



PartDistillation: Learning Parts from Instance Segmentation (TUE-PM-289)



Jang Hyun Cho



Philipp Krähenbühl



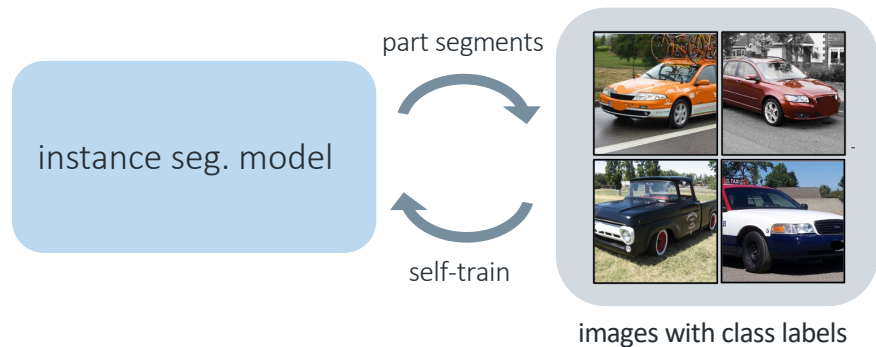
Vignesh Ramanathan

What is PartDistillation



PartDistillation is an unsupervised self-training method for object part segmentation task.

What is PartDistillation



For each object class, **PartDistillation** generate part segments as pseudo-labels and self-train.

What is PartDistillation



instance seg. model



No Part Labels

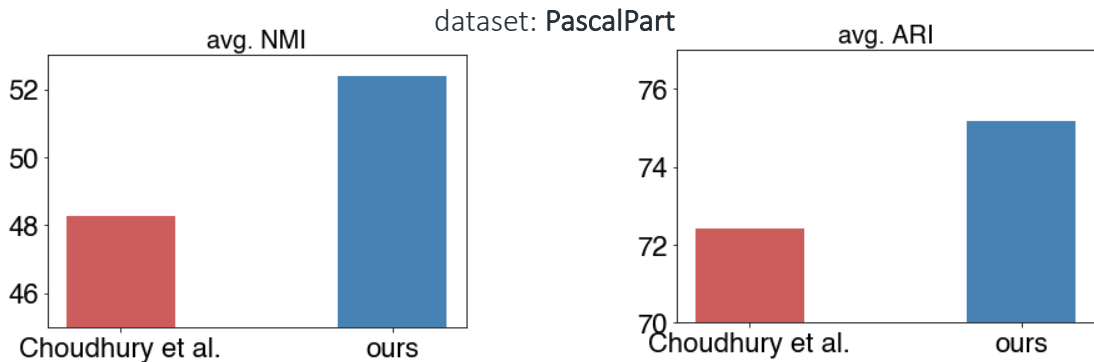
PartDistillation builds an accurate and diverse part segmentation model without any part labels.

Part Segmentation for over 10K Object Classes



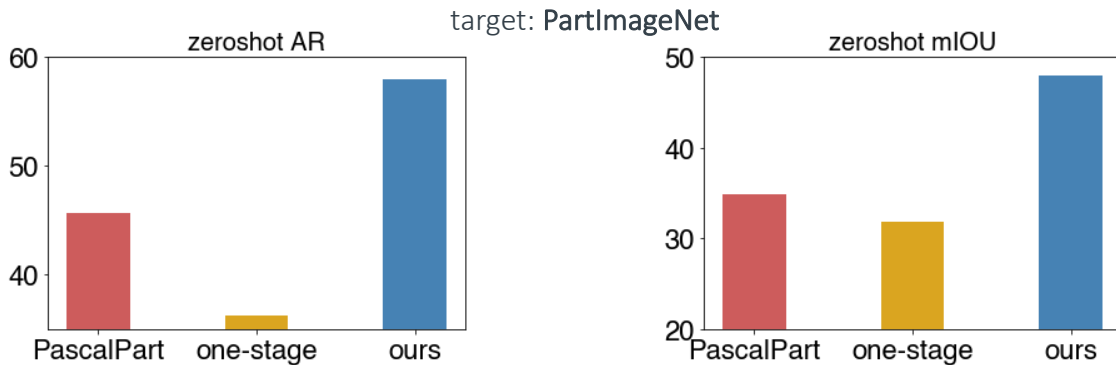
PartDistillation discovers novel parts for over 10K object categories.

