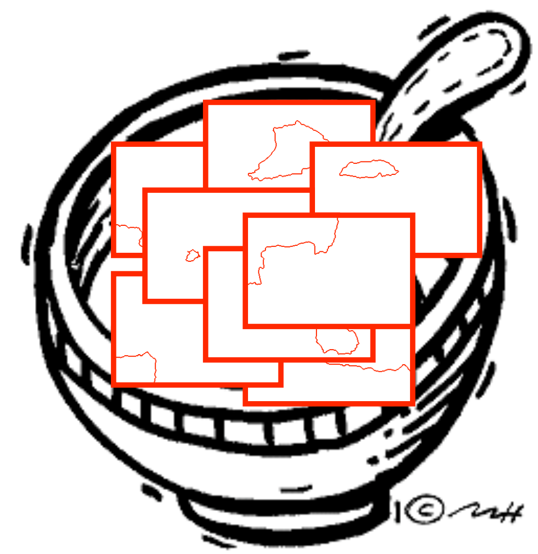
Mean Shift

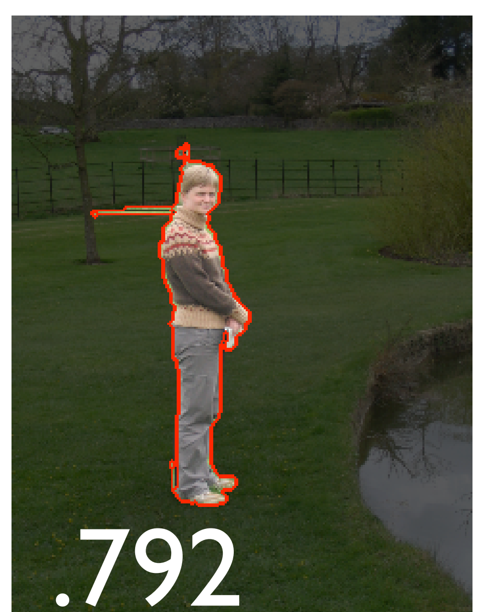
FH

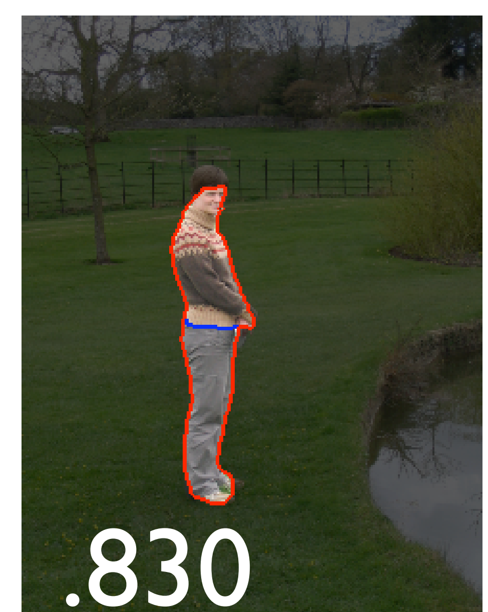
NCuts



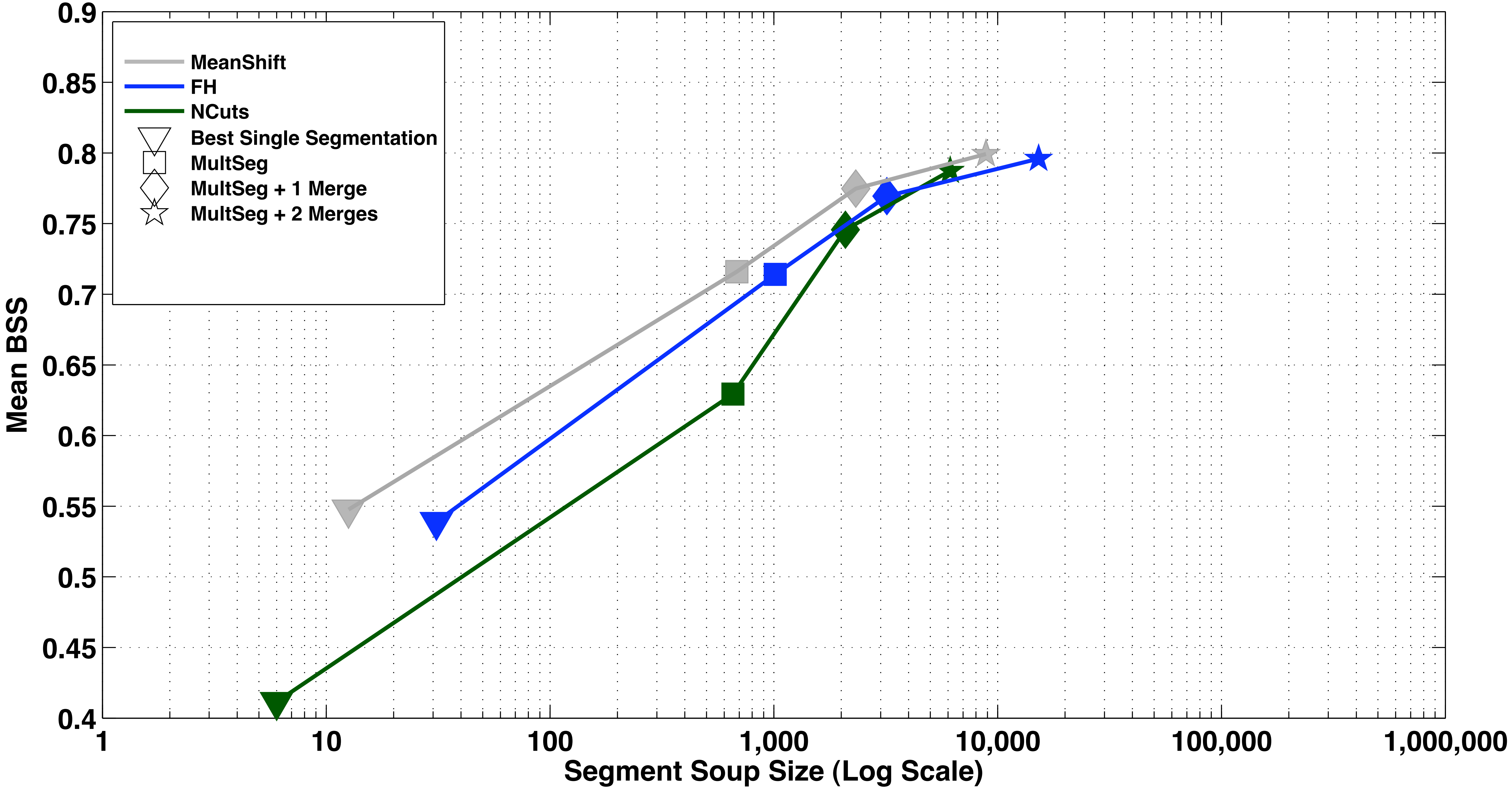






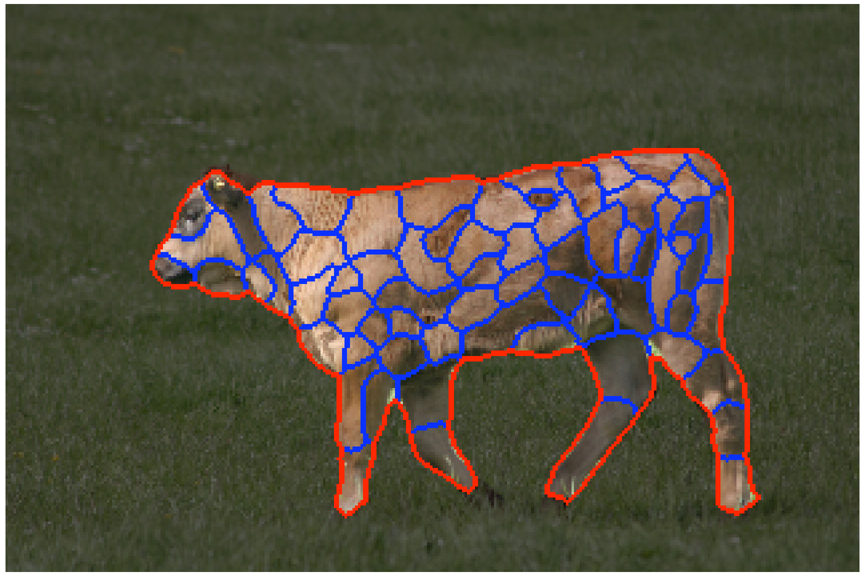


Quantitative Results



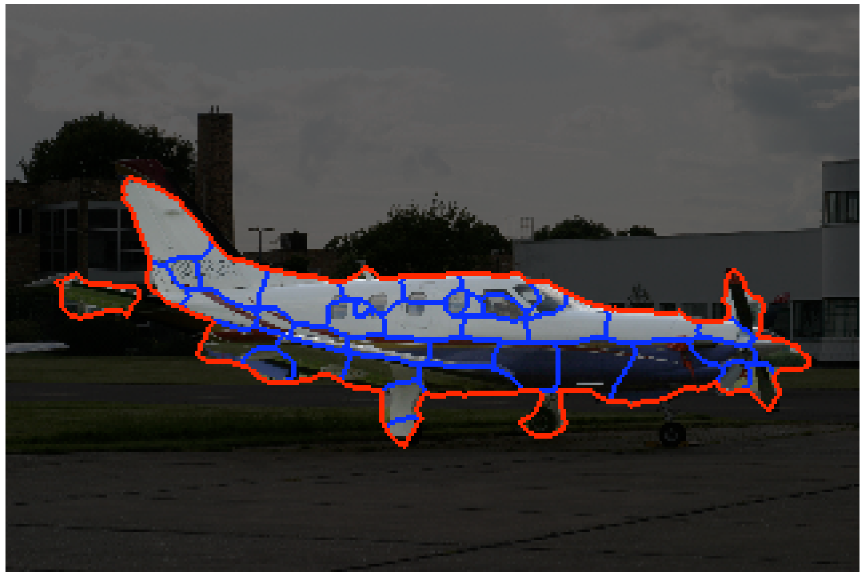
Upper-Bound: Superpixels

- Create superpixels with NCuts and K=200 (Ren & Malik 2003)
- Consider all merges of superpixels
- Infeasible in practice



