

Interactive Foreground Segmentation in Images and Videos

Suyog Jain

1

Foreground Segmentation

Generate pixel level foreground masks for objects in a given image or video



2

Why is Foreground Segmentation useful?

Better Visual Search



Results from AlchemyAPI search

Many irrelevant images appear in search result

3

Why is Foreground Segmentation useful?

Better Visual Search



Results from AlchemyAPI search

Search can focus on the object of interest

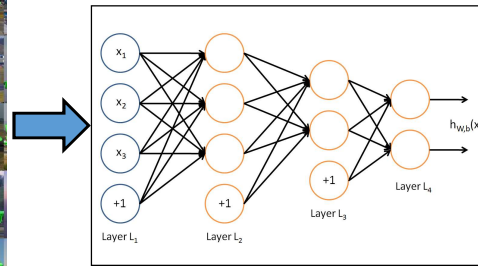
4

Why is Foreground Segmentation useful?

Training object recognition systems



Training Images



Recognition System

Can benefit from well segmented objects during training

5

Why is Foreground Segmentation useful?

Computer Graphics

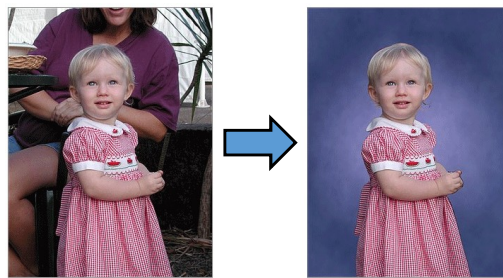


Image & Video editing

Need accurate foreground segmentations

6

Why is Foreground Segmentation useful?

Computer Graphics



3D Reconstruction

Can benefit from foreground segmentations

[Snavely, ICCV 2009]

7

Human-Machine Collaboration



Good at perception and can easily identify the foreground regions

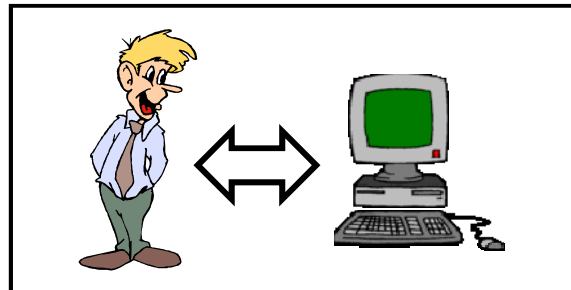


Good at processing large volumes of data at the lowest level of details very efficiently

Complementary strengths

8

Human-Machine Collaboration



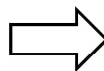
Bringing them together can lead to systems which can be accurate and cost effective

9

Interactive Foreground Segmentation



**Human
input**

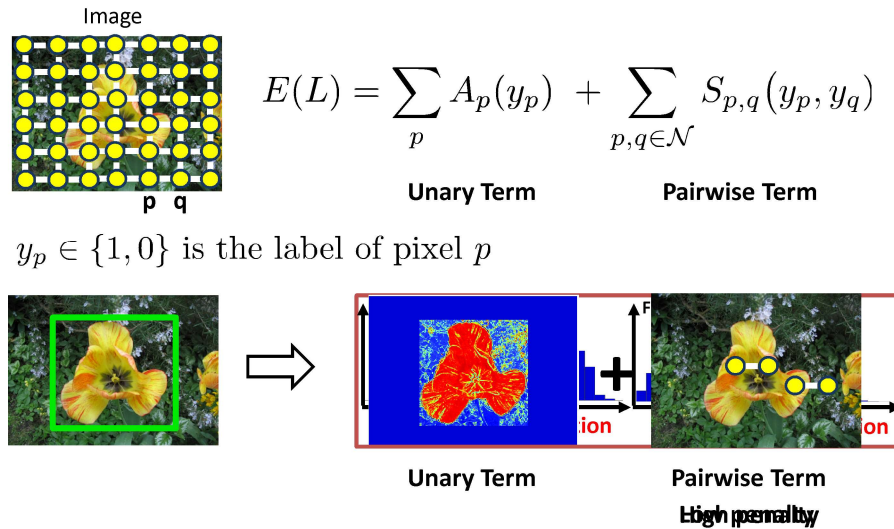


**Segmentation
algorithm**

[Boykov 2001, Zabih 2001, Gulshan 2010, Kohli 2008]

10

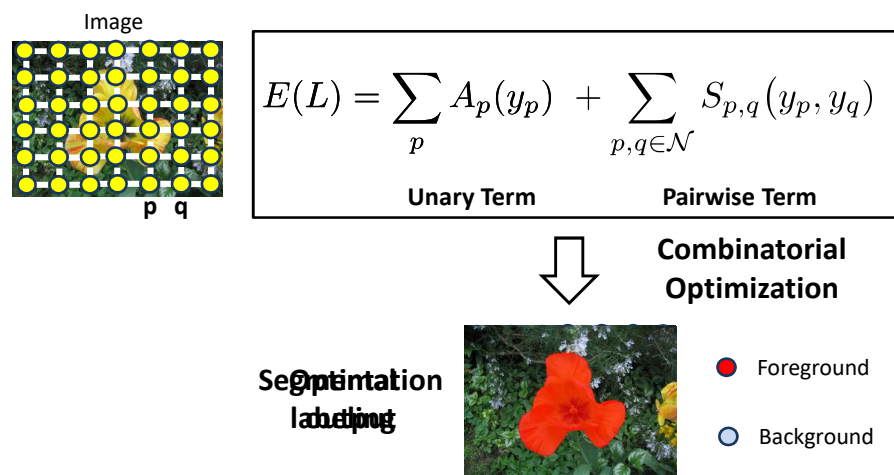
MRF-Segmentation Model



[Boykov 2001, Zabih 2001, Gulshan 2010, Kohli 2008]

11

MRF-Segmentation Model



[Boykov 2001, Zabih 2001, Gulshan 2010, Kohli 2008]