Main Results



PartDistillation makes a SOTA unsupervised part segmenter.

Main Results

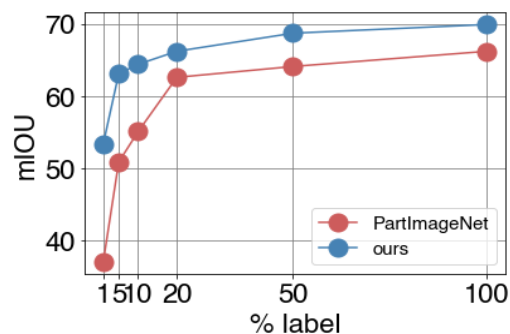
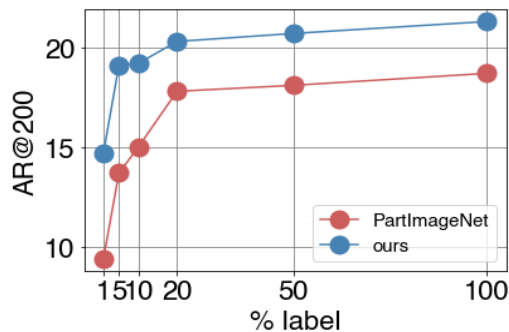


PartDistillation generalizes better than supervised training.

Main Results

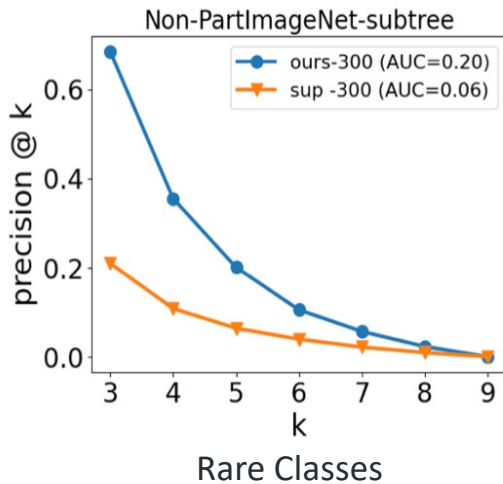
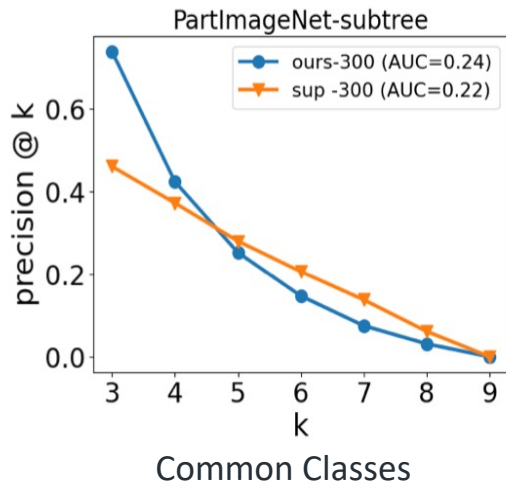


target: CityScapes



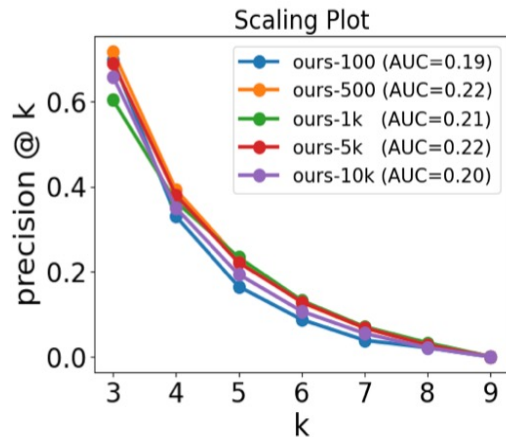
PartDistillation is more label efficient than supervised training.

