LEDetection: A Simple Framework for Semi-Supervised Few-Shot Object Detection

Phi Vu Tran

LexisNexis® RISK SOLUTIONS

How Do Unlabeled Data Boost Few-Shot Detection?

1) Unlabeled data *close the gap* on proposal recall between the fully and semi-supervised detectors in low-label regimes.

2) Unlabeled data *improve proposal quality* by producing diverse proposals (green boxes) on novel unseen objects {boat, bus, car}.

3) Proposal quality is *highly correlated* with semi-supervised few-shot object detection, as evaluated on the MS-COCO benchmark.

Learning with Consistent Multi-View Proposals

We introduce **SoftER Teacher** to capture superior representations with unlabeled data for effective few-shot detection via Proposal Consistency Learning.

Robust Detection From a *Single* Label (Red Inset)

Experimental Results: Balancing the Trade-off Between Novel Detection and Base Forgetting

VOC07 test – Split 1 Method	Backbone	Base AP_{50}	Base AR_{50}	Base AP_{50} (15 Classes) 1-Shot	5-Shot	10-Shot	Novel AP_{50} (5 Classes) 1-Shot	5-Shot	10-Shot	Overall AP_{50} (20 Classes) 1-Shot	5-Shot	10-Shot
MPSR* [ECCV'20]	R-101	80.8	–	61.5	69.7	71.6	**42.8**	55.3	61.2	56.8	66.1	69.0
Retentive R-CNN* [CVPR'21]	R-101	80.8	–	80.9	80.8	80.8	42.4	53.7	56.1	71.3	74.0	74.6
TFA [ICML'20]	R-101	80.8	–	77.6 ± 0.2	77.4 ± 0.3	77.5 ± 0.2	25.3 ± 2.2	47.9 ± 1.2	52.8 ± 1.0	64.5 ± 0.6	70.1 ± 0.4	71.3 ± 0.3
SoftER Teacher (Ours)	R-101	**85.0**	**93.1**	**83.3 ± 0.4**	**84.6 ± 0.2**	**84.9 ± 0.2**	35.5 ± 3.7	**60.3 ± 2.8**	**65.9 ± 1.6**	**71.3 ± 1.2**	**78.5 ± 0.8**	**80.2 ± 0.4**
Faster R-CNN (Our Impl.)	R-50	81.7	88.0	82.0 ± 0.2	82.4 ± 0.1	82.3 ± 0.1	27.9 ± 3.2	52.1 ± 2.1	58.2 ± 1.6	68.5 ± 0.8	74.9 ± 0.5	76.2 ± 0.4
Soft Teacher (Our Impl.)	R-50	**85.0**	90.9	**84.6 ± 0.4**	**85.2 ± 0.1**	**85.0 ± 0.1**	29.5 ± 4.4	53.8 ± 2.6	60.8 ± 1.5	**70.8 ± 1.2**	77.4 ± 0.7	79.0 ± 0.3
SoftER Teacher (Ours)	R-50	84.7	**92.3**	83.1 ± 0.2	84.4 ± 0.2	84.6 ± 0.2	**31.3 ± 4.3**	**57.8 ± 2.6**	**62.9 ± 1.7**	70.2 ± 1.2	**77.8 ± 0.7**	**79.2 ± 0.4**

* indicates results were reported on a single sample run (hence the lack of error bars), which may have been over-estimated due to the high variance of few-shot training samples.

Presented at the 27th International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain.