



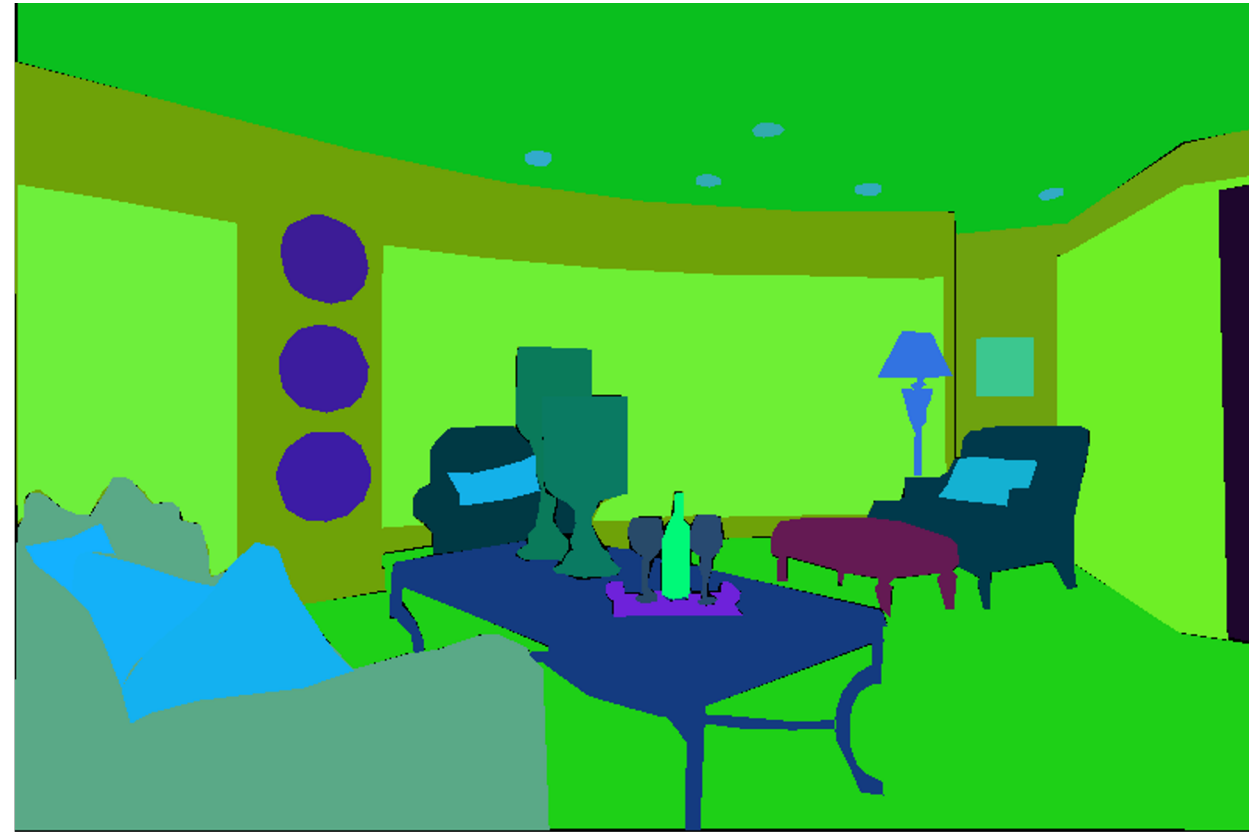
Scaling up instance annotation via label propagation

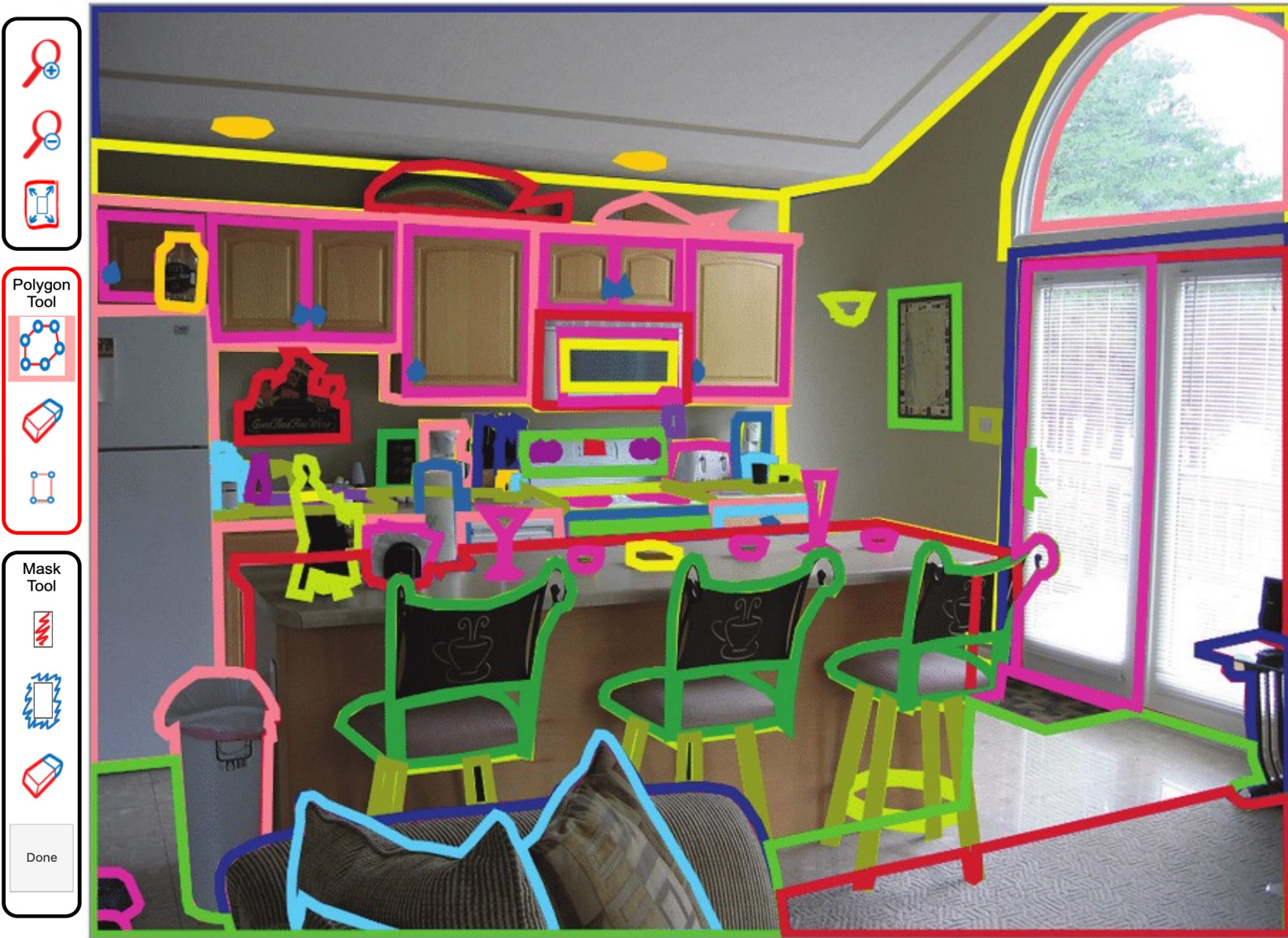
Dim P. Papadopoulos*, Ethan Weber*, Antonio Torralba