12

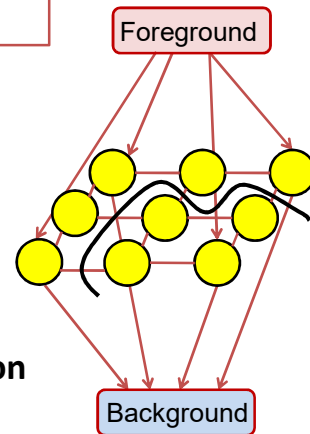
MRF-Segmentation Model

$$E(L) = \sum_p A_p(y_p) + \sum_{p,q \in \mathcal{N}} S_{p,q}(y_p, y_q)$$

$y_p \in \{1, 0\}$ is the label of pixel p

**Can be solved efficiently using
Max flow algorithms**

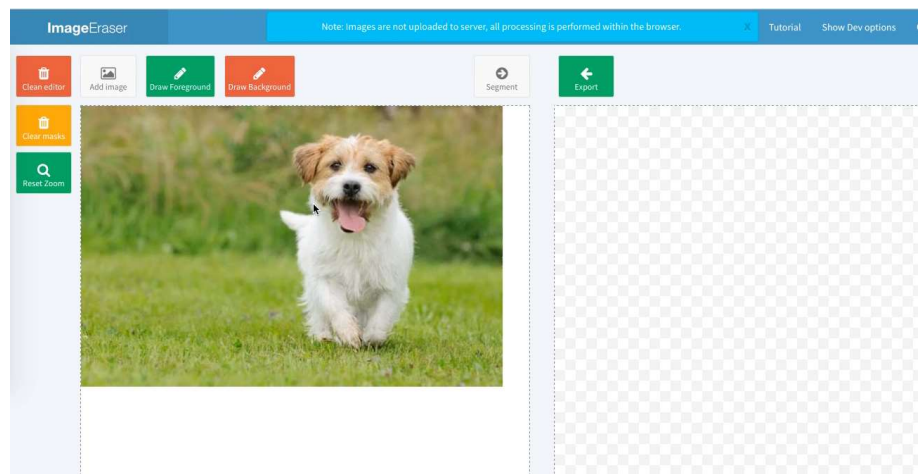
Also known as Graph Cuts Segmentation



[Boykov 2001, Zabih 2001, Gulshan 2010, Kohli 2008]

13

Demo



14

Video Object Segmentation

Generate pixel level foreground masks for an object(s) across the frames of a video

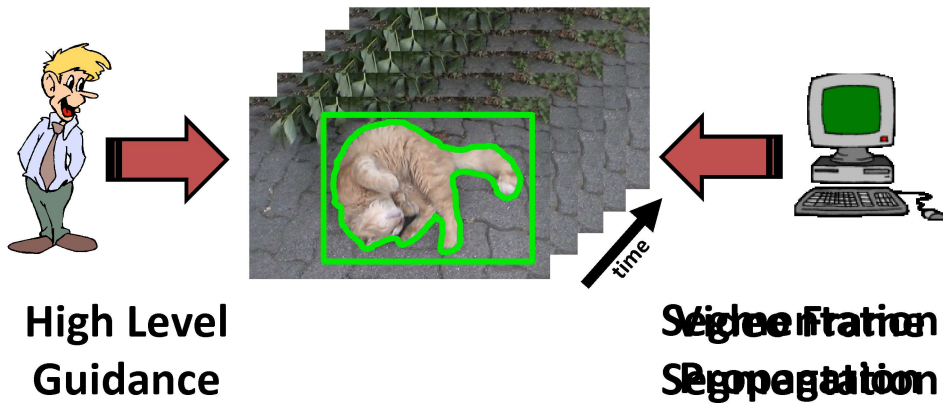


Need only a couple of clicks!

15

Interactive Video Segmentation

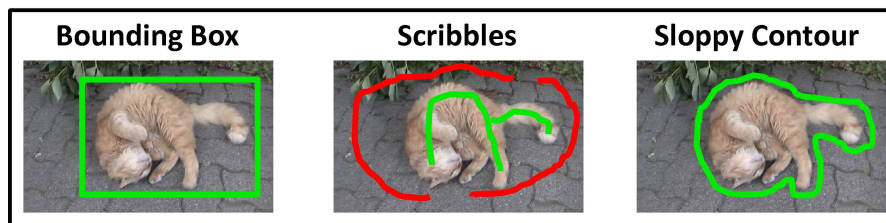
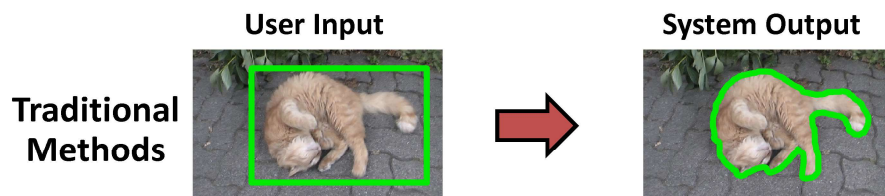
Bring the complementary strengths of humans and machines together.



16

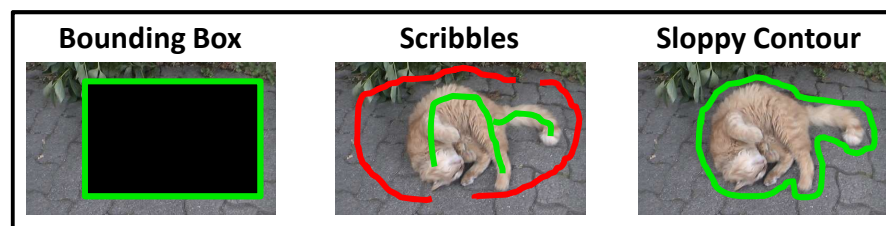
Interactive Video Segmentation

Get user input first and then generate a segmentation hypothesis



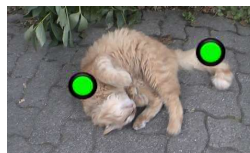
[Boykov 2001, Zabih 2001, Rother 2004, Kohli 2008]

Interactive Video Segmentation



VS.

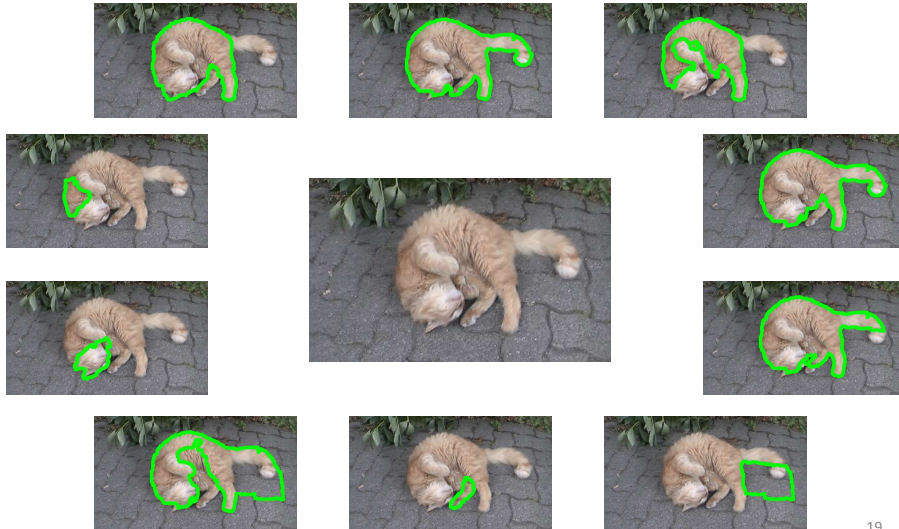
Point Clicks?



