Towards Open World Object Detection

K J Joseph[‡Δ], Salman Khan[Δ], Fahad Shahbaz Khan[Δ], Vineeth N Balasubramanian[‡]
[‡]Indian Institute of Technology Hyderabad, [Δ]MBZ University of AI, Abu Dhabi.
{cs17m18p100001,vineethnb}@iith.ac.in,{salman.khan,fahad.khan}@mbzuai.ac.ae

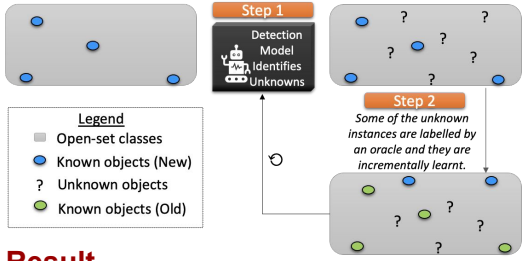
Highlights

We develop a practical extension of object detection problem which tasks the model to identify unknowns and incrementally learn them. Our methodology is based on:

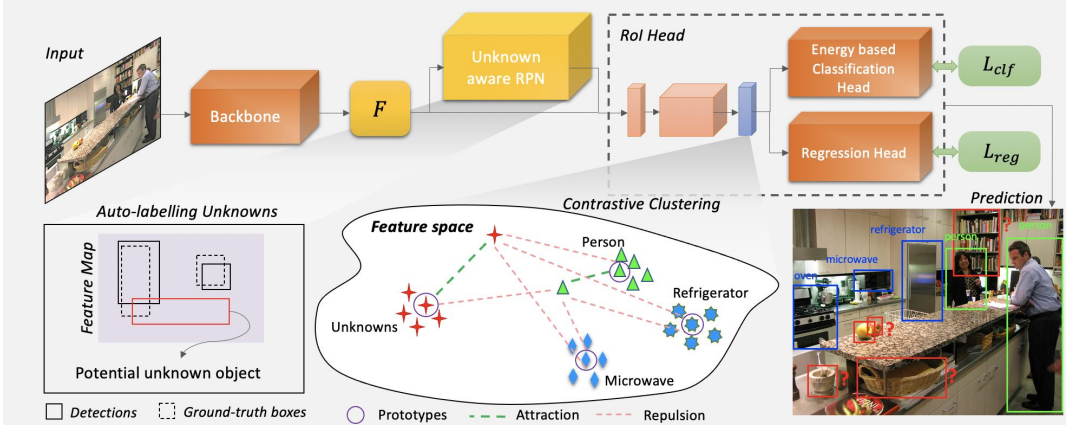
- Contrastive clustering to separate unknowns
- Unknown-aware RPN
- Energy based Unknown Identifier

We introduce experimental settings and evaluation protocols for the proposed problem.

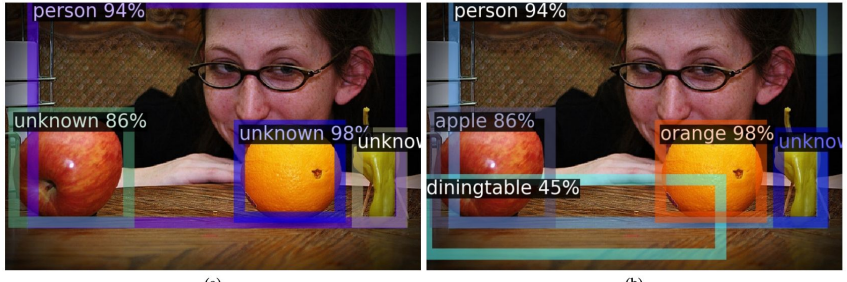
Our Novel Problem Setting



ORE: Open world object detector



Qualitative Result



(a) (b)

Detections on the left are produced by our method after learning a set of few classes which does not include edibles. We are able to identify them and correctly labels them as *unknown*. When the model is eventually taught to detect fruits, these instances are labelled correctly as in (b); without forgetting to detect *person*. An unidentified class instance still remains, and is successfully labelled *unknown*.

Quantitative Results

Task IDs (→)	Task 1			Task 2					Task 3					Task 4		
	WI	A-OSE	mAP (↑)	WI	A-OSE	mAP (↑)			WI	A-OSE	mAP (↑)			mAP (↑)		
	(↓)	(↓)	Current known	(↓)	(↓)	Previously known	Current known	Both	(↓)	(↓)	Previously known	Current known	Both	Previously known	Current known	Both
Oracle	0.02004	7080	57.76	0.0066	6717	54.99	30.31	42.65	0.0038	4237	40.23	21.51	30.87	32.52	19.27	31.71
Faster-RCNN	0.06991	13396	56.16	0.0371	12291	4.076	25.74	14.91	0.0213	9174	6.96	13.481	9.138	2.04	13.68	4.95
Faster-RCNN + Finetuning	Not applicable as incremental component is not present in Task 1			0.0375	12497	51.09	23.84	37.47	0.0279	9622	35.69	11.53	27.64	29.53	12.78	25.34
ORE	**0.02193**	**8234**	**56.34**	**0.0154**	**7772**	52.37	25.58	**38.98**	**0.0081**	**6634**	37.77	12.41	**29.32**	30.01	13.44	**26.66**

Here we showcase how ORE performs on Open World Object Detection. Wilderness Impact (WI) and Average Open Set Error (A-OSE) quantify how ORE handles the unknown classes (gray background), whereas Mean Average Precision (mAP) measures how well it detects the known classes (white background). We see that ORE consistently outperforms the Faster R-CNN based baseline on all the metrics. Task 1 contains all 20 classes from Pascal VOC, while the remaining 60 classes from MS COCO are semantically grouped to form the next three tasks. Kindly refer to our paper for a more detailed analysis.