Instance Segmentation

Image

Prediction

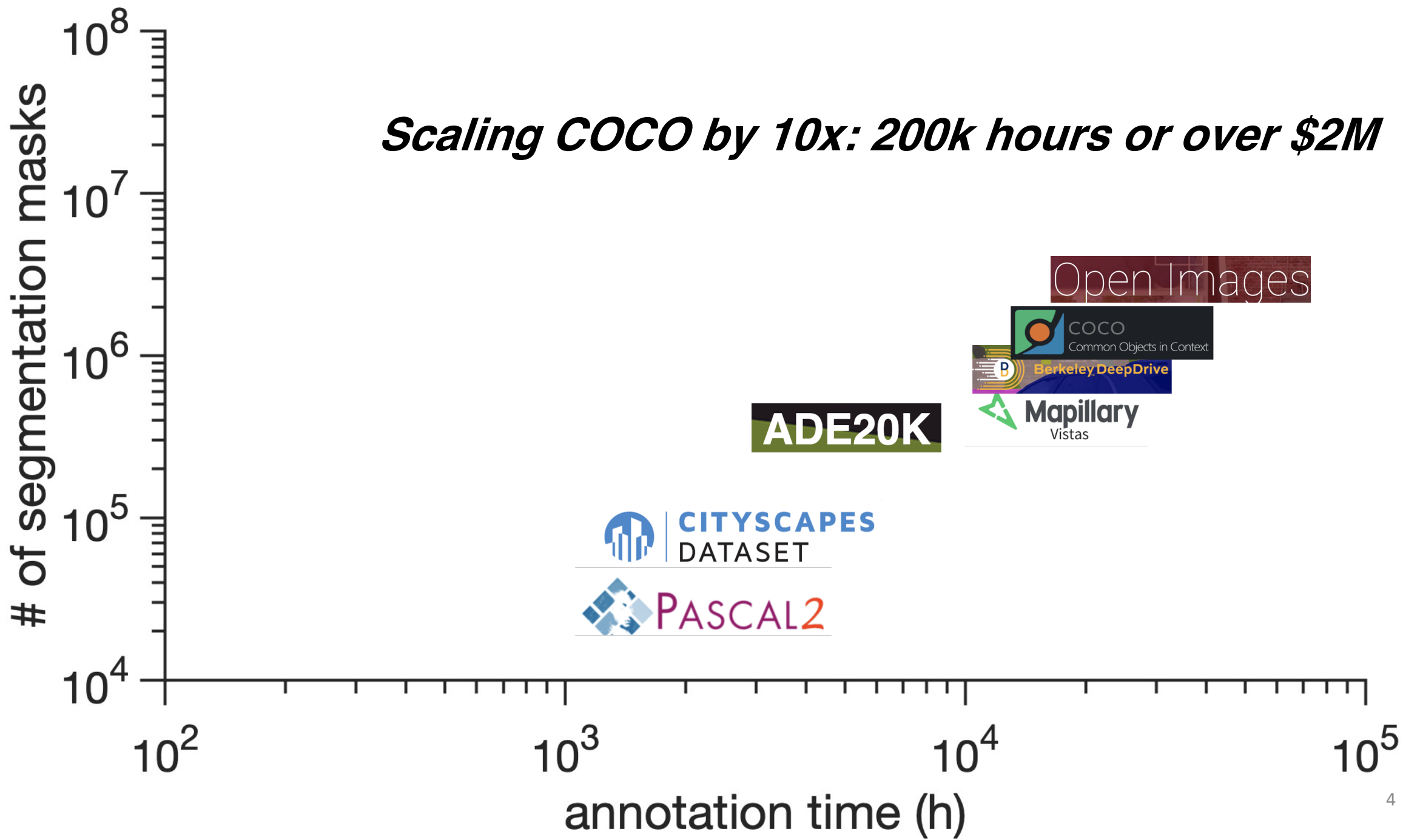




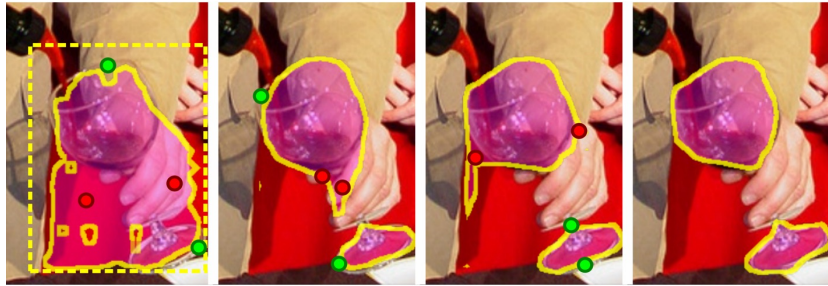
COCO:
80sec per object

CityScapes:
1.5 hours per image

Scaling COCO by 10x: 200k hours or over \$2M ●



Interactive Segmentation



[Xu CVPR 16, Li CVPR 18,
Chen CVPR 18, Benenson CVPR 19]