Manual Evaluation Beyond Common Objects



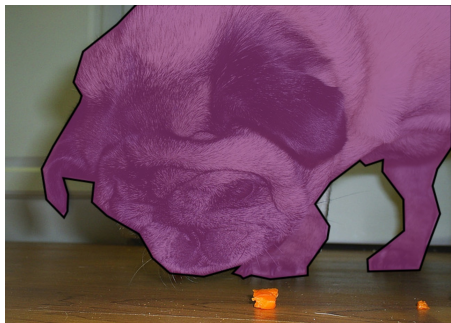
PartDistillation generalizes beyond common object categories.

Manual Evaluation Beyond Common Objects

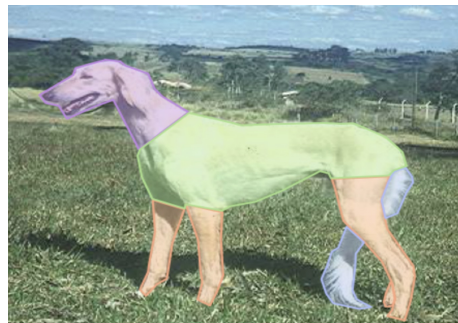


PartDistillation scales well to 10K categories.

Object Part Segmentation



Object Instance Segmentation



Object Part Segmentation



Part segmentation offers more fine-grained understanding of object layout.

Images from COCO and PartImageNet

Object Part Segmentation



Part segmentation benchmarks focus on only the common object categories.

Images from PascalPart, PartImageNet, and LVIS

Object Part Segmentation



LVIS (1200+ categories, 2M+ masks)



PartImageNet (150 categories, 100k masks)

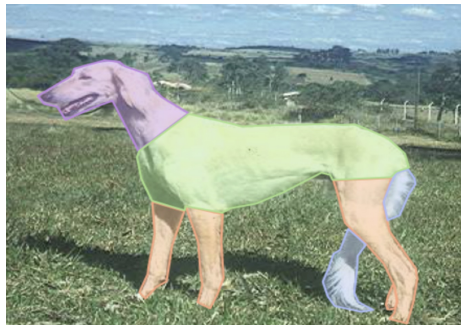
Existing datasets are orders of magnitude smaller than corresponding instance segmentation datasets.

Images from PascalPart, PartImageNet, and LVIS

Object Part Segmentation



PascalPart dog (9 parts)



PartImageNet dog (4 parts)

Parts are harder to detect, annotate, and properly define.

Images from PascalPart, PartImageNet, and LVIS

Why PartDistillation?



Supervised Learning Cannot Scale

Existing part segmentation datasets only cover < 100 coarse object categories, with different granularity of parts.

Supervised Models Do Not Generalize

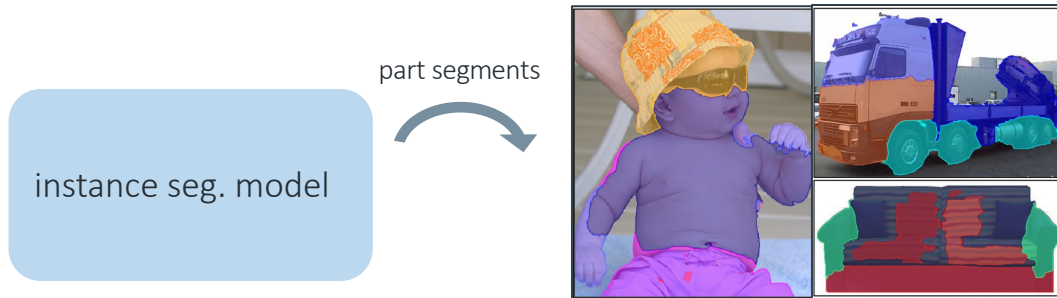
Supervised models do not generalize well across different object classes (e.g., dog legs → human legs)

Why PartDistillation?



PartDistillation addresses these problems through large-scale unsupervised learning.

Part from Instance Segmentation Model



Part from Instance Segmentation Model



non-transformer

Part Segments

Given an instance segmentation mask, **cluster** pixel-level features within the mask into K **part segments**.