[Boykov 2001, Zabih 2001, Rother 2004, Kohli 2008]

Our Idea - Flip the process

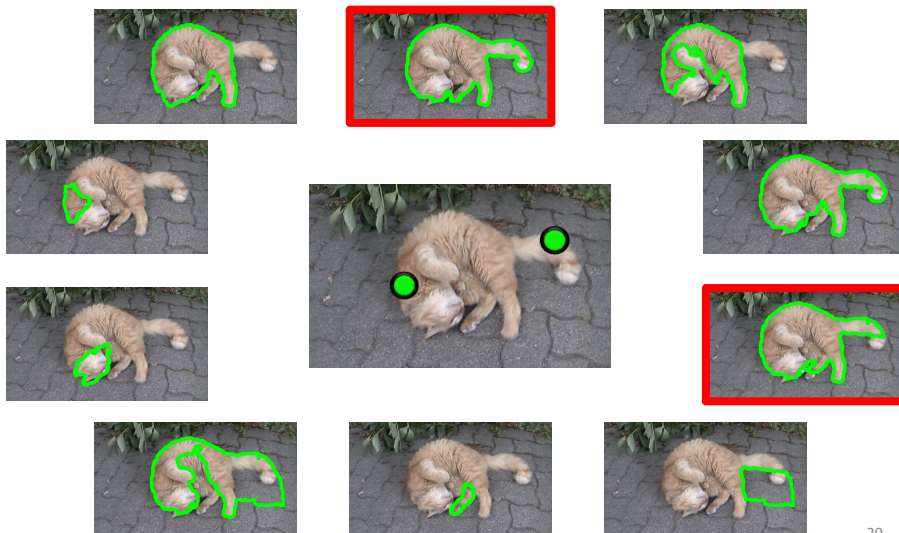
Pre-generate thousands of segmentations with **no human input**.



19

Our Idea - Flip the process

Use **boundary clicks** to quickly “carve” out the accurate ones.

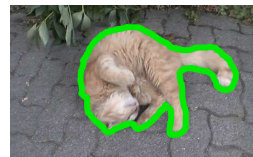


20

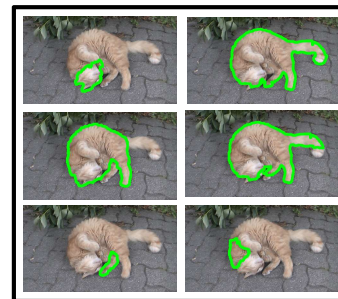
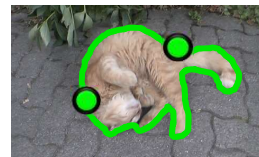
Interactive Video Segmentation

More accurate segmentation with less annotation cost.

Traditional Methods



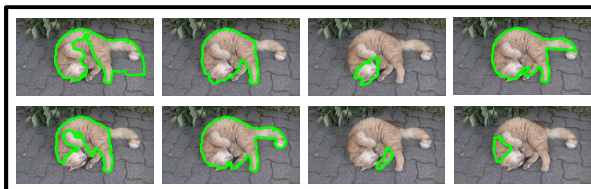
Ours



[Boykov 2001, Zabih 2001, Rother 2004, Kohli 2008]

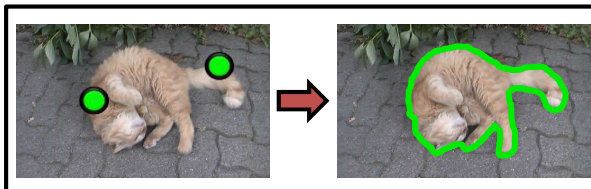
Overview

1.



Region Proposals

2.



Click Carving

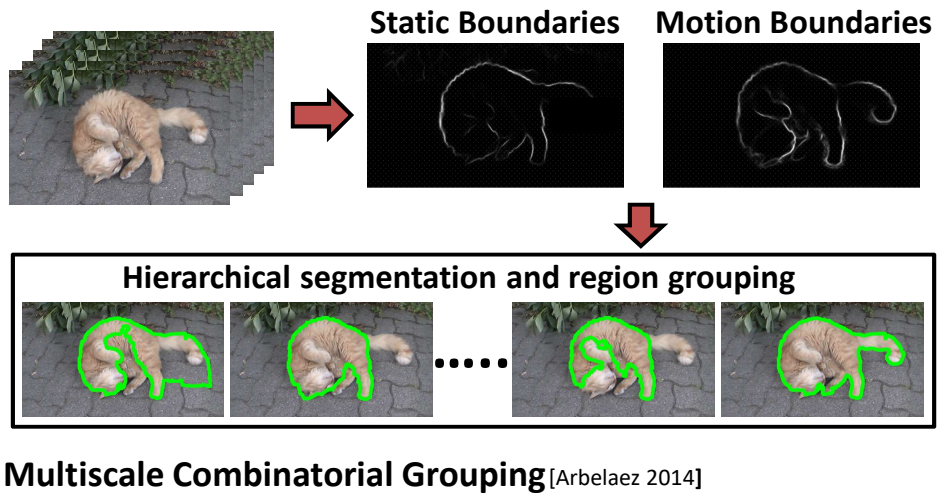
3.



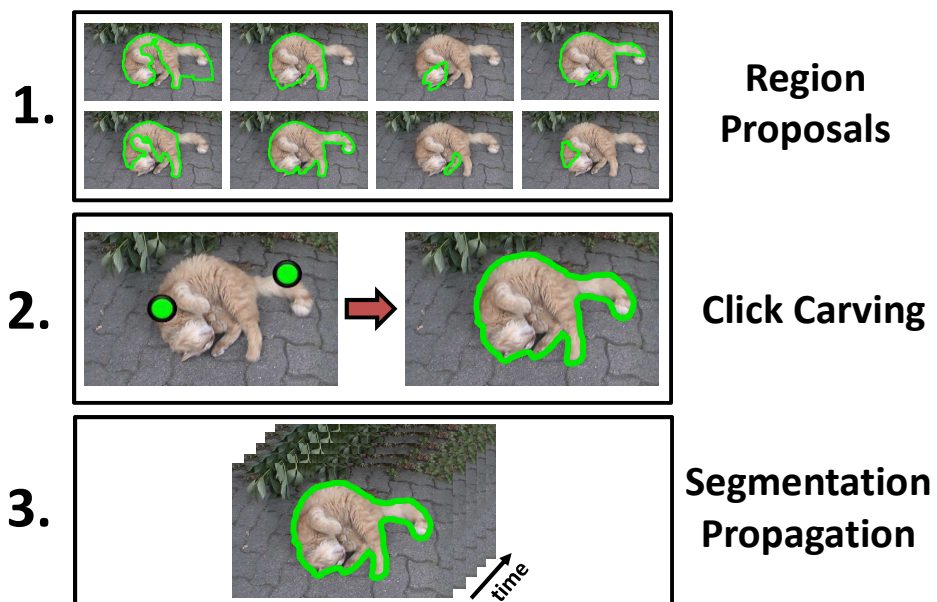
Segmentation Propagation

Region Proposals

Use perceptual grouping cues to generate thousands of object segmentations with **no human input**.



Overview



Click Carving

1 1 1
1 1 1
0 1
1 0 0

 Votes

25

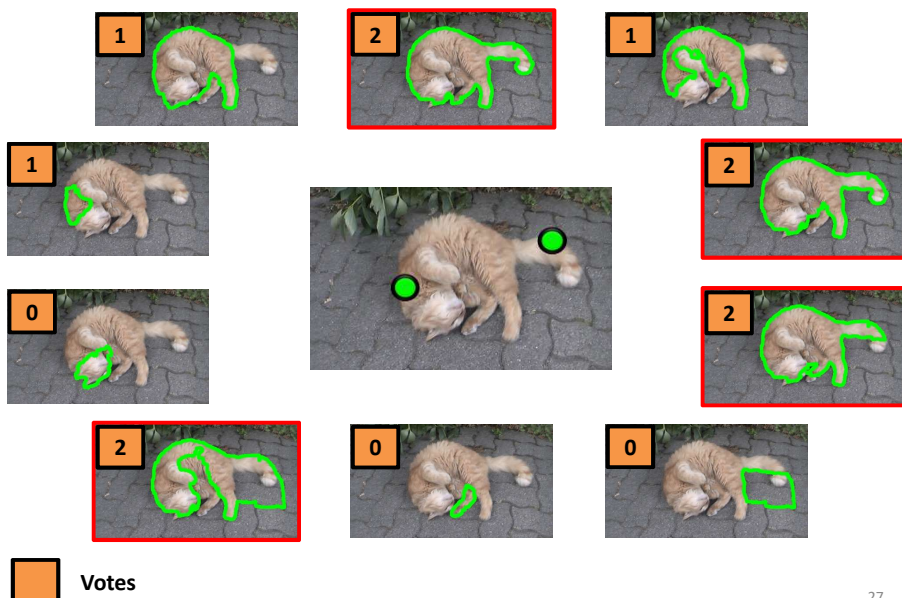
Click Carving

1 1 1
1 1 1
0 1
1 0 0

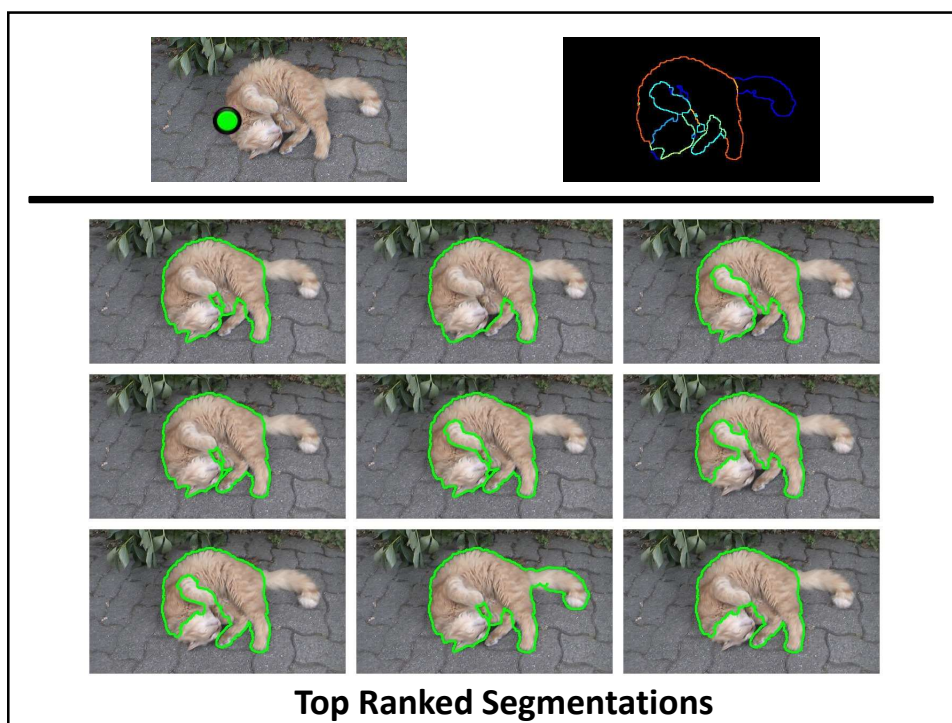
 Votes