[Lin CVPR 16, Bearman ECCV 16,
Agustsson CVPR 19]

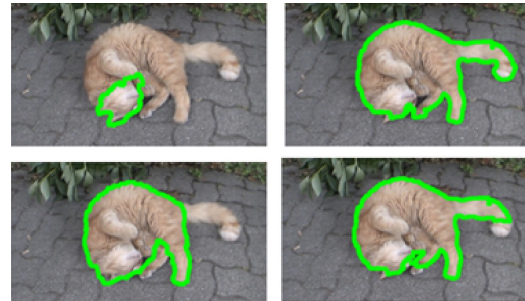
cheap

expensive

Binary Binary ++ Bounding box Pick among N Click/Scribble/ Edit polygon



[Papadopoulos CVPR 16]



[Jain IJCV 19]



[Castrejon CVPR 17, Acuna CVPR 18,
Ling CVPR 19]

+ Cut the unit cost of each annotation

- Cost is linear to the number of instances

Motivation

At large scale, instances with similar appearance and similar segmentation prediction



Motivation

Cluster



Motivation

Cluster, Annotate



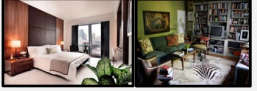
Motivation

Cluster, Annotate and Propagate



Method Overview

Detection



$$s_1=0.95 f_1$$



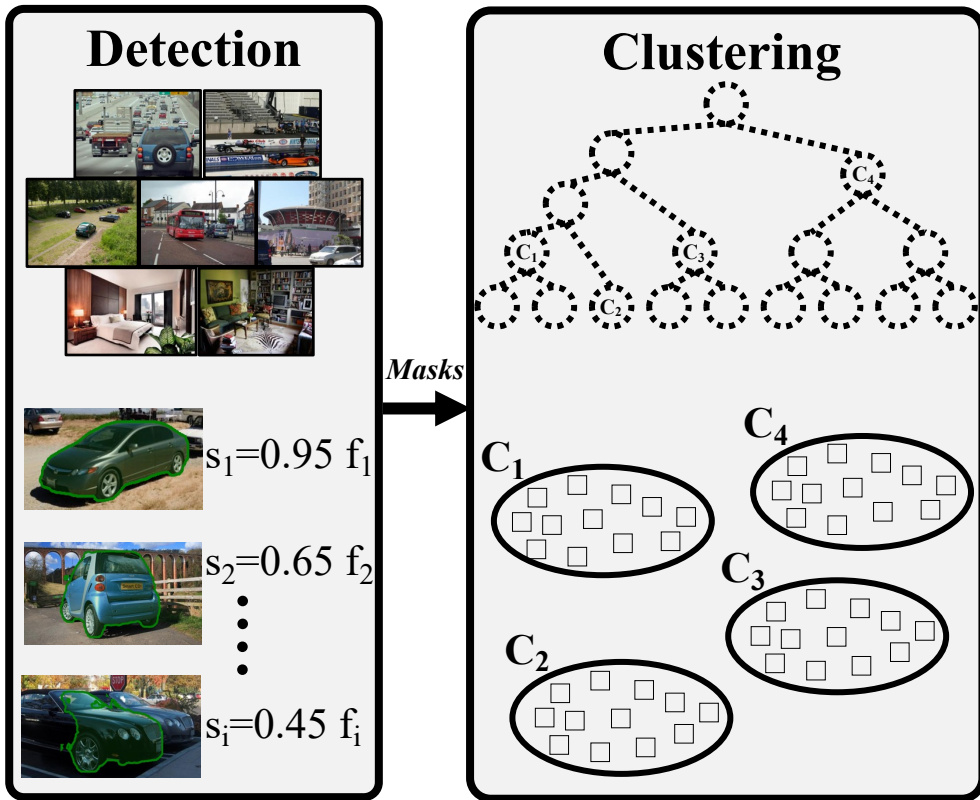
$$s_2=0.65 f_2$$

⋮

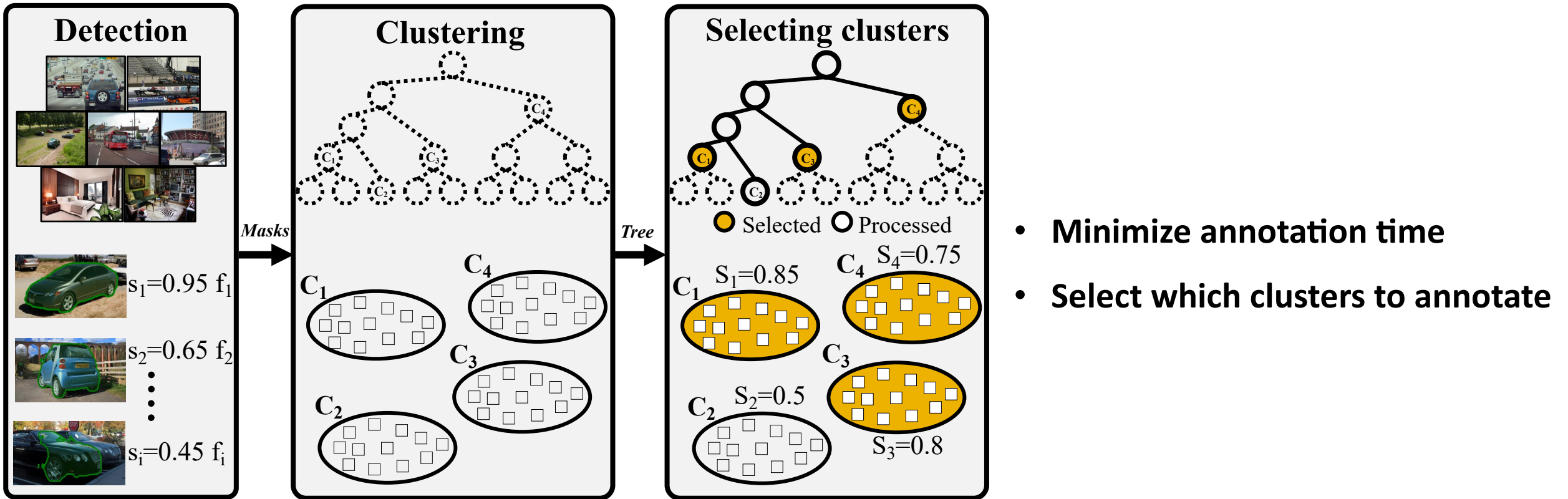


$$s_i=0.45 f_i$$

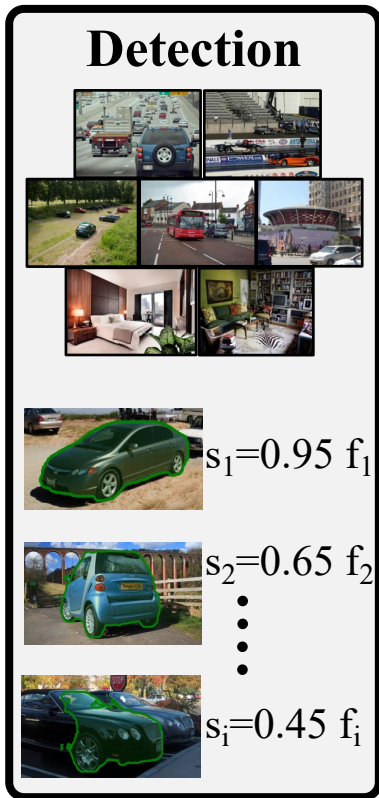
Method Overview



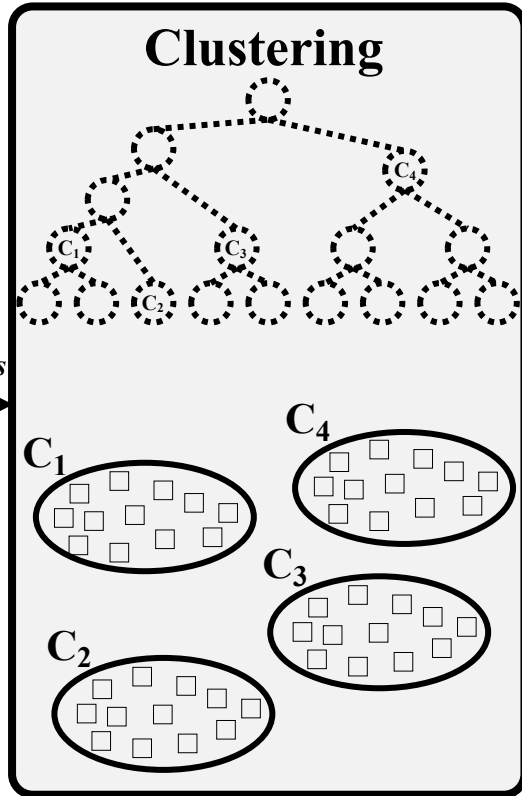
Method Overview



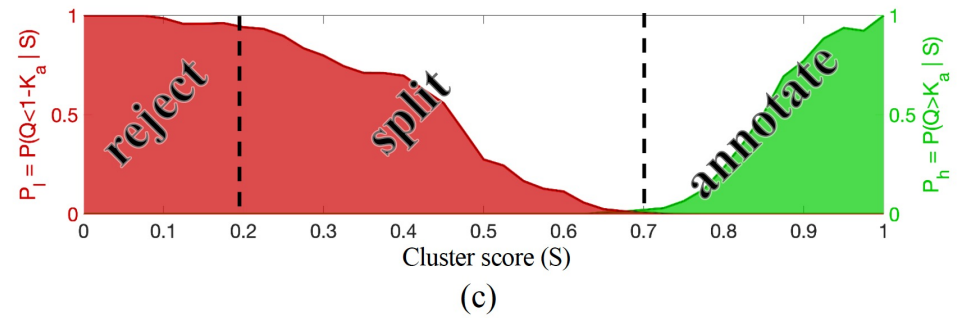
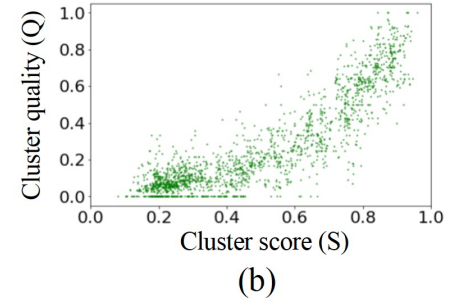
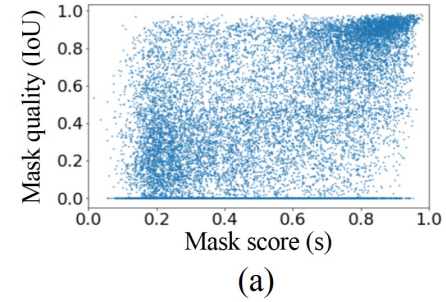
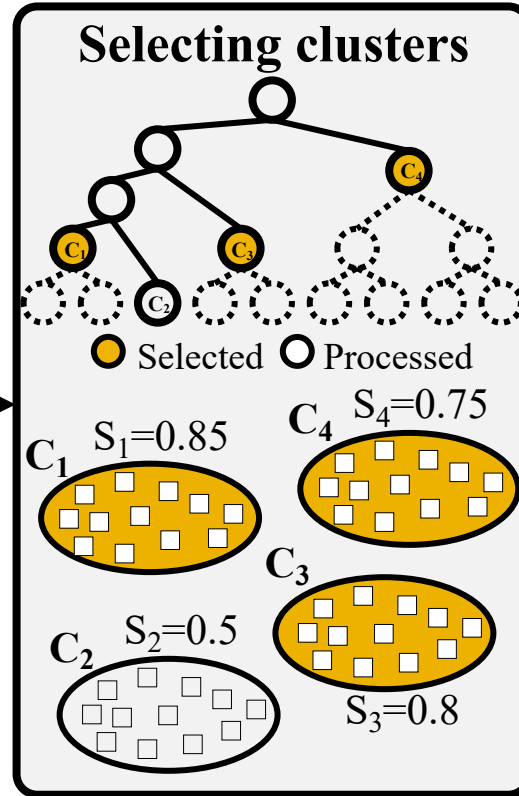
Method Overview



