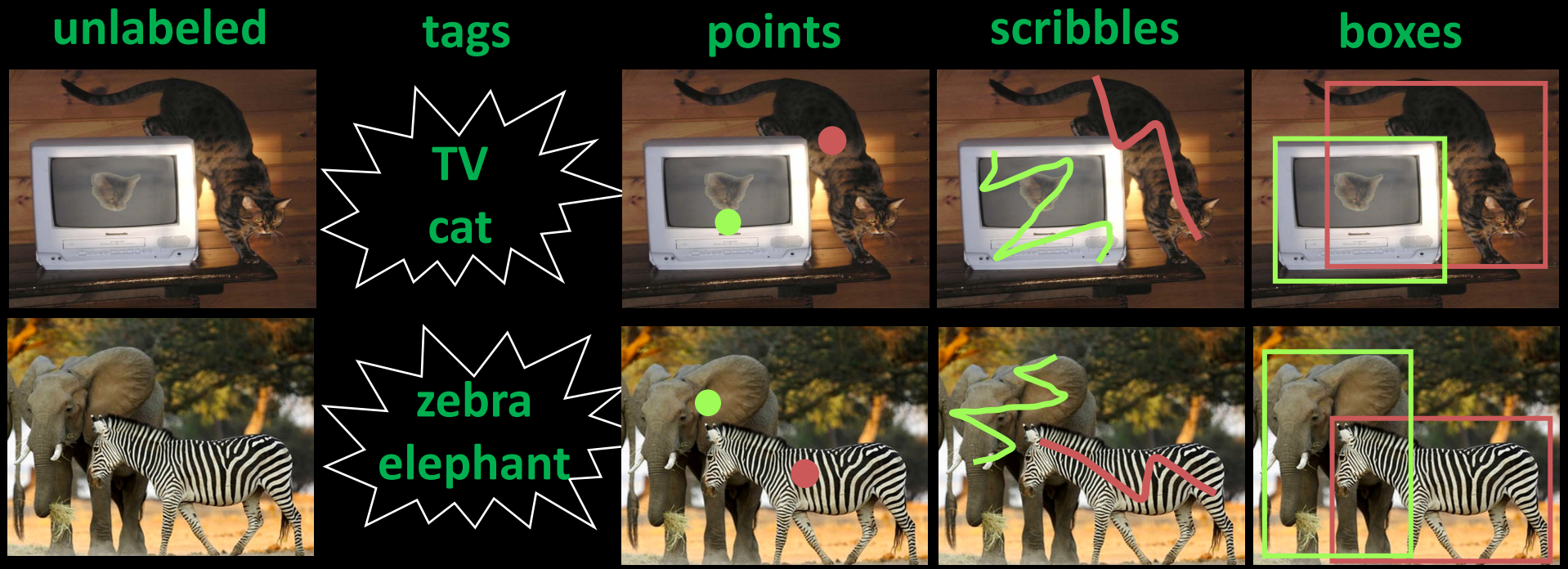
UFO^2: A Unified Framework towards Omni-supervised Object Detection

Zhongzheng Ren, Zhiding Yu, Xiaodong Yang, Ming-Yu Liu,
Alexander G. Schwing, Jan Kautz

ECCV 2020

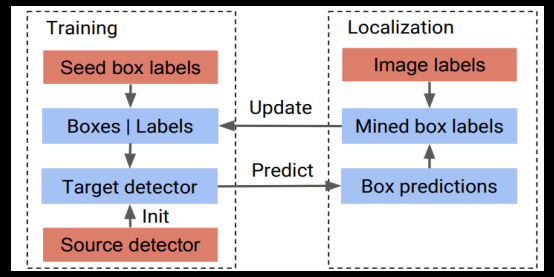
Omni-supervised Object Detection



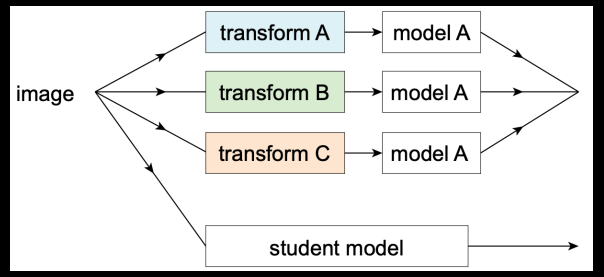
weakly-supervised — commonly used in e.g., segmentation — strongly-supervised

