

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

Instance segmentation

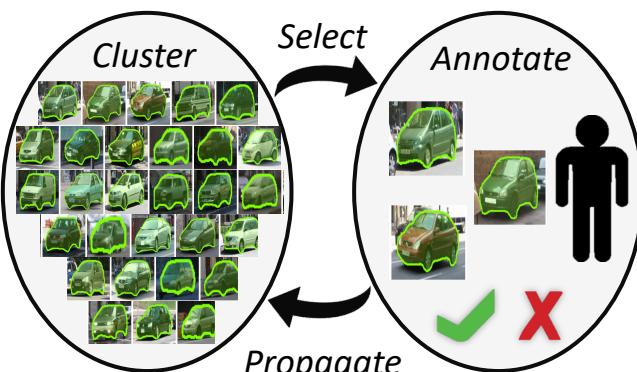
Obtaining training data



- Manual annotation: expensive (80s per object)
- Interactive segmentation: faster human interaction

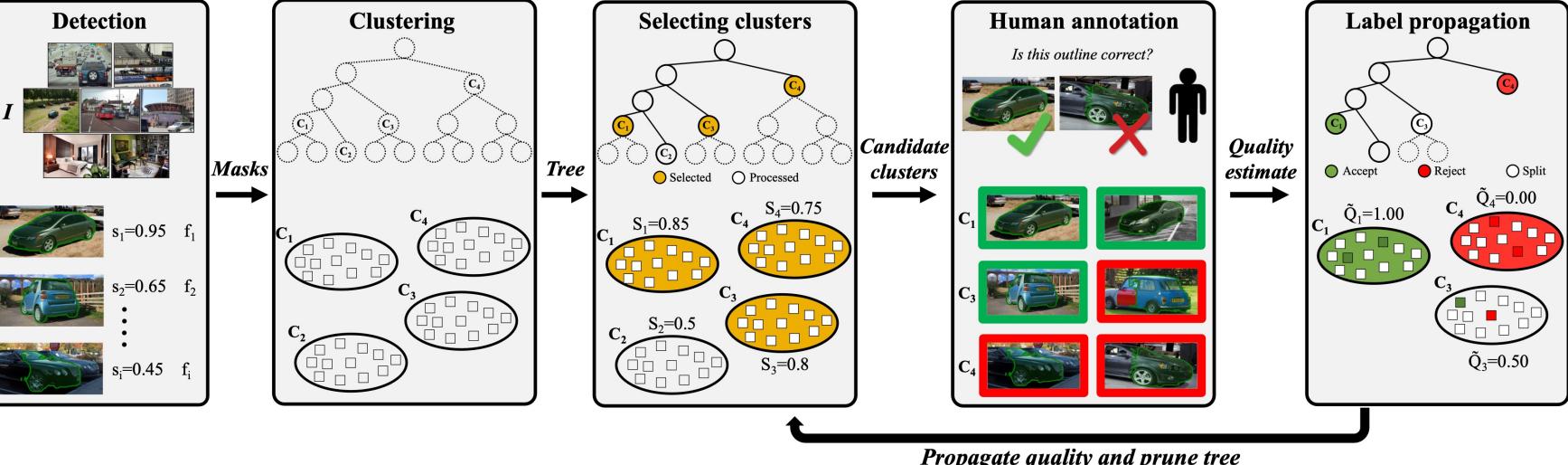
Cost linear to the number of instances

Our method



Cost linear to the number of clusters

Method



Propagate quality and prune tree

Experiments

Simulation (ADE20K)

- Train initial model on 10K images
 - Unlabeled pool of 10K images
 - Feature representation
 - Searching the tree
 - Selecting clusters
 - Initial training set
 - Number of samples
 - Quality of clusters
 - IoU quality of masks
- (a) Clustering: # of high quality clusters vs # of annotated clusters. Curves show performance for Random, Mask score s, Backbone features, MaskIoU features, Mask attention features, and Mask attention + Mask score s.
- (b) Searching: Annotation quantity vs # of annotated clusters. Curves show performance for Universal Thresholding, BFS, DFS, DFS + Mask score s, DFS + Mask IoU (ours), Mask attention features, and Heuristics early Mask IoU (crude).
- (c) Selecting clusters: Annotation quantity vs # of annotated clusters. Curves show performance for Annotating all clusters, $\rho_{\text{IoU}} = 0.5$, $\rho_{\text{IoU}} = 0.7$, and $\rho_{\text{IoU}} = 0.9$.

Real large-scale (Places)

- Quantity (1M masks)
- Usefulness (12.7% mAP)

	train	AP	Time (h)
ADE 1K	8.1	217h	
ADE 1.4K	9.6	290h	
Ours	12.7	290h	

- Quality (81.4% mIoU)



<http://scaling-anno.csail.mit.edu/>