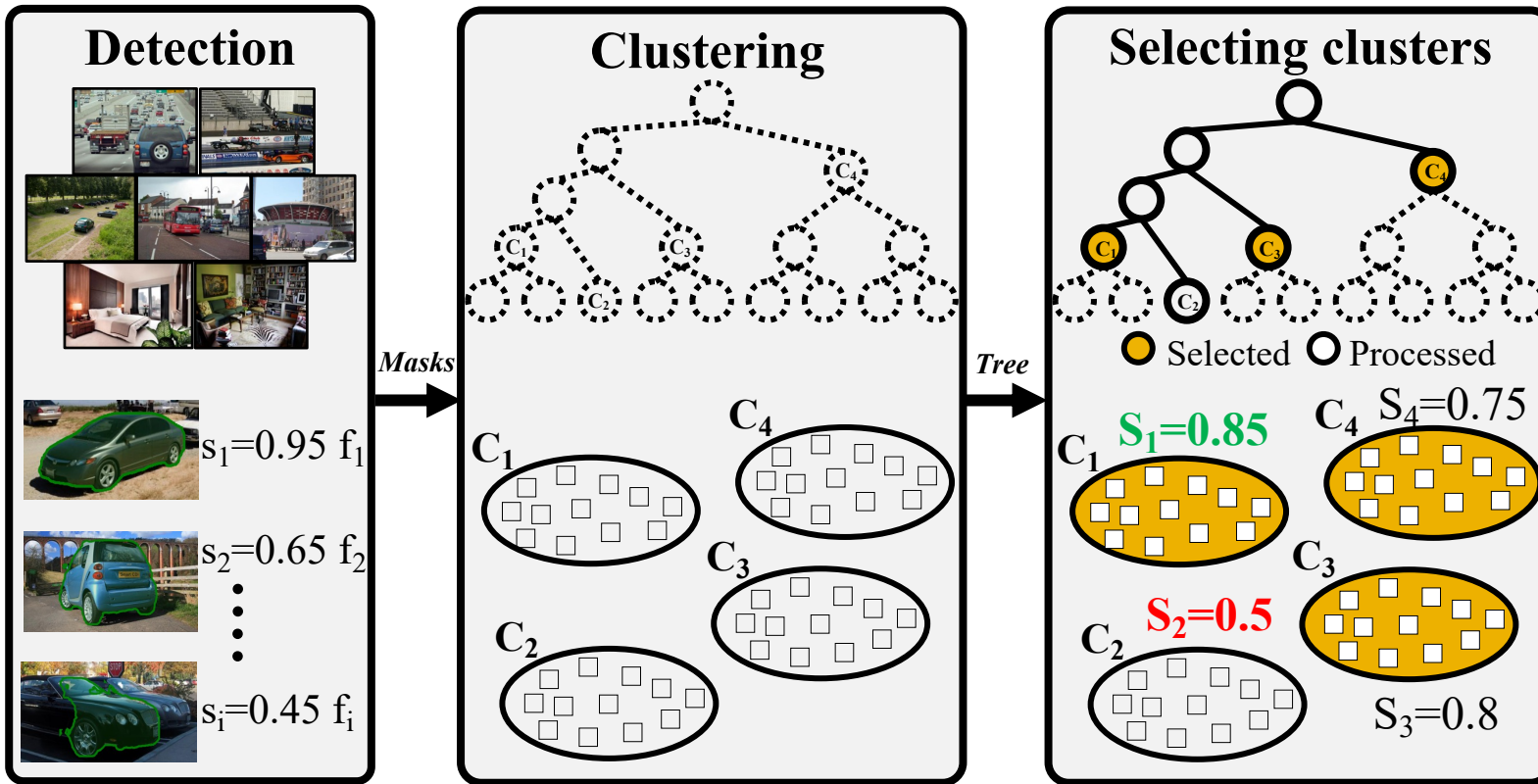
Masks



Tree

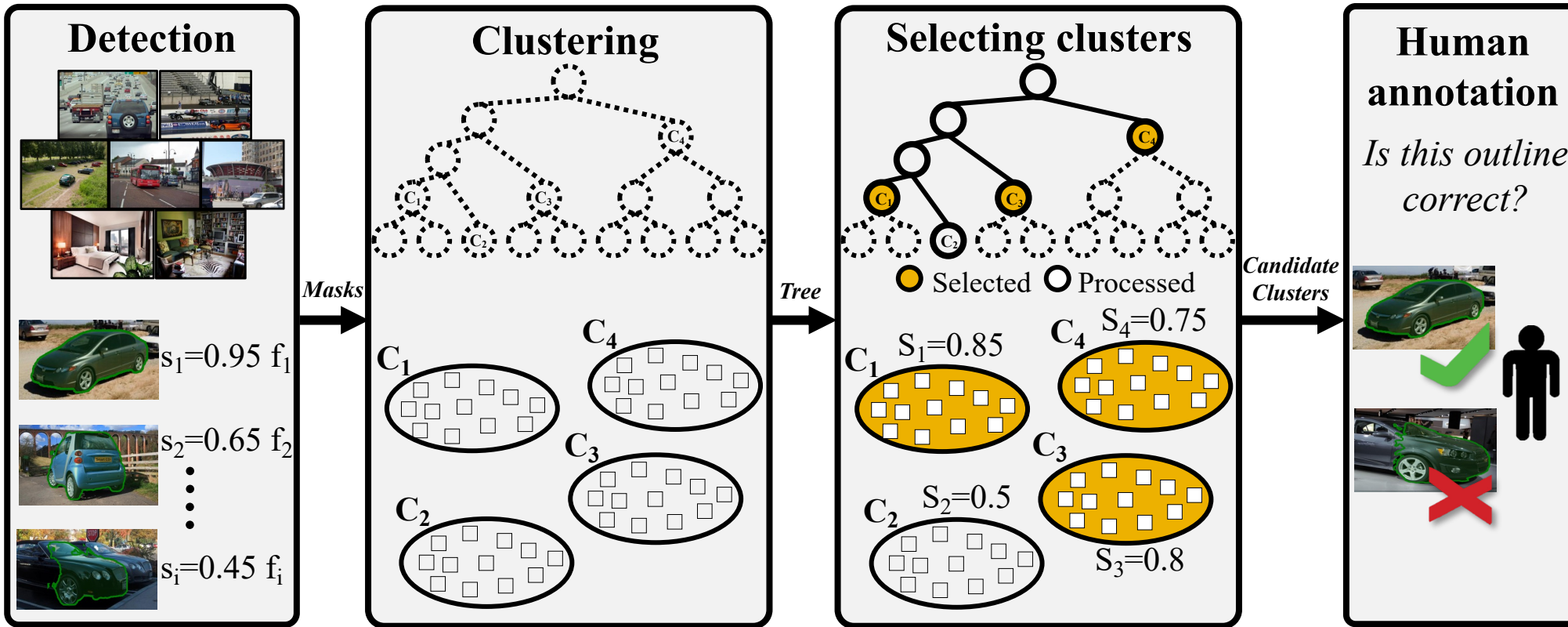


Method Overview

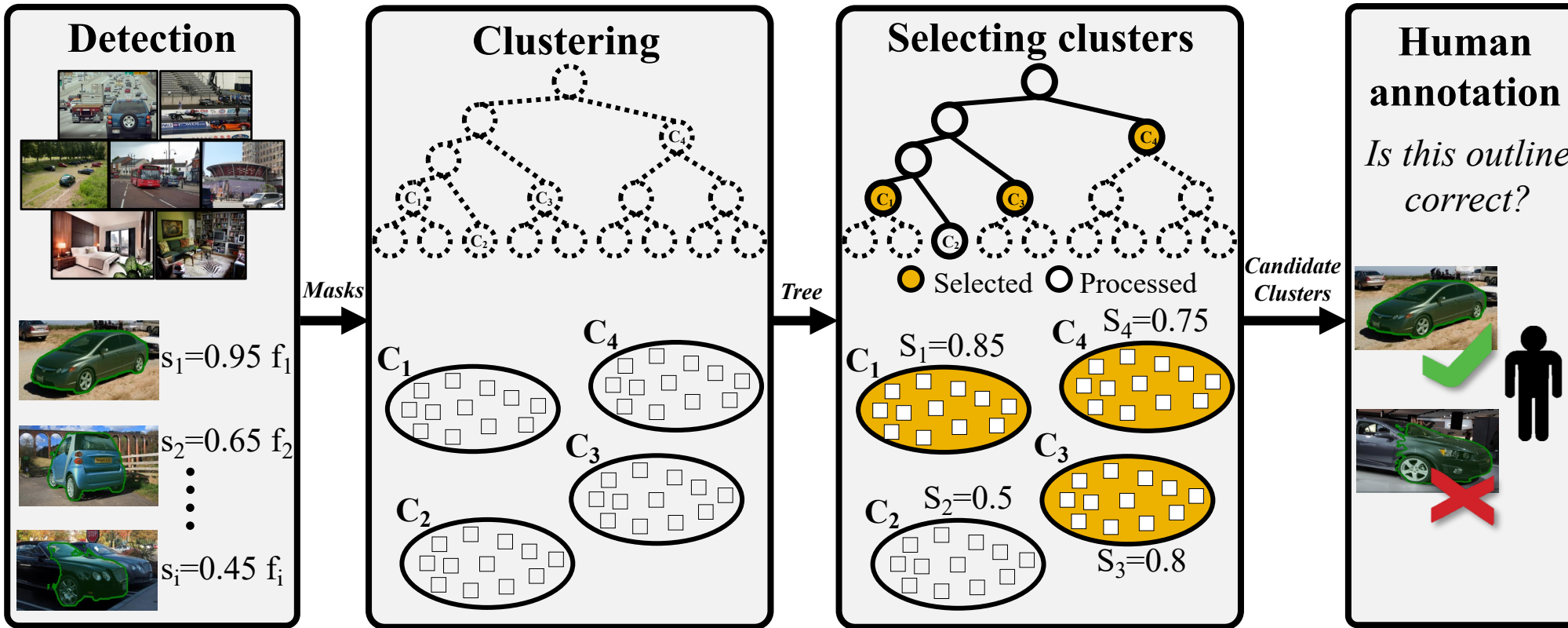


- *Prioritize search over cluster with **high** estimation score*
- *Split and Skip clusters with **low** score without annotation*

Method Overview



Method Overview



Minimum number of annotated masks per cluster

Method Overview

Detection

$s_1=0.95 f_1$
 $s_2=0.65 f_2$
 \vdots
 $s_i=0.45 f_i$