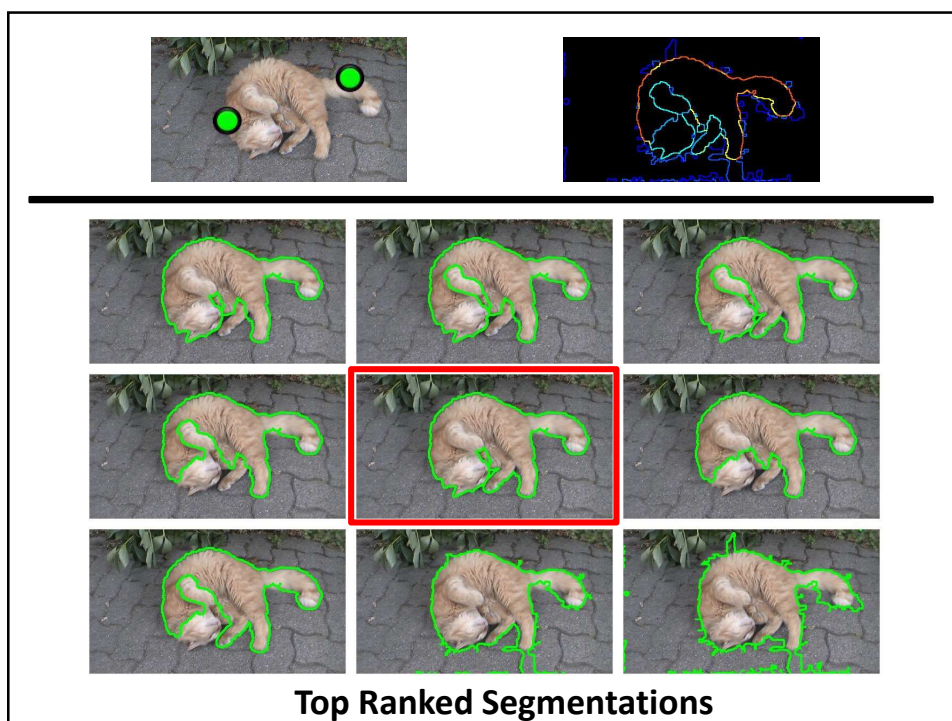
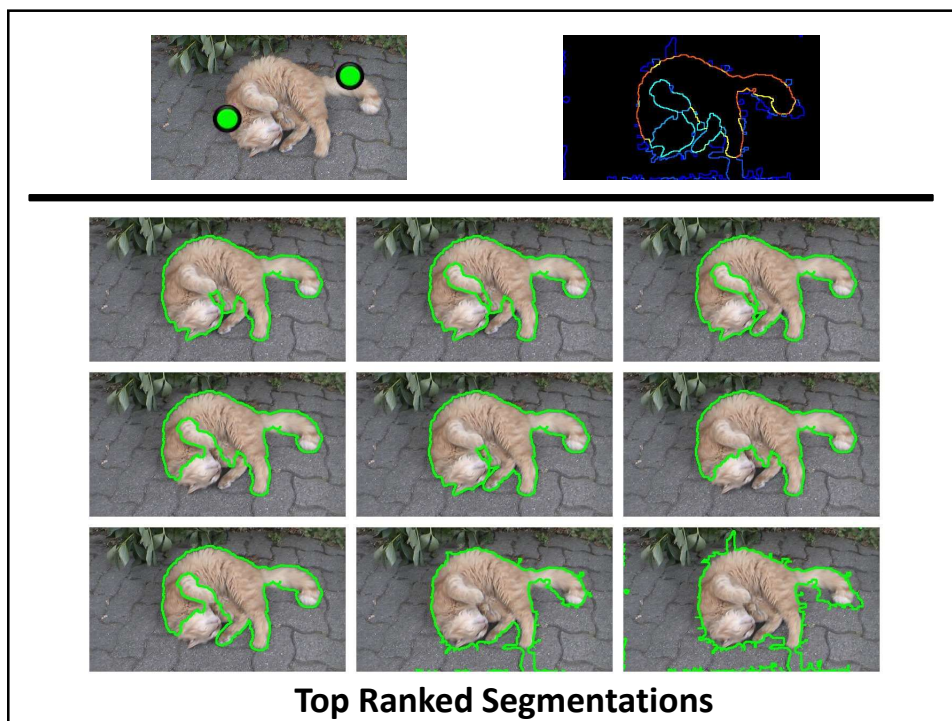
26

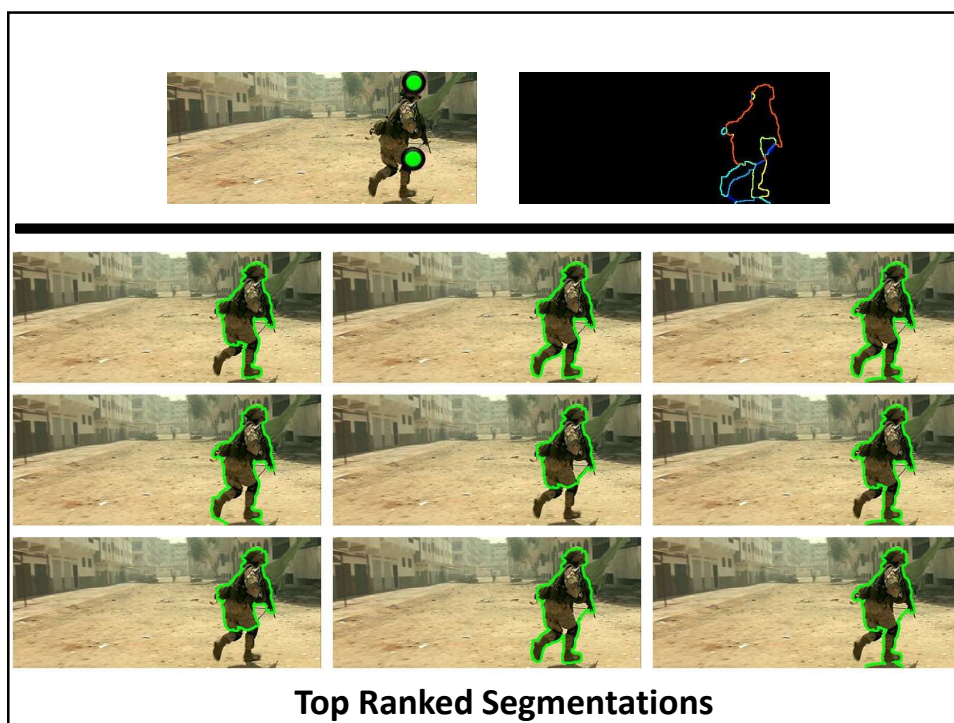
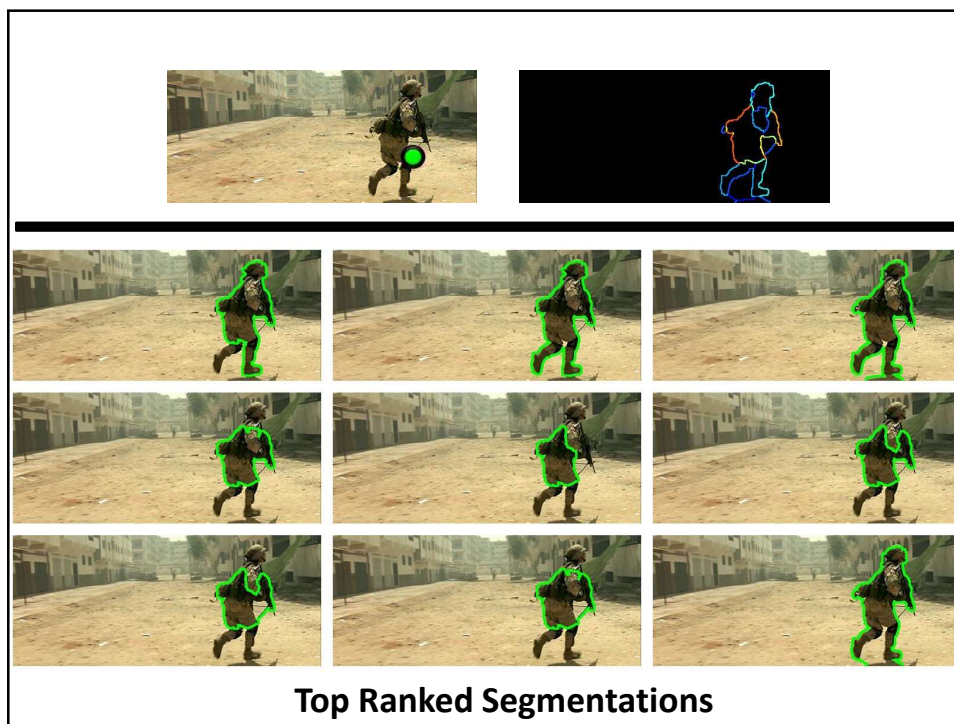
Click Carving