omni-supervised

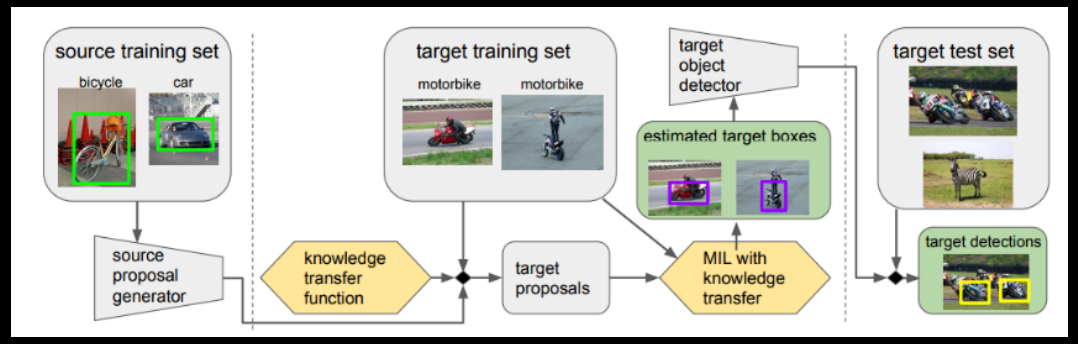
Omni-supervised Object Detection



(Gao et al., 2018)



(Radosavovic et al., 2019)



(Uijlings et al., 2018)

Semi-supervised Object Detection

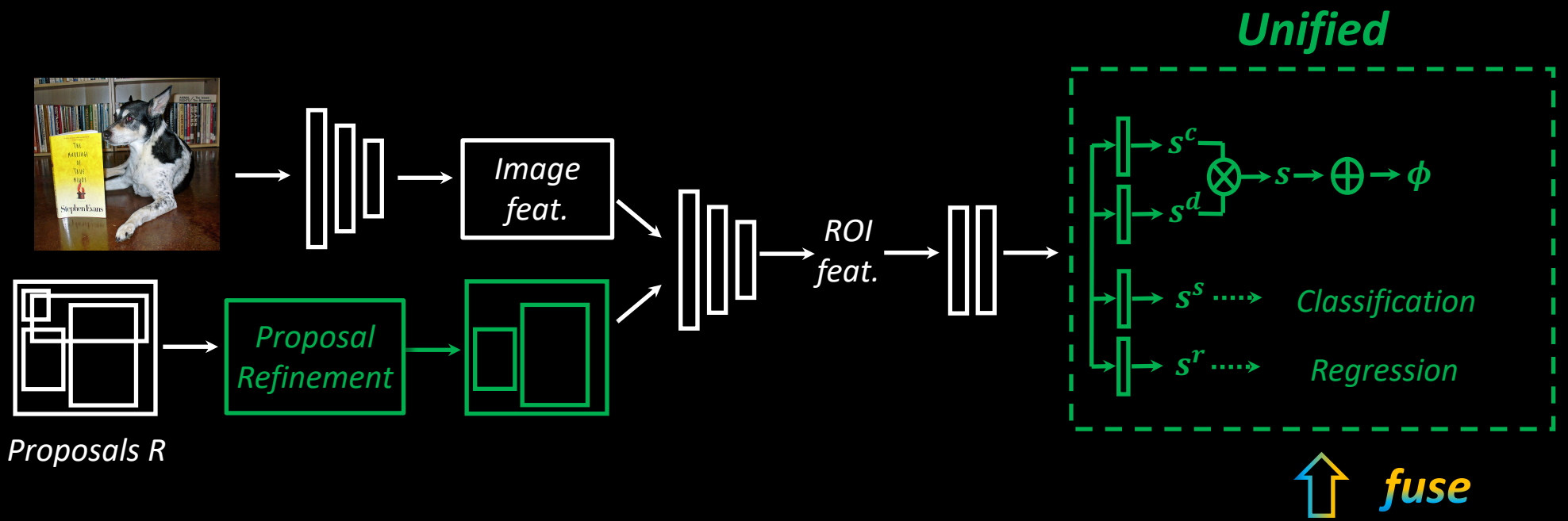
Prior work

- **Stage-wise training; pipelined**
- **Require some strong labels**