Masks

Clustering

C_1 C_2 C_3 C_4

Tree

Selecting clusters

● Selected ○ Processed

C_1 $S_1=0.85$ C_4 $S_4=0.75$
 C_2 $S_2=0.5$ C_3 $S_3=0.8$

Candidate Clusters

Human annotation

Is this outline correct?

✓
✗

\tilde{Q}


Label propagation

● Accept ● Reject ○ Split

C_1 $\tilde{Q}_1=1.00$ C_4 $\tilde{Q}_4=0.00$
 C_3 $\tilde{Q}_3=0.50$

Method Overview

Detection



Masks

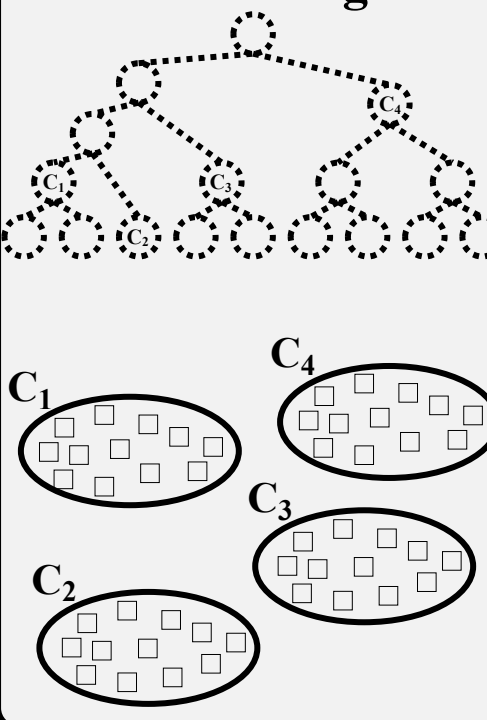
$s_1=0.95 f_1$

$s_2=0.65 f_2$

...

$s_i=0.45 f_i$

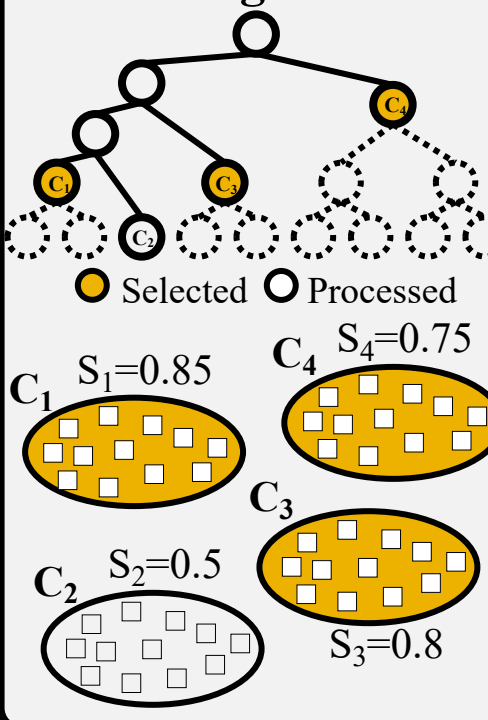
Clustering



Tree

C_1 C_2 C_3 C_4

Selecting clusters



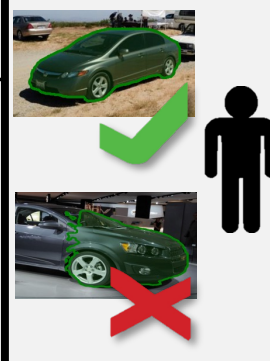
Candidate Clusters

● Selected ○ Processed

$C_1 S_1=0.85$ $C_2 S_2=0.5$ $C_3 S_3=0.8$ $C_4 S_4=0.75$

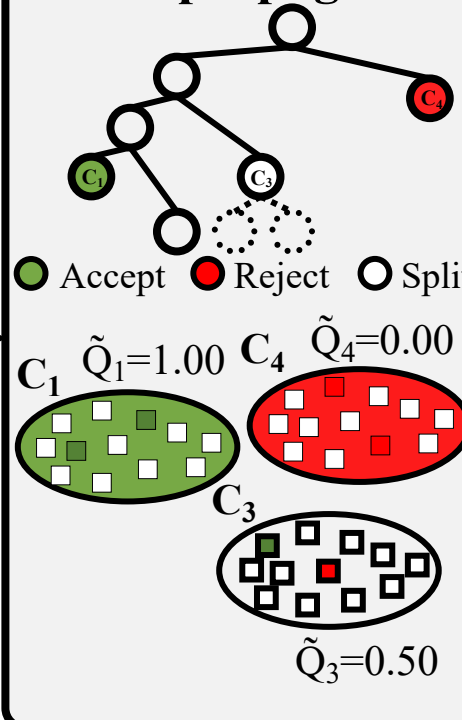
Human annotation

Is this outline correct?



Candidate Clusters

Label propagation



\tilde{Q}

● Accept ● Reject ○ Split

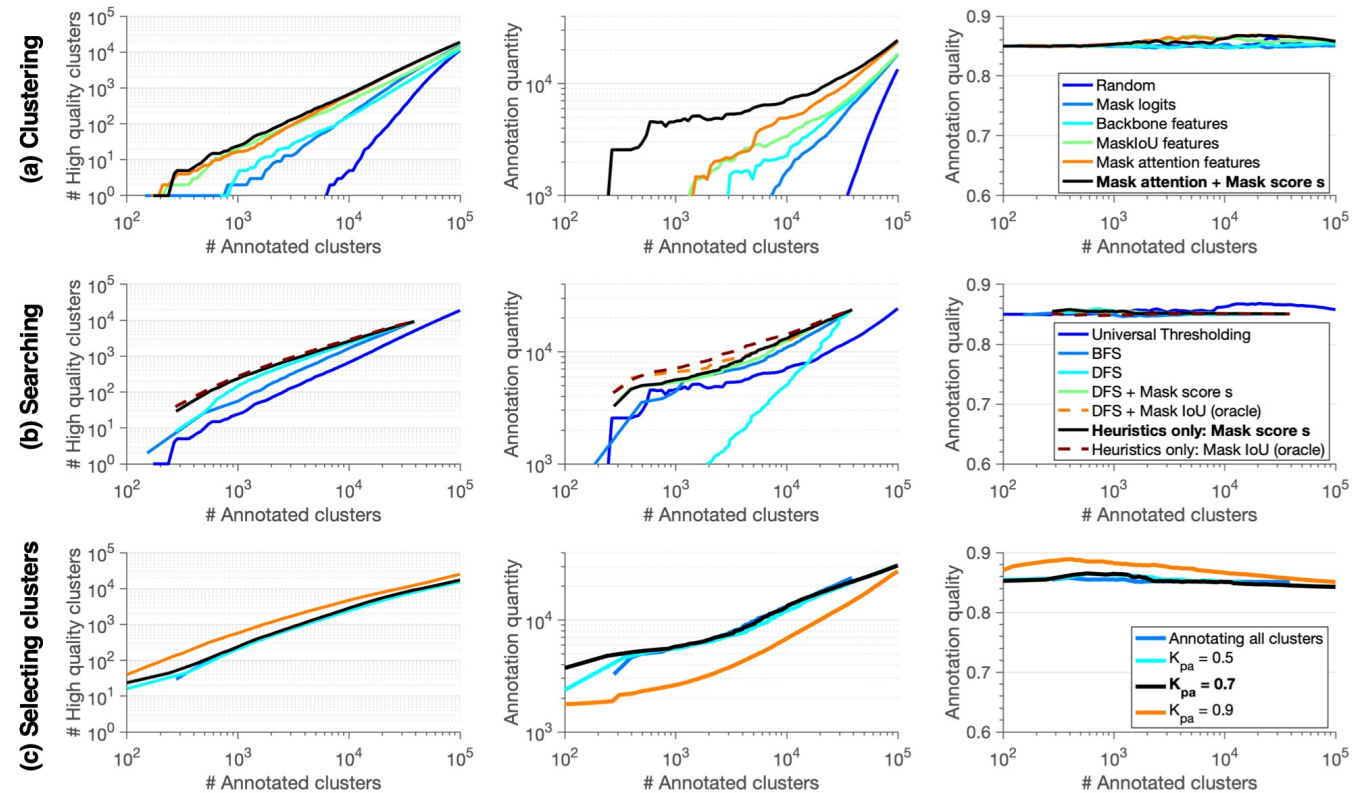
$C_1 \tilde{Q}_1=1.00$ $C_2 \tilde{Q}_2=0.50$ $C_3 \tilde{Q}_3=0.50$ $C_4 \tilde{Q}_4=0.00$

Propagate quality and prune tree

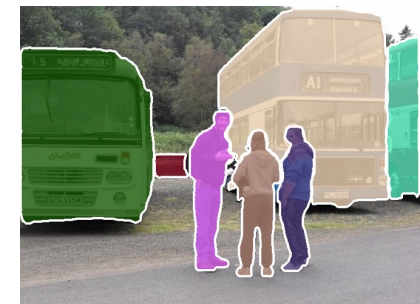
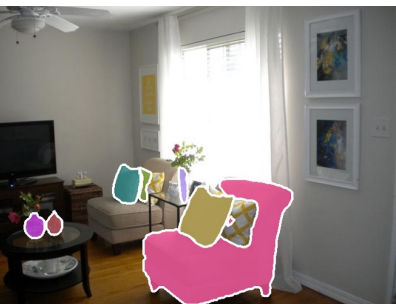
Simulated experiments on ADE20K, COCO, OpenImages



- Feature representation
- Searching the tree
- Selecting clusters
- Size of the initial training set
- Number of samples
- Quality of clusters
- IoU quality of masks



Places Annotations

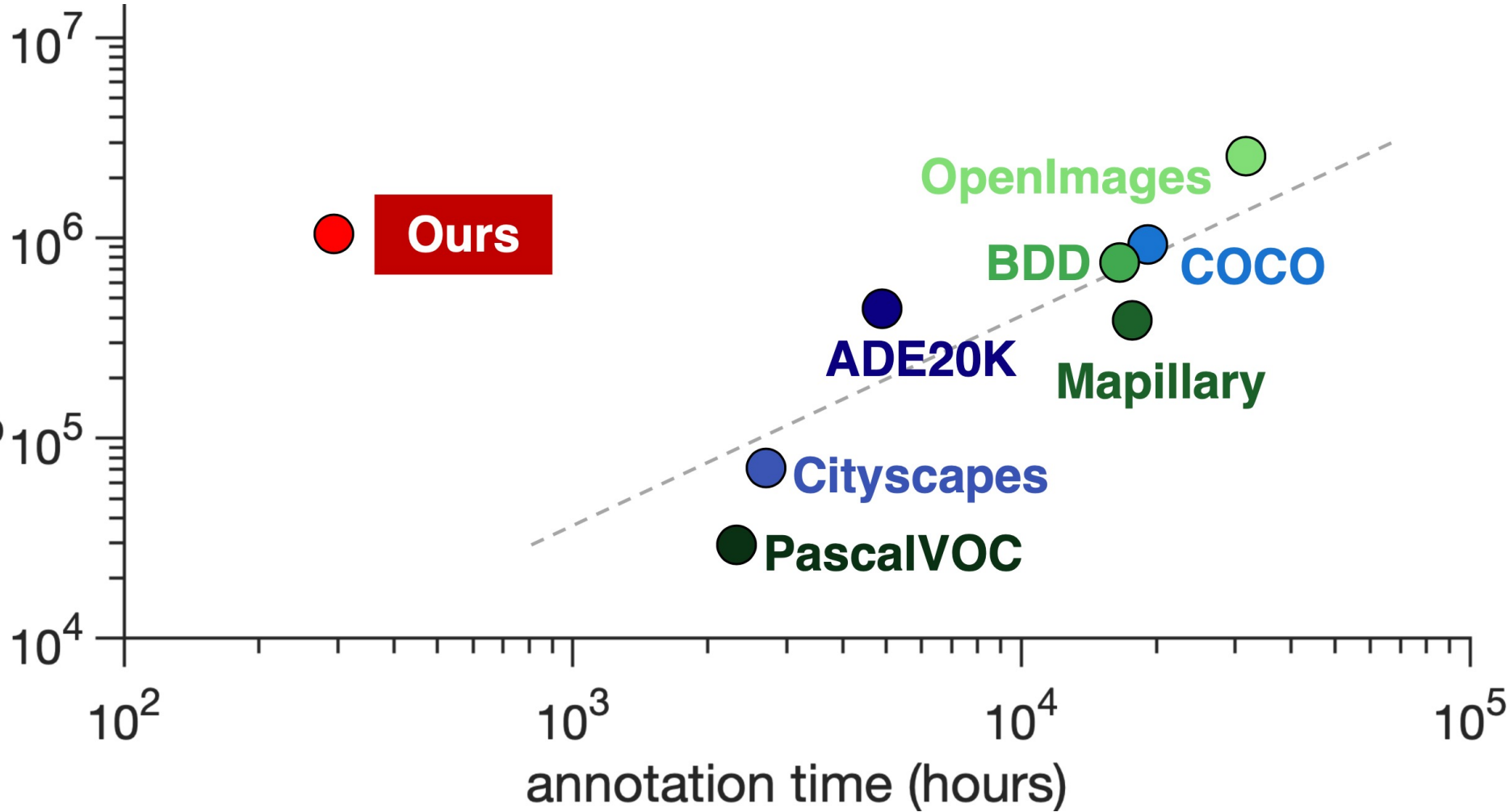


1M annotation masks

Places Annotations: Quantity and cost



of segmentation masks



Annotation speed-up in perspective

Manual annotation: 1 car

Proposed method: 76 cars !!!



Places Annotations: Quality and usefulness

- **Quality**



Manually annotate 1,142 Places images:

mIoU=81.4%

- **Usefulness**

Training set	Annotation	Time	mAP
ADE 1K	Manual	217h	8.1
ADE 1.4K	Manual	290h	9.6
Places 1M	Ours	290h	12.7

Code, models and data available online!!!
<http://scaling-anno.csail.mit.edu/>

Thank you !!!

2021 **ICCV** OCTOBER 11-17
VIRTUAL