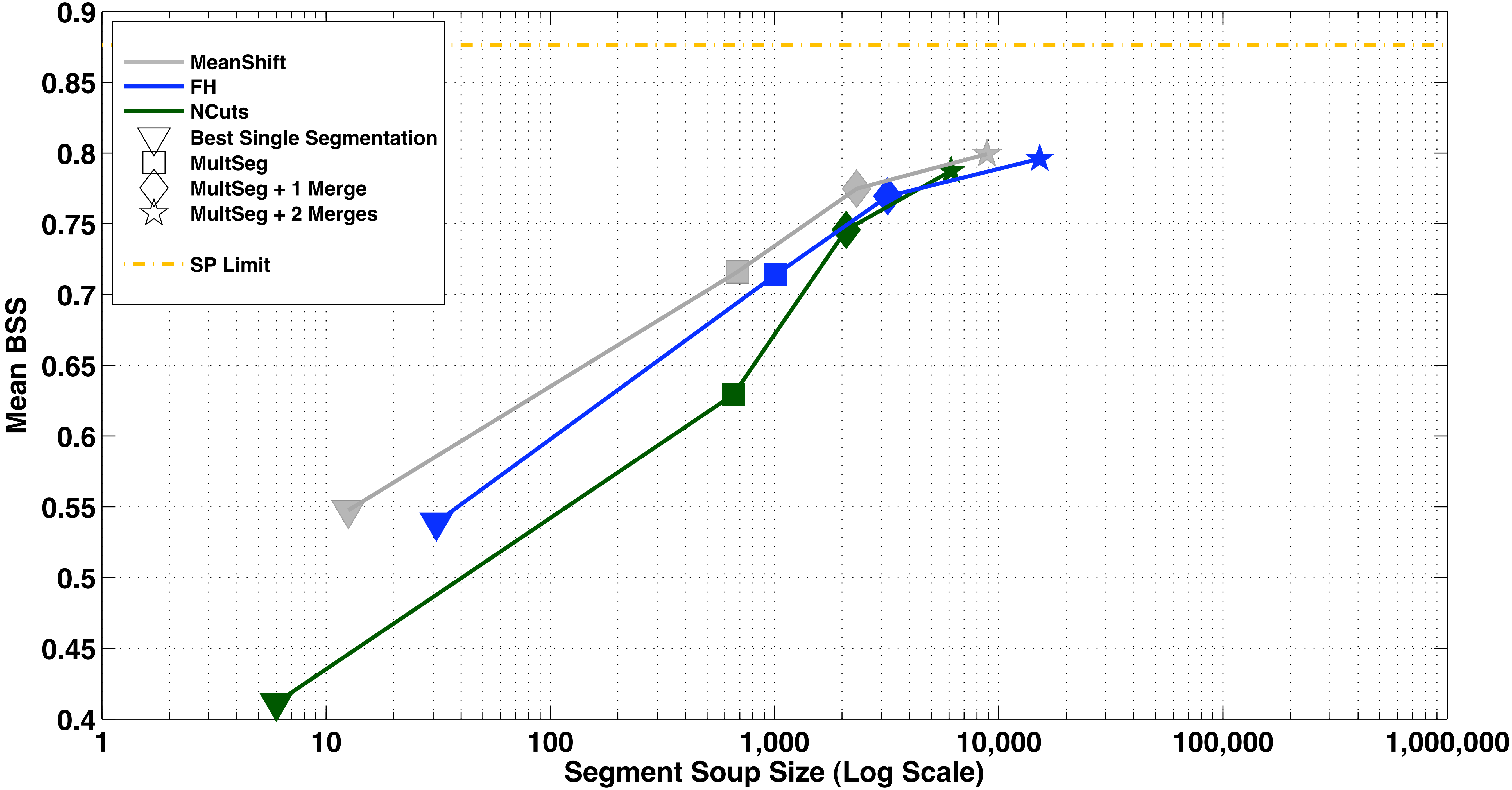
Superpixel Limit .917

Superpixel Limit .932



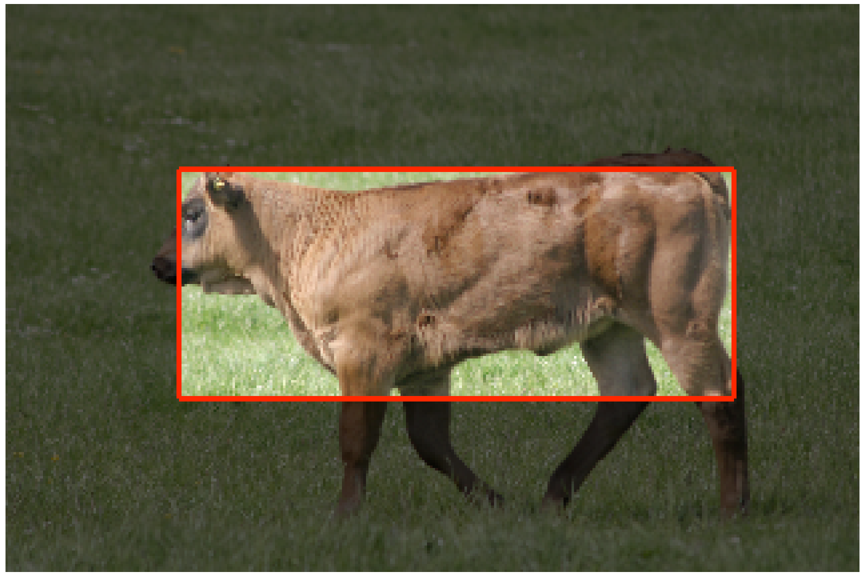
Superpixel Limit .825

Quantitative Results



Upper-Bound: Rectangular Windows

- Consider the best* rectangular spatial support
- Infeasible in practice

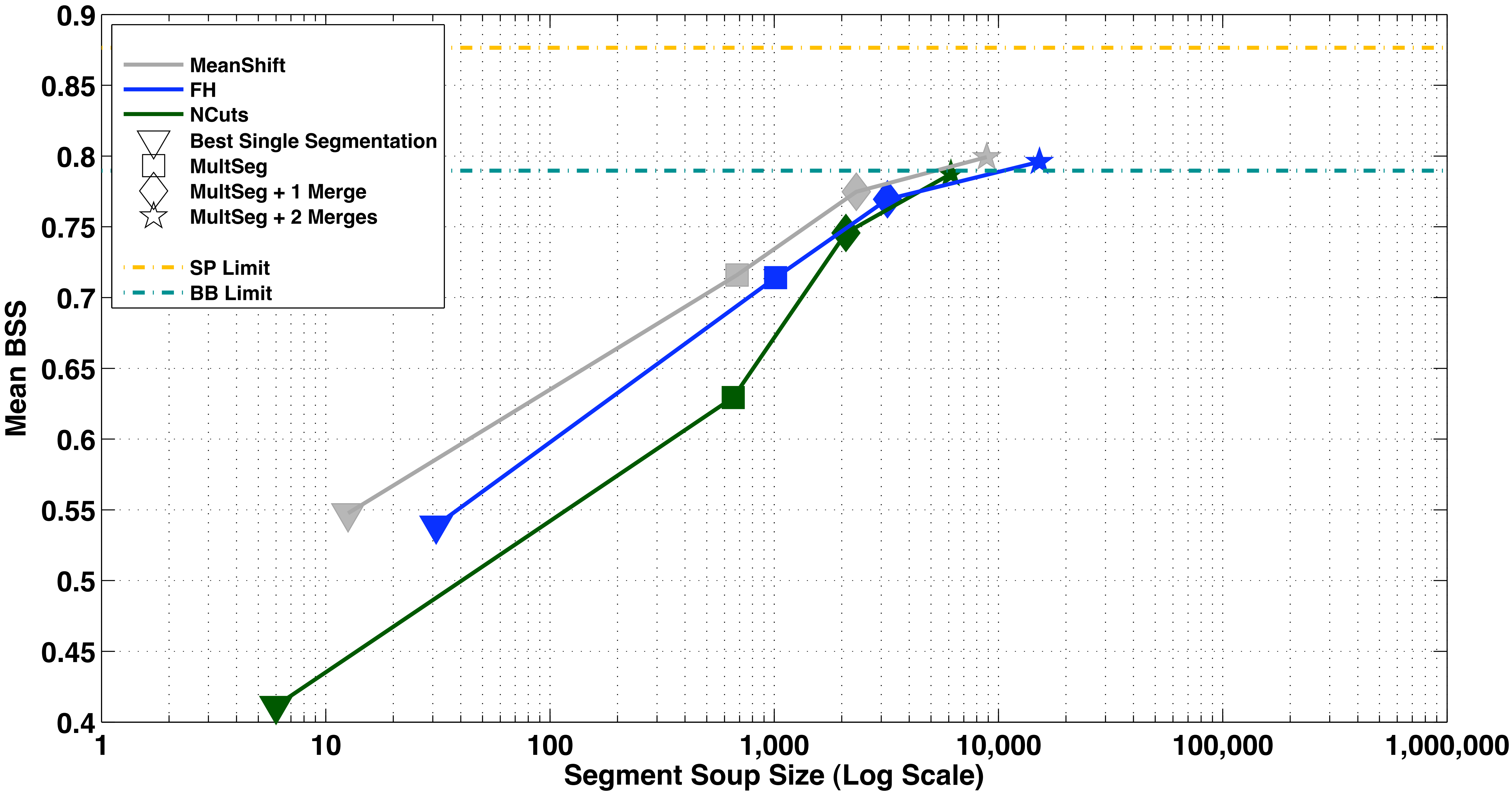


Rectangular Limit .682

Rectangular Limit .909

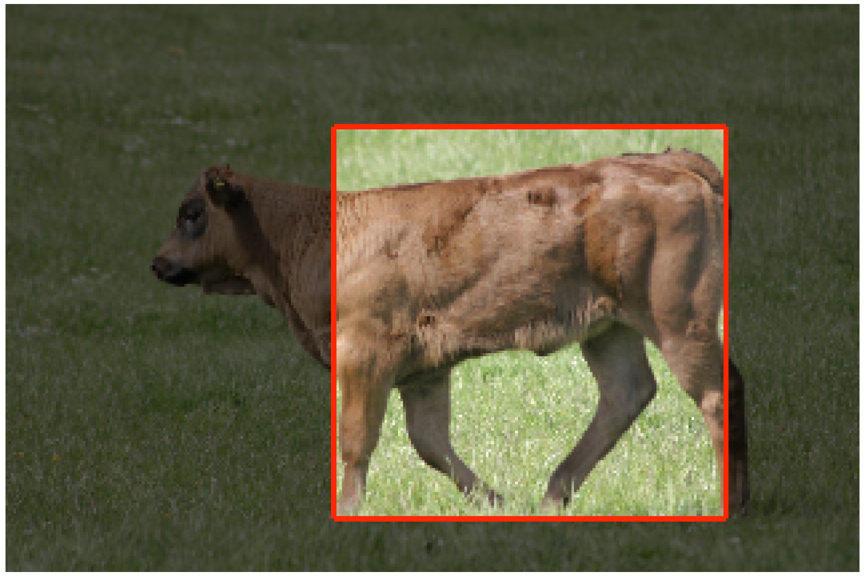
Rectangular Limit .616

Quantitative Results



Viola-Jones Sliding Windows

- Generate soup of segments by sliding square windows
- Often used in practice

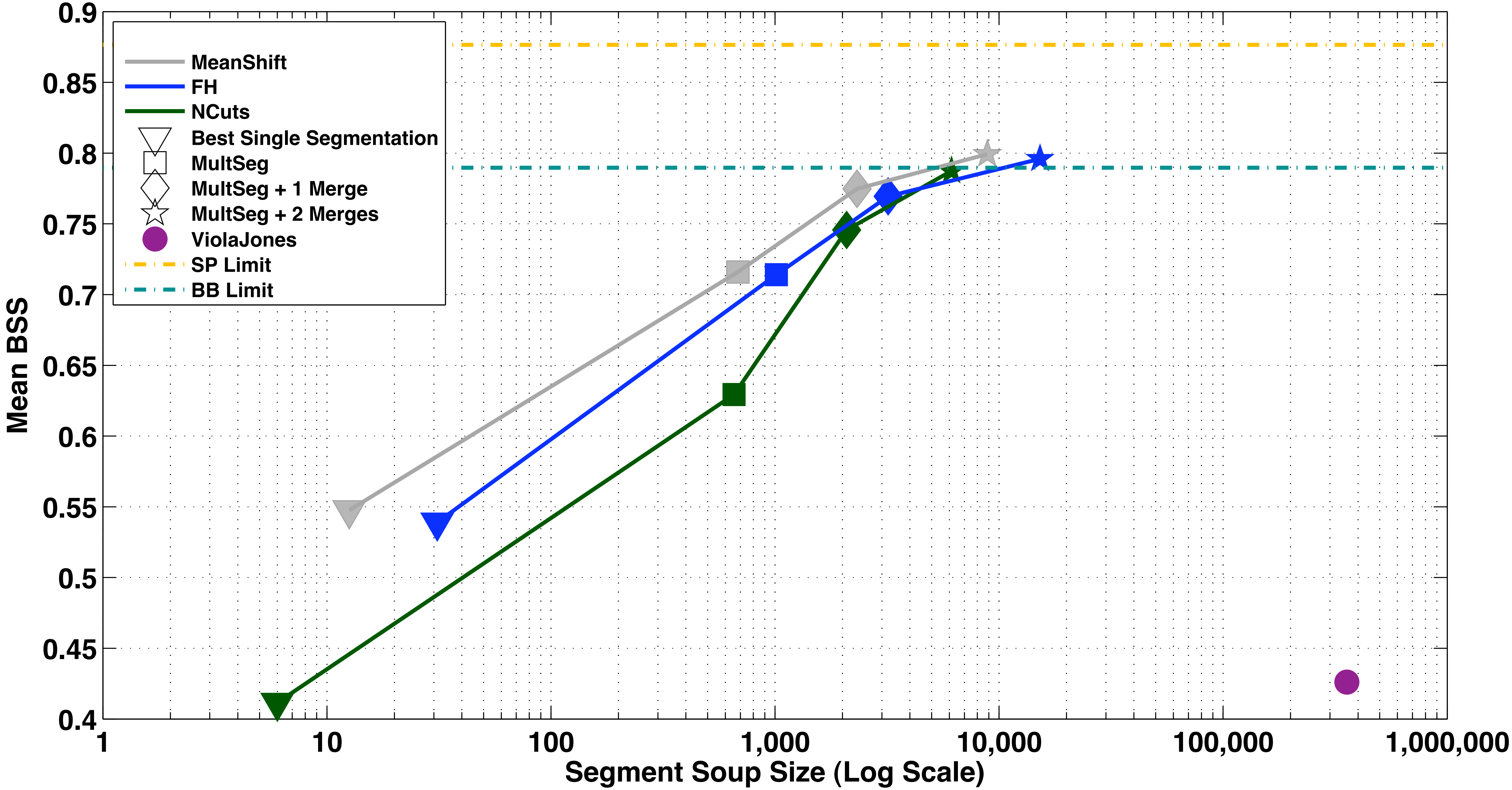


Square .555

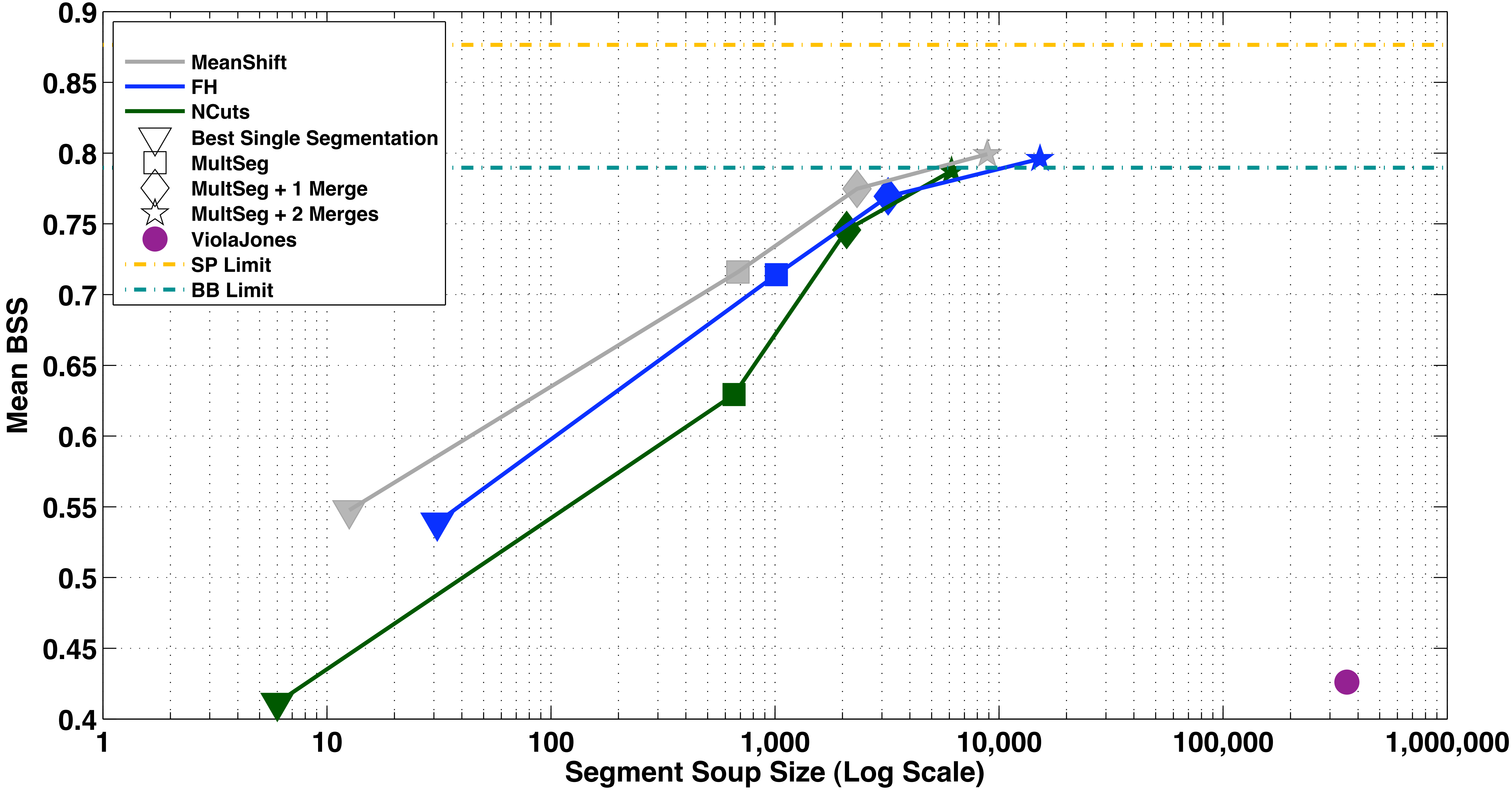
Square .495

Square .301

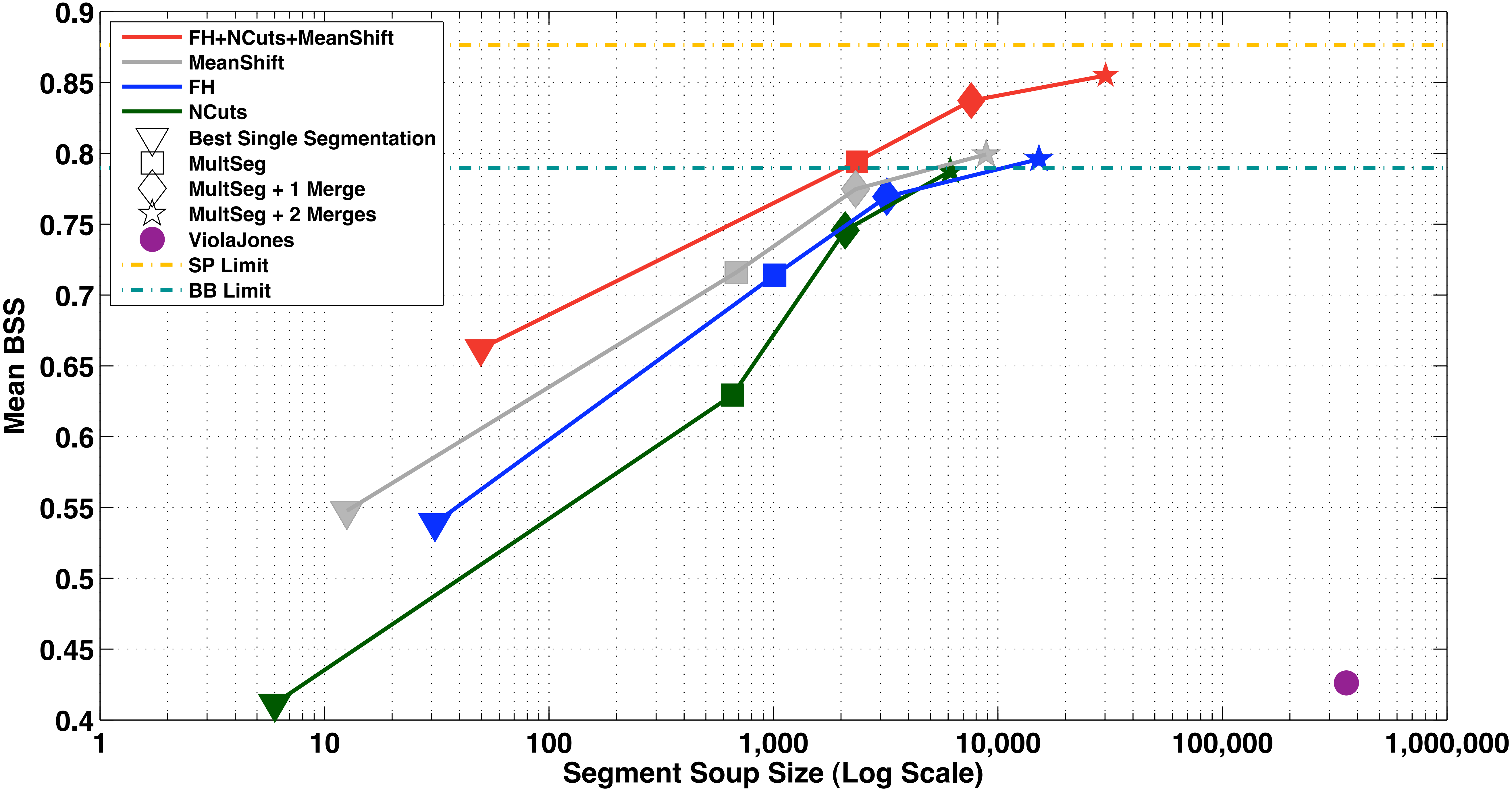
Comparing to Limits



Which Segmentation Algorithm is the best?



Which Segmentation Algorithm is the best?



Conclusions

Conclusions

- Correct Spatial Support is important for recognition

Conclusions

- Correct Spatial Support is important for recognition
- Multiple Segmentations are better than one

Conclusions

- Correct Spatial Support is important for recognition
- Multiple Segmentations are better than one
- Mean-Shift is better than FH or NCuts, but together they do best

Conclusions

- Correct Spatial Support is important for recognition
- Multiple Segmentations are better than one
- Mean-Shift is better than FH or NCuts, but together they do best
- Segment merging can benefit any segmentation

Conclusions

- Correct Spatial Support is important for recognition
- Multiple Segmentations are better than one
- Mean-Shift is better than FH or NCuts, but together they do best
- Segment merging can benefit any segmentation
- "Segment Soup" is large, but not catastrophically large

Conclusions

- Correct Spatial Support is important for recognition
- Multiple Segmentations are better than one
- Mean-Shift is better than FH or NCuts, but together they do best
- Segment merging can benefit any segmentation
- "Segment Soup" is large, but not catastrophically large

Questions?