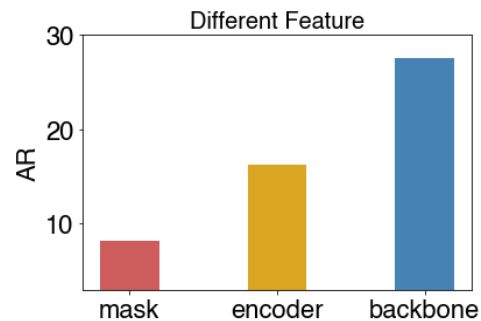
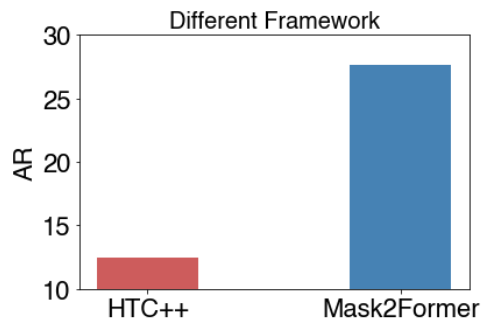
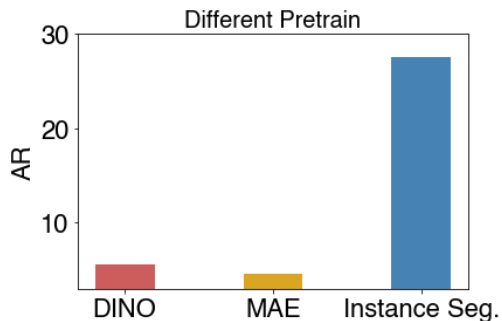


transformer

Observation

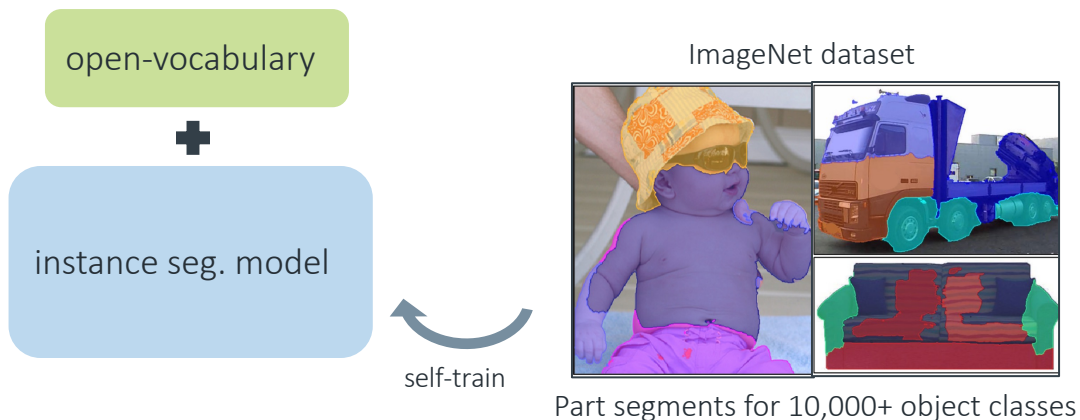
A **transformer** trained for **instance segmentation** task contains surprisingly accurate **part** information.

