The Pursuit of Knowledge: Discovering and Localizing Novel Categories using Dual Memory

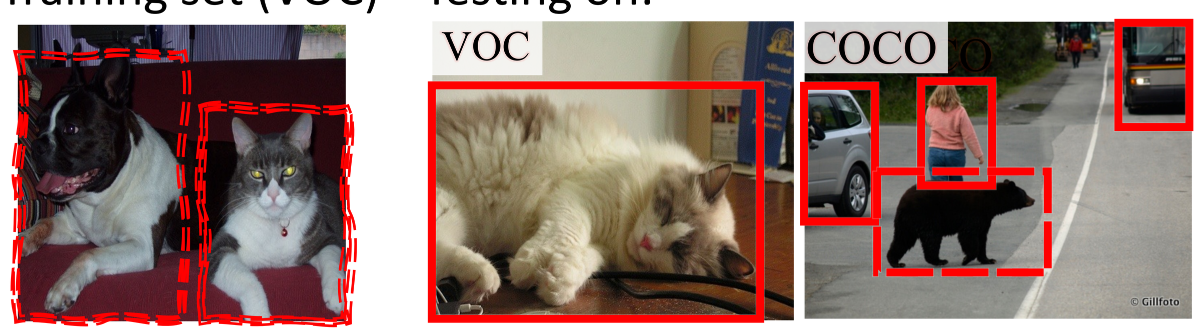
Sai Saketh Rambhatla[1], Rama Chellappa[1,2] and Abhinav Shrivastava[1]
[1]University of Maryland, College Park [2]Johns Hopkins University

2021 ICCV OCTOBER 11-17 VIRTUAL

Goal: Discover novel objects and learn models to detect them without human supervision.

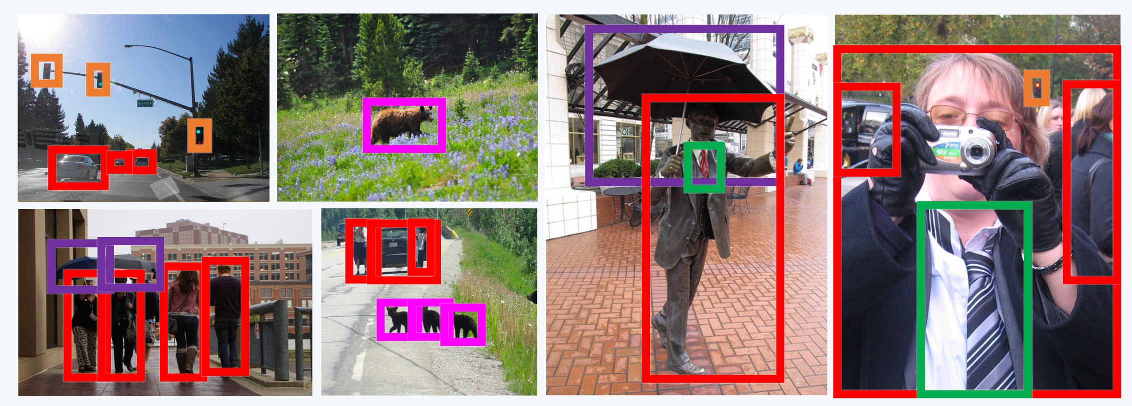
Shortcomings of standard supervised paradigm

Training set (VOC) Testing on:



Detectors trained on a labeled dataset (VOC)
Fail to detect novel objects (e.g., additional objects in COCO)