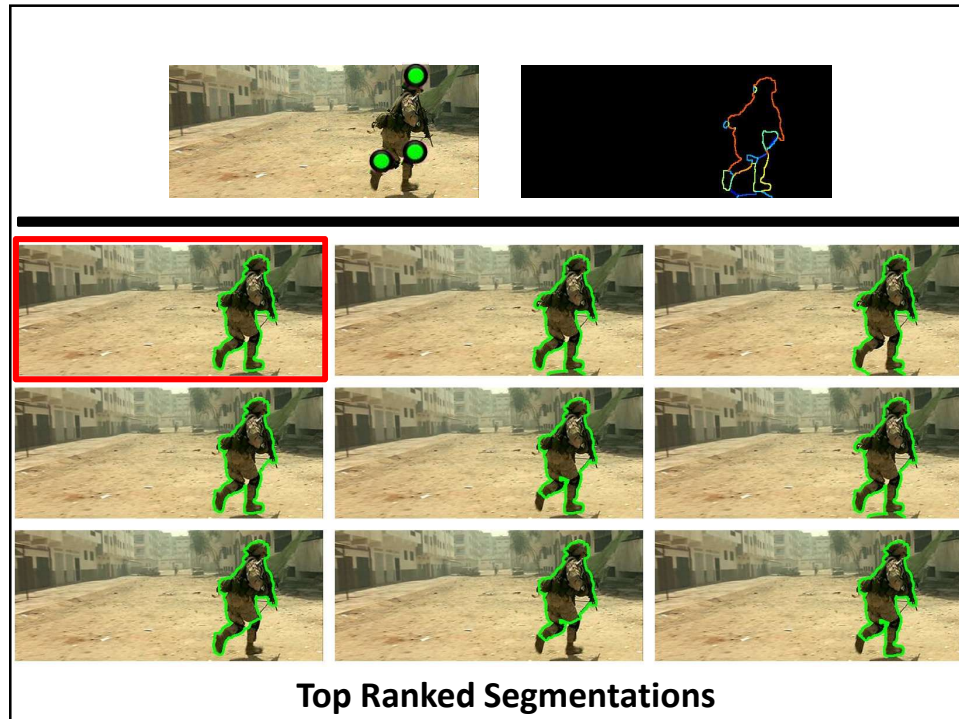


27



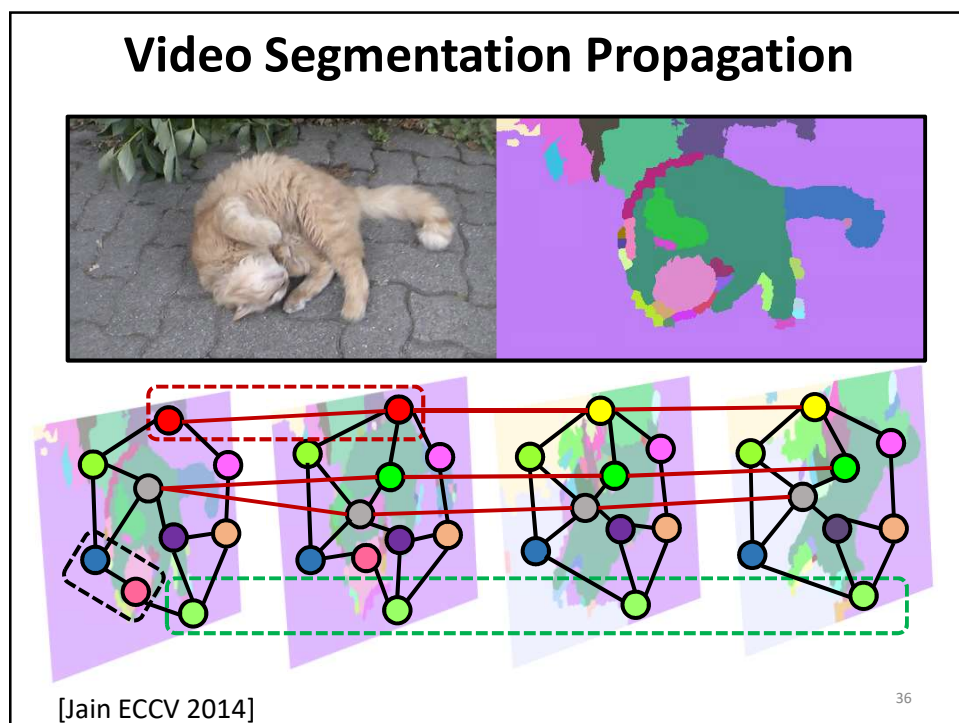
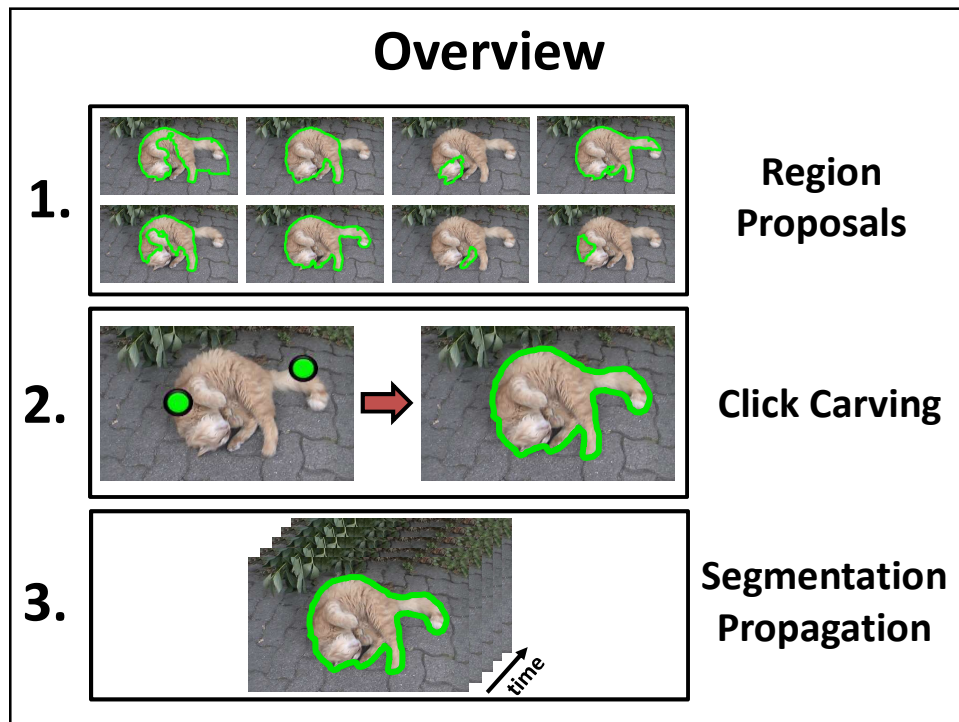






Click Carving – User Interface

Click Carving: Segmenting
Objects in Video
with Point Clicks



Results

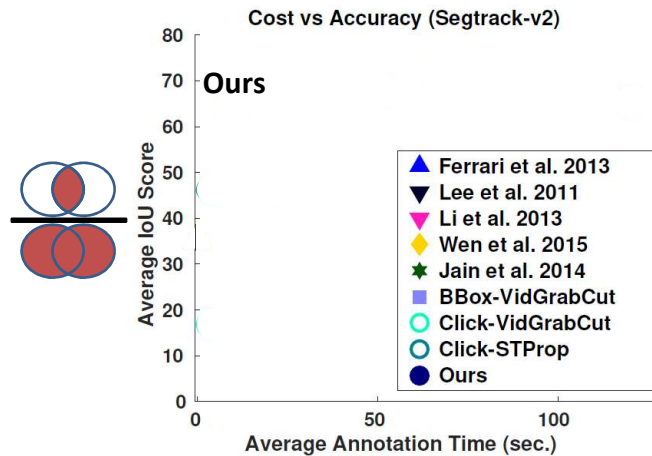
37

Experimental Setup

- Evaluate on 3 challenging video segmentation datasets:
 - Segtrack-v2 [Li et al. 2013]
 - VSB 100 [Sundber et al. 2011]
 - iVideoSeg [Nagaraja et al. 2015]
- User study:
 - 3 annotators with a max annotation budget of 10 clicks.
 - Record number of clicks, time spent and best object mask chosen by the annotator.
- Compare with several existing methods which use different amount of human annotation.

38

Click Carving Results

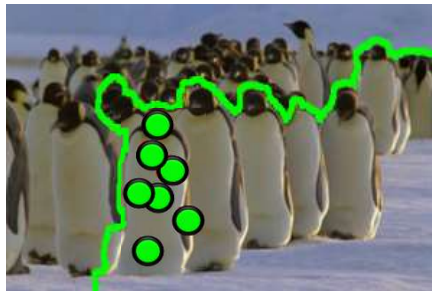


Excellent cost vs. accuracy tradeoff

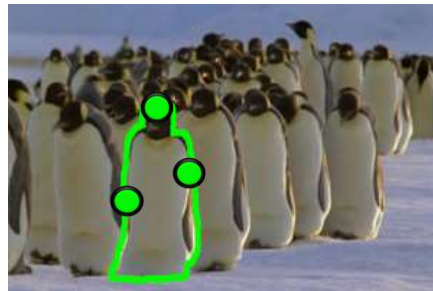
Click Carving Results

Why use boundary clicks?

Interior Clicks



Boundary Clicks



Boundary clicks are far more discriminative than interior clicks.

Click Carving Results



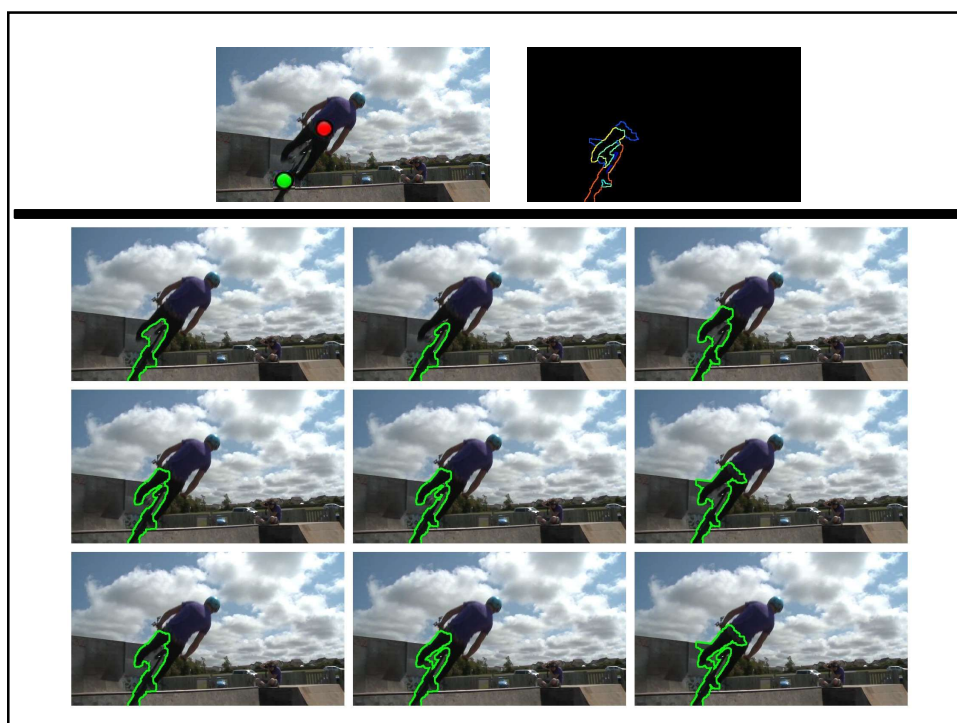
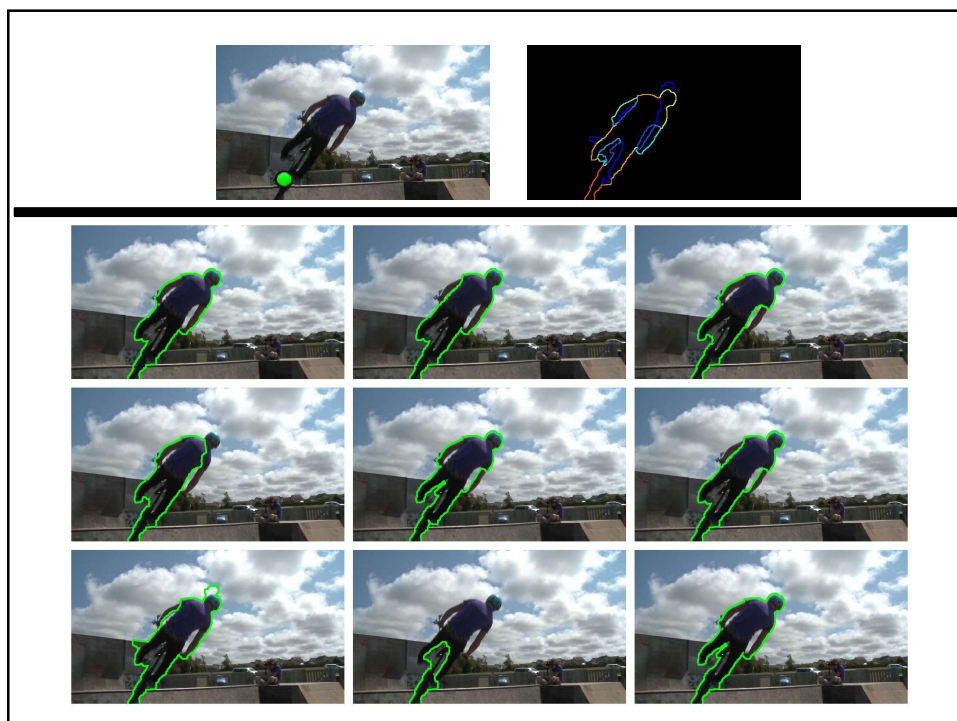
Using only 1-2 clicks

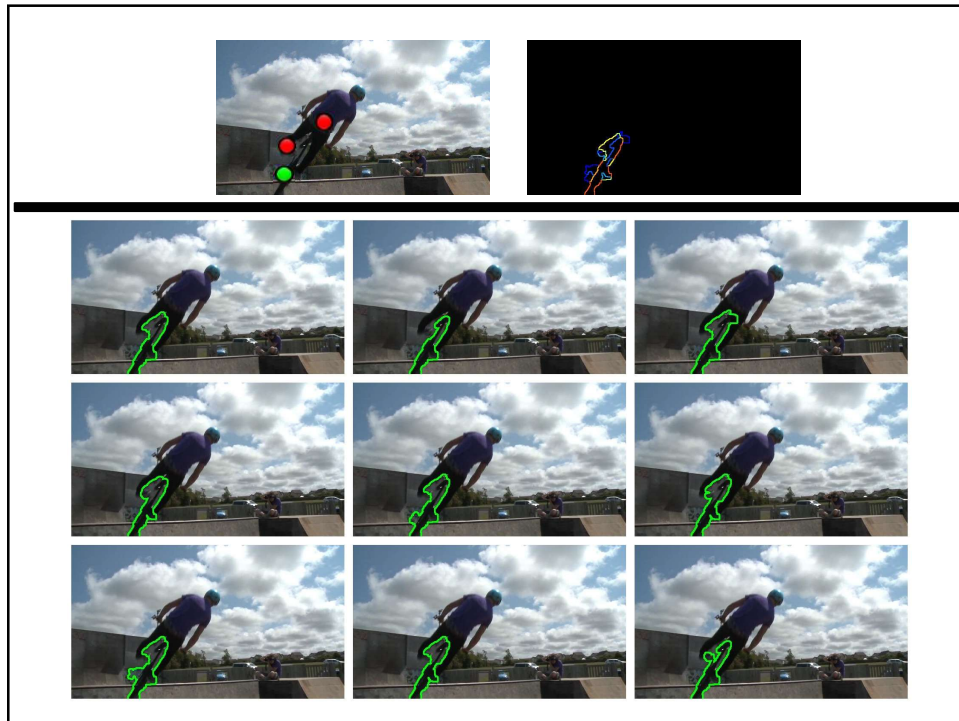
41

Additional Features

- Negative Clicks

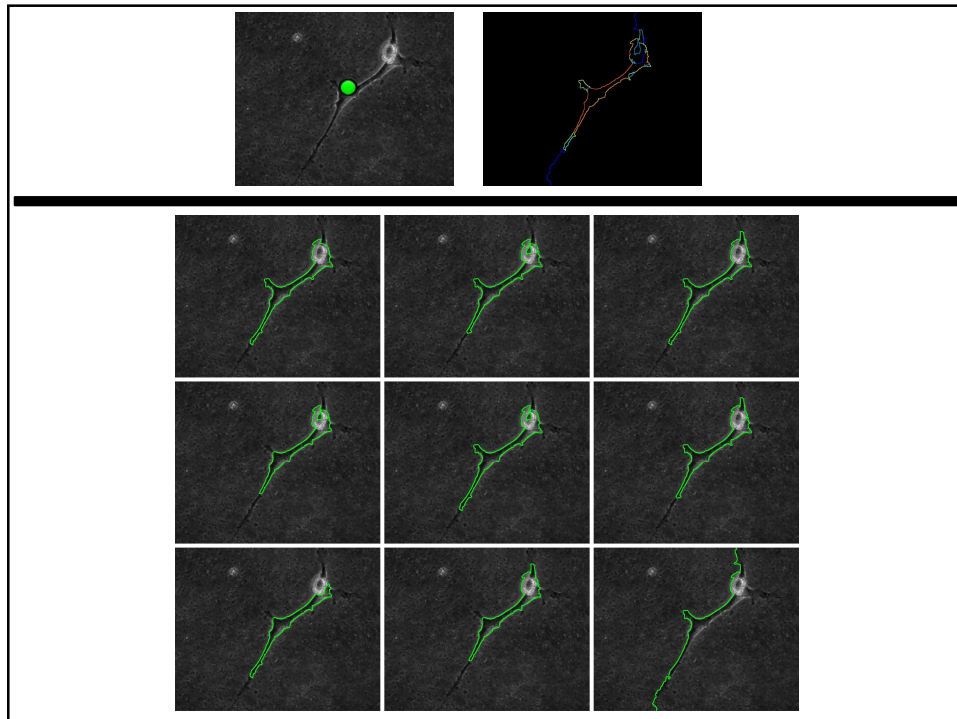
42





Additional Features

- Negative Clicks
- Biomedical Images



Demo

- Click Carving Demo
 - [Demo1](#)
 - [Demo2](#)
- Pixel Objectness
 - <http://vision.cs.utexas.edu/projects/pixelobjectness/>
- FusionSeg – video segmentation
 - <http://vision.cs.utexas.edu/projects/fusionseg/>

Questions

More details at:

<https://www.cs.utexas.edu/~suyog/>

49