Part from Instance Segmentation Model



Mask2Former trained for instance segmentation with Swin Transformer backbone

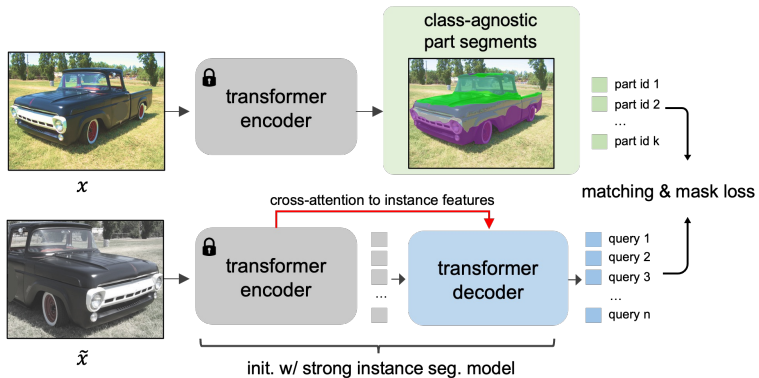
Part from Instance Segmentation Model



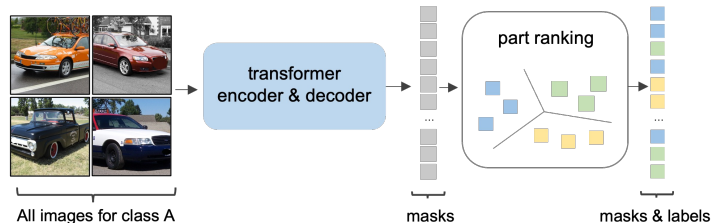
PartDistillation



PartDistillation

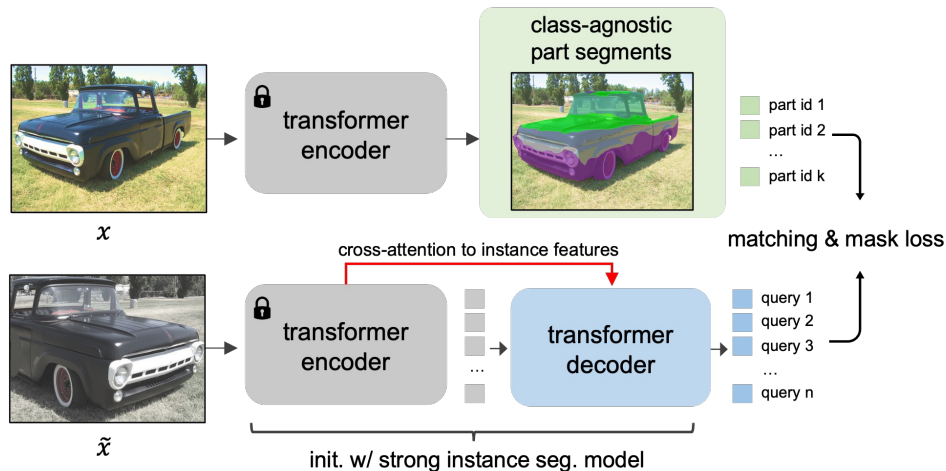


First Stage: **Part-proposal Learning**



Second Stage: **Part Ranking**

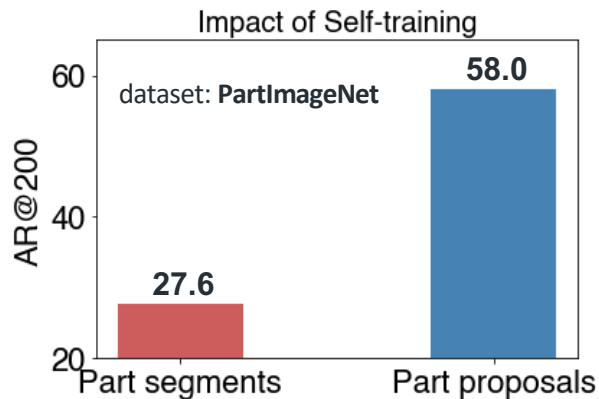
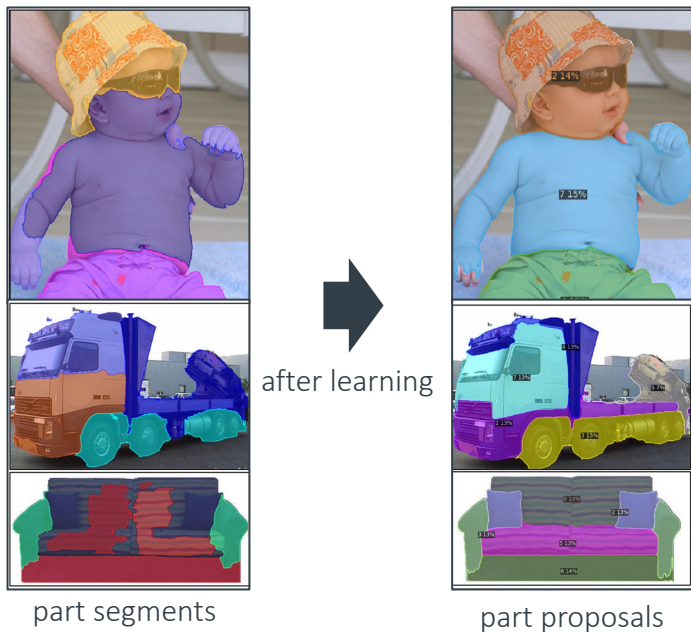
Part-proposal Learning



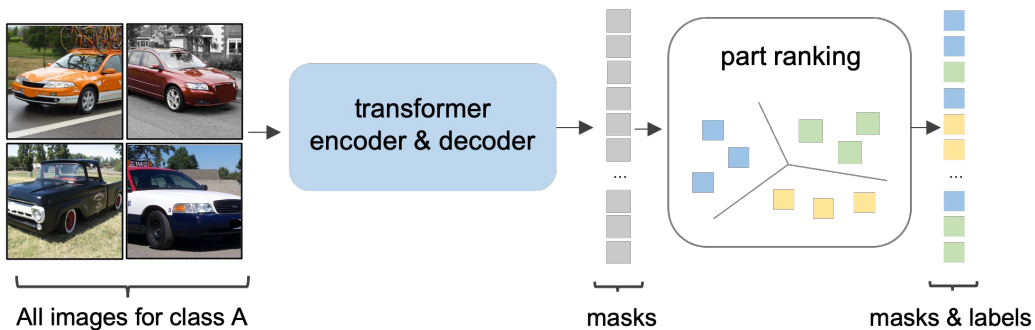
First Stage: **Part-proposal Learning**

Class-agnostic **part proposals** and **part-level features** with queries

Part-proposal Learning



Class-specific Part Ranking



Second Stage: Part Ranking

Class-specific **part segmentation** by part ranking over dataset
 Part-level features are **post-processed** and **clustered** for each object class
Part ranking function as cluster density estimate assigns labels
 Finally, **self-train** all-together with **class-specific part masks**!

Part Segmentation for over 10K Object Classes



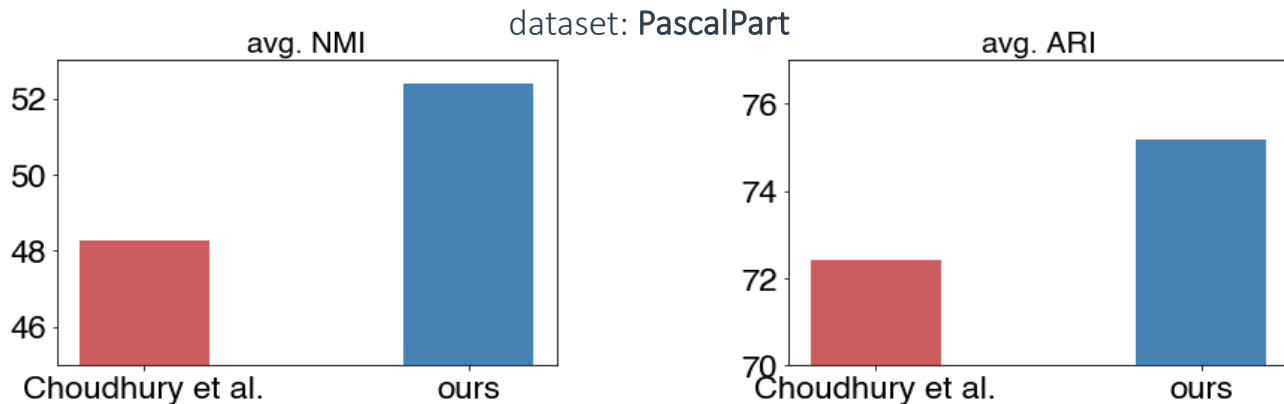
PartDistillation on ImageNet-21K

We filter-out ImageNet classes that do not have clear object definition (e.g., food, fungus, event, act, etc.). The remaining 12K classes are used for PartDistillation. Each object class has 8 part clusters over dataset.

Getting Instance Segmentation for 10K Objects

We use an open-vocabulary object detector Detic to localize for each object class by using the class name (e.g., "a chair.").

Unsupervised Benchmark

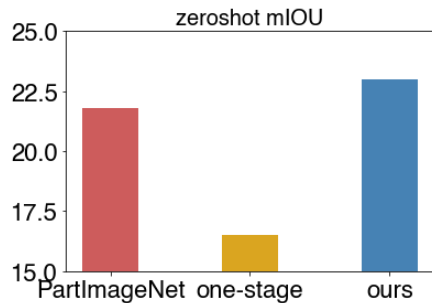
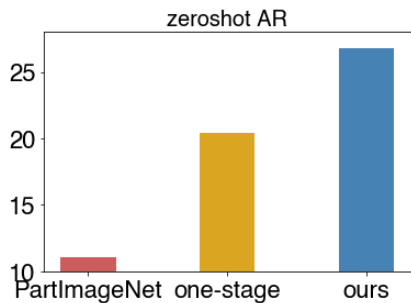


PartDistillation is a SOTA unsupervised part segmentation method.

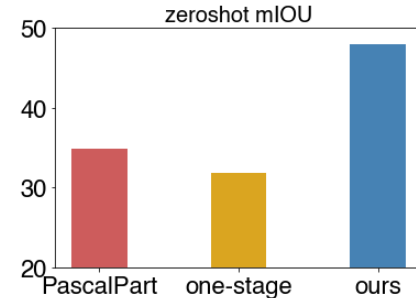
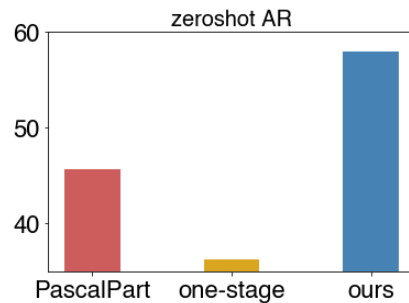
Zero-shot Benchmark



dataset: PascalPart



dataset: PartImageNet

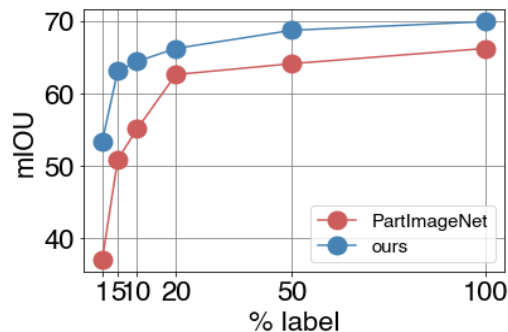
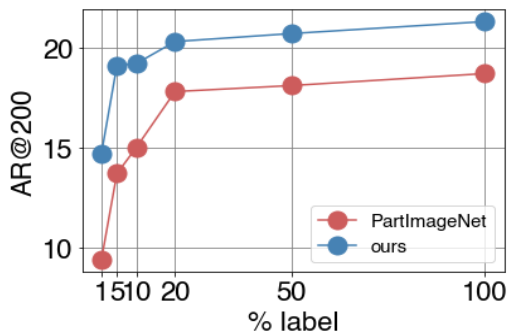


PartDistillation generalizes better than supervised training.

Few-shot Benchmark



dataset: CityScapes



PartDistillation is more label efficient than supervised training.

Manual Evaluation

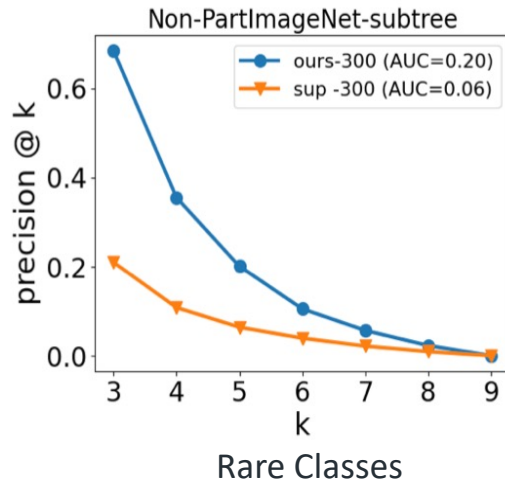
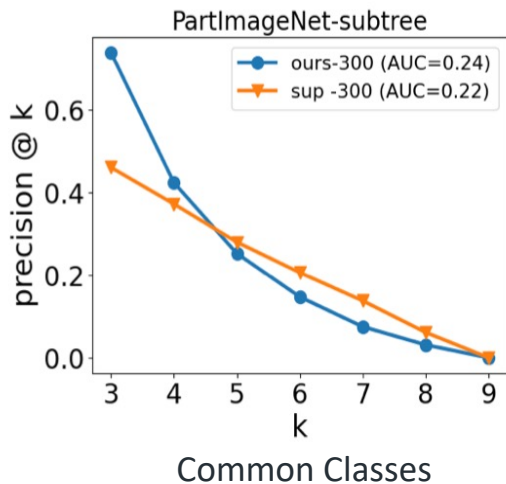


Existing benchmark can only evaluate on **common objects**

Conduct **manual evaluation** and compare **PartDistillation** as well as a SOTA **unsupervised** method and **supervised** baseline.

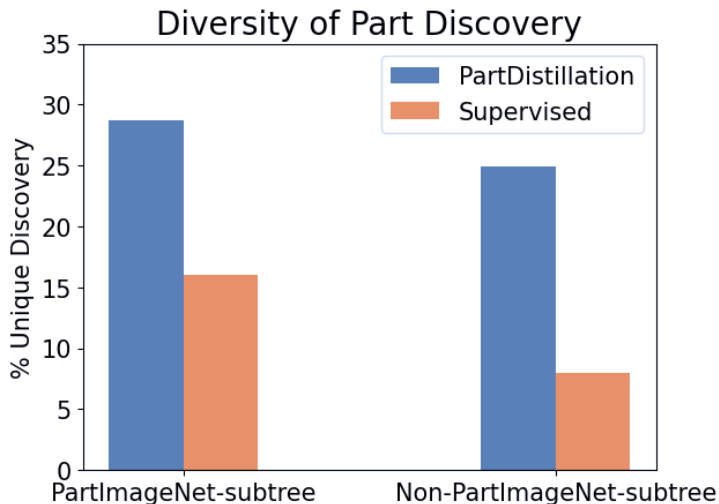
We measure **precision**, **diversity**, and **generality**.

Manual Evaluation Beyond Common Objects



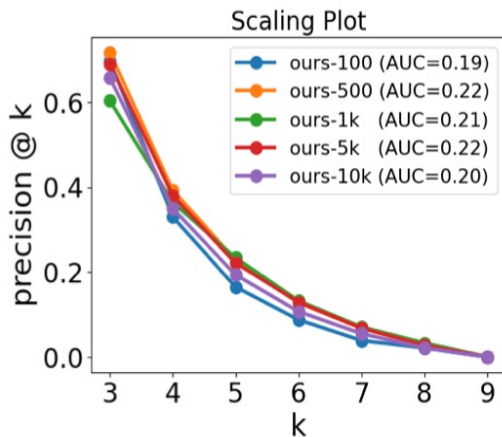
PartDistillation generalizes beyond common object categories.

Manual Evaluation Beyond Common Objects



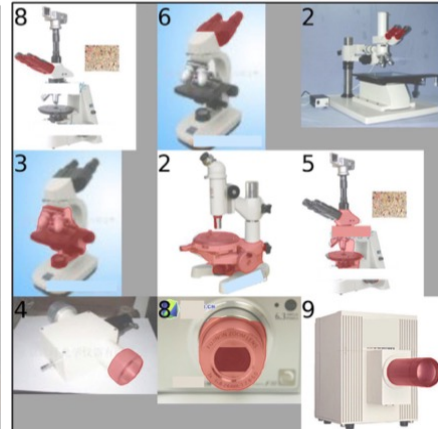
PartDistillation is more diverse than supervised training.

Manual Evaluation Beyond Common Objects



PartDistillation scales well to 10K categories.

Qualitative Results



Qualitative Results



Want to Know More?

When: Tuesday June 20, 4:30 pm – 6pm PDT

Where: West Building Exhibit Halls ABC #289



virtual poster



project page



Thank you!



Jang Hyun Cho



Philipp Krähenbühl



Vignesh Ramanathan