Our Discovery and Localization Benchmark & Framework

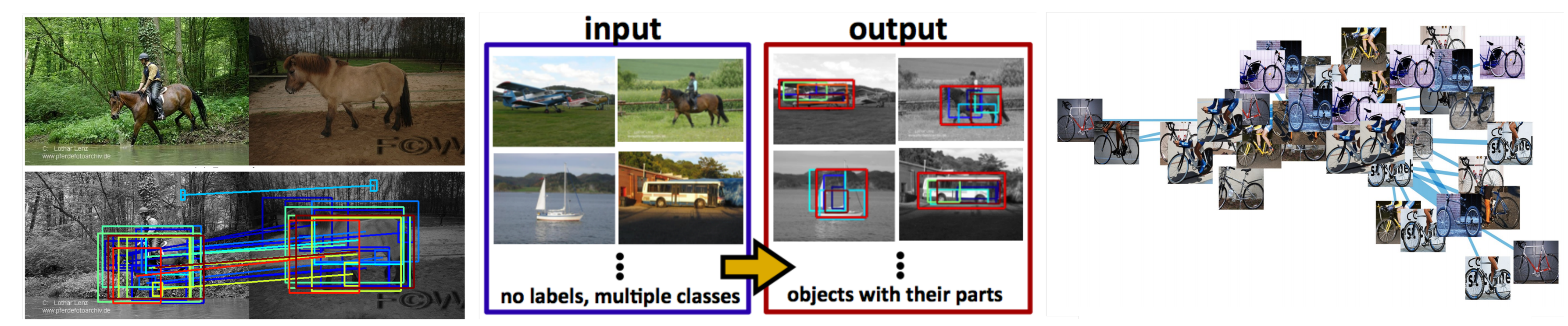


Discovered and localized examples of *unknown* classes (from COCO). In this illustration these correspond to (clockwise): traffic-light, tie, umbrella, bear

★ **Large-scale, realistic benchmark** for object discovery & localization

★ **Scalable** never-ending **in-the-wild** concept discovery framework

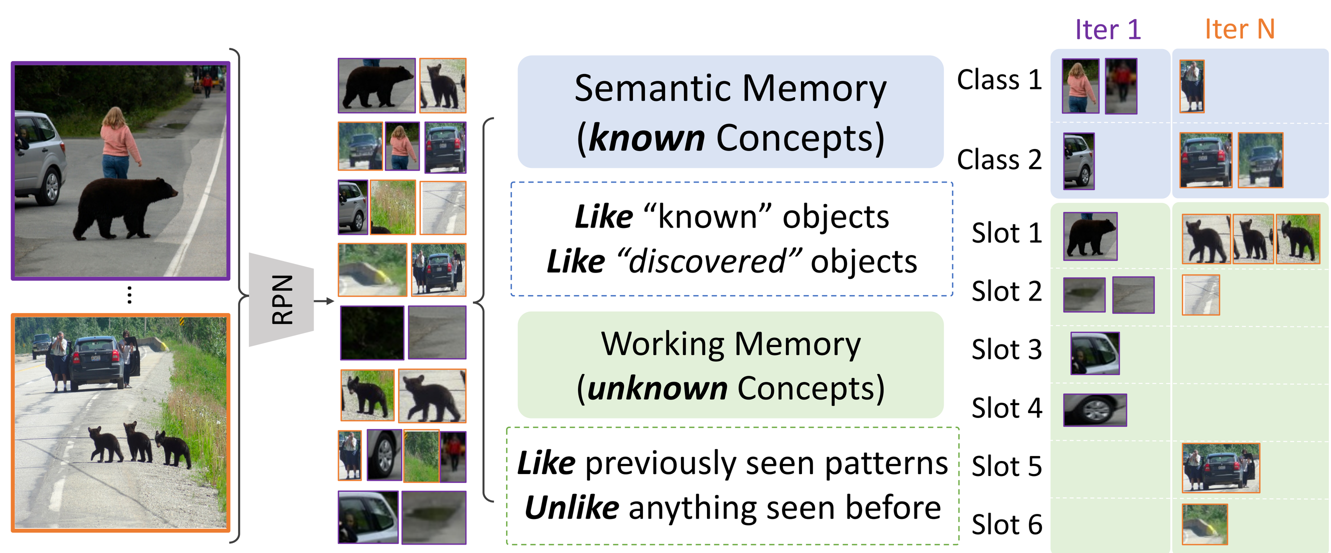
Shortcomings of contemporary discovery methods



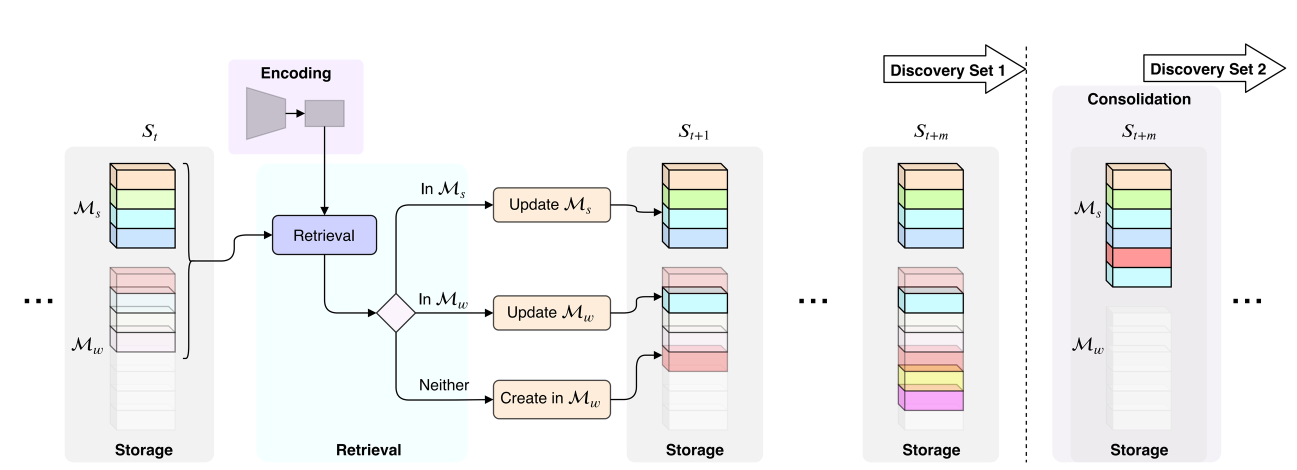
Tailored for co-segmentation/localization Not scalable

[Cho et al., CVPR 2015] [Vo et al., CVPR 2019] [Vo et al., ECCV 2020]

Dual Memory Framework for Unsupervised Object Discovery



Framework: Iterative, Online, and Scalable



Encoding: Region Proposals

Storage: Dual Memory

Retrieval: Decide seen/unseen?

Dual Memory

Contains clusters ("slots"), represented as a Centroid or a Classifier

Centroid
- Fast updates
- Inaccurate
- Cos. similarity

Classifier
- Slow updates
- Accurate
- Cls. score

Semantic Memory
- Long-term memory
- "Semantic Prior" init.
- Reliable associations
- Infrequent updates
- ➤ Classifier

Working Memory
- Short-term memory
- Null init.,
- Unreliable associations
- Frequent updates
- ➤ Centroid

Benchmark Details and Results

Benchmark

Labeled datasets:
- ImageNet
- Pascal VOC

In-the-wild discovery dataset: COCO

Salient features:
- Large-scale
- Different distribution of labeled & discovery set
- Localization

Metrics

Our benchmark:
- Purity vs. Coverage
- mAP for learned detectors
- # of objects discovered

Other methods*:
- CorLoc
- CorRet
- Det-Rate

*None evaluates discovery performance

Large-scale Object discovery on the entire COCO train2014 (80k images). Comparisons with scalable clustering methods using AuC for *unknown* classes.

Method	AuC@0.5	AuC@0.2	#disc. objs
K-means	3.34	7.23	42
FINCH	3.03	6.99	42
Ours	**3.60**	**9.11**	**46**

Smaller-scale Object discovery on subsets of COCO train2014. Comparison with contemporary discovery methods using AuC for *unknown* classes.

Method	Conf.	#imgs.	CorLoc	CorRet	DetRate
Vo et. al	CVPR'19	2.5k	6.62	80.00	4.73
Vo et. al†	CVPR'19	2.5k	6.34	70.00	5.17
Ours		2.5k	43.00	64.22	48.56
Vo et. al	ECCV'20	20k	15.77	100	11.56
Ours		20k	41.41	64.60	46.81

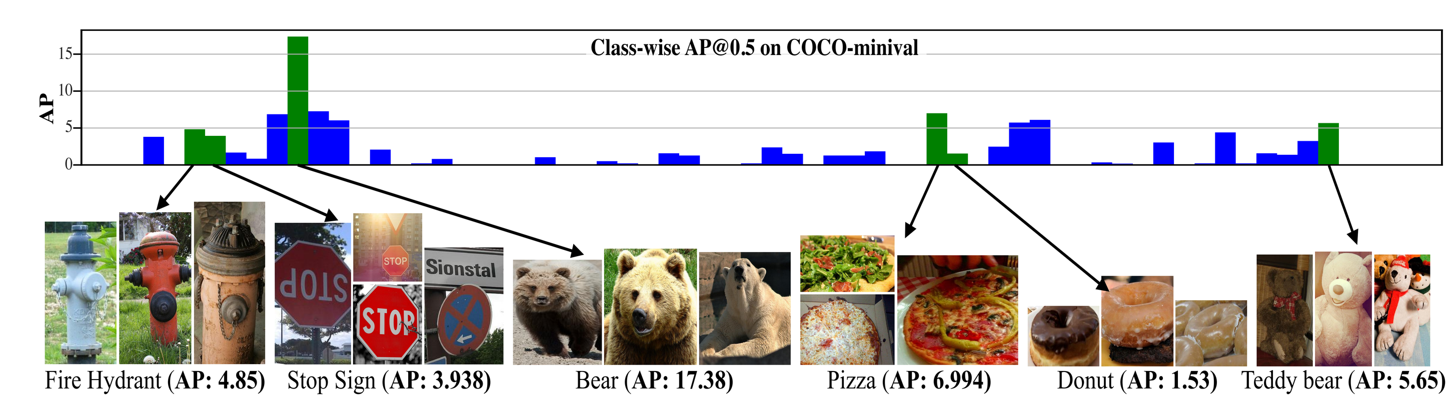
†: OSD with ResNet-101 Faster R-CNN proposals and classification-head features (same as Ours).

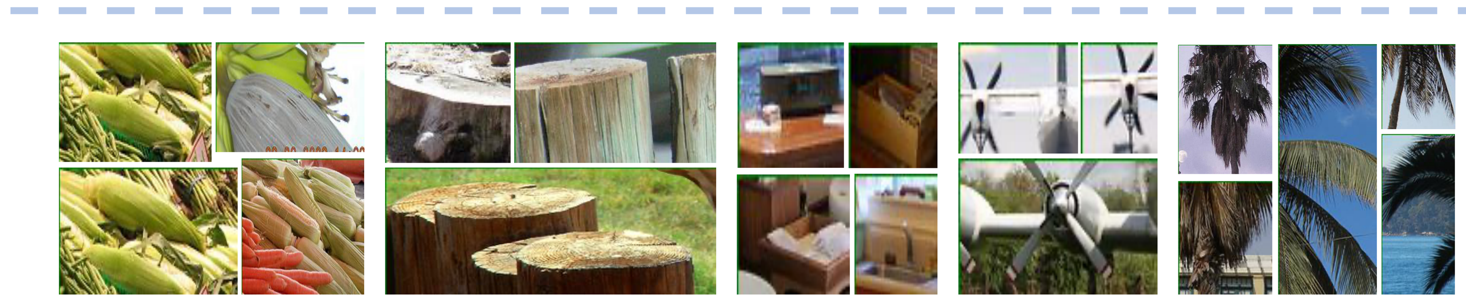
Detection performance (mAP) for object detectors on COCO minival, trained using *oracle* labels for clusters.

Classes	GT-IoU: 0.5		GT-IoU: 0.2	
	AP@0.5	AP@0.2	AP@0.5	AP@0.2
All (80)	2.69	4.44	2.62	4.37
Novel (60)	1.87	3.50	1.76	3.42
Novel†	5.23	6.47	5.45	6.40

†: mAP of classes with AP greater than chance.

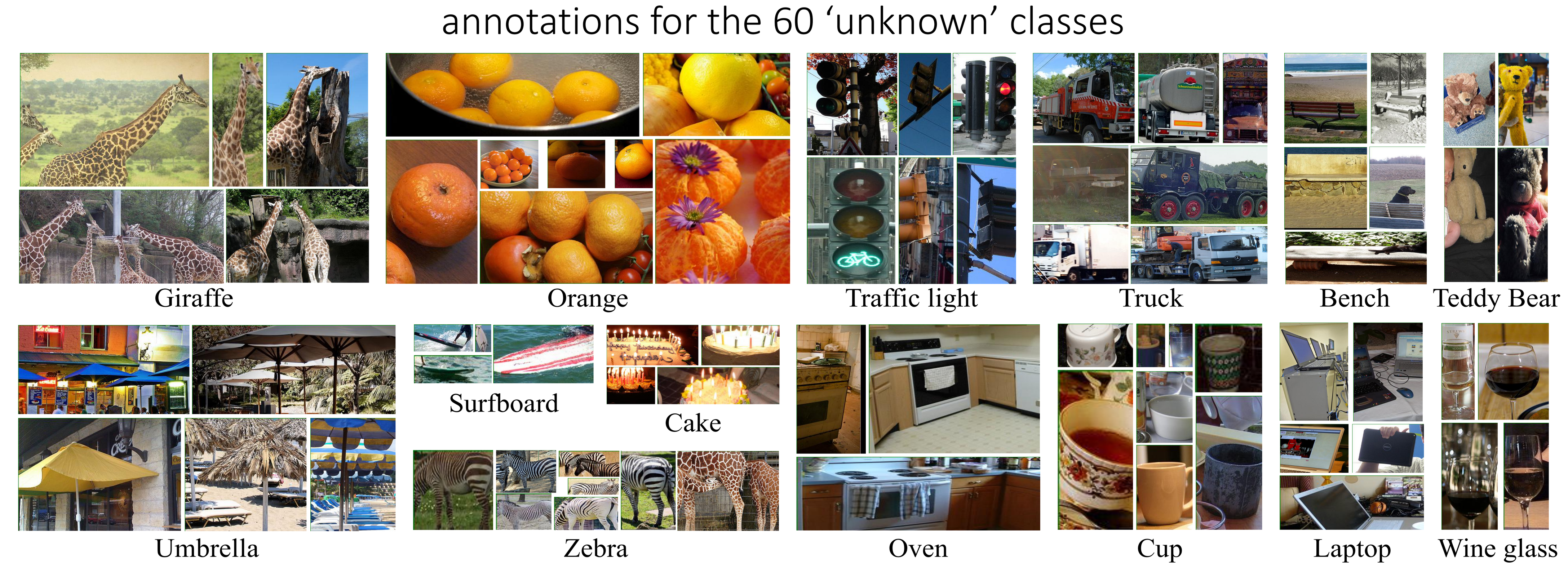
Sample detections and class-wise AP on COCO minival using our object detectors trained novel classes using *oracle* labels.





Concepts discovered by our approach which we cannot evaluate since they are unlabeled

Concepts discovered by our approach which we can evaluate using the ground-truth annotations for the 60 'unknown' classes



Visit the website for more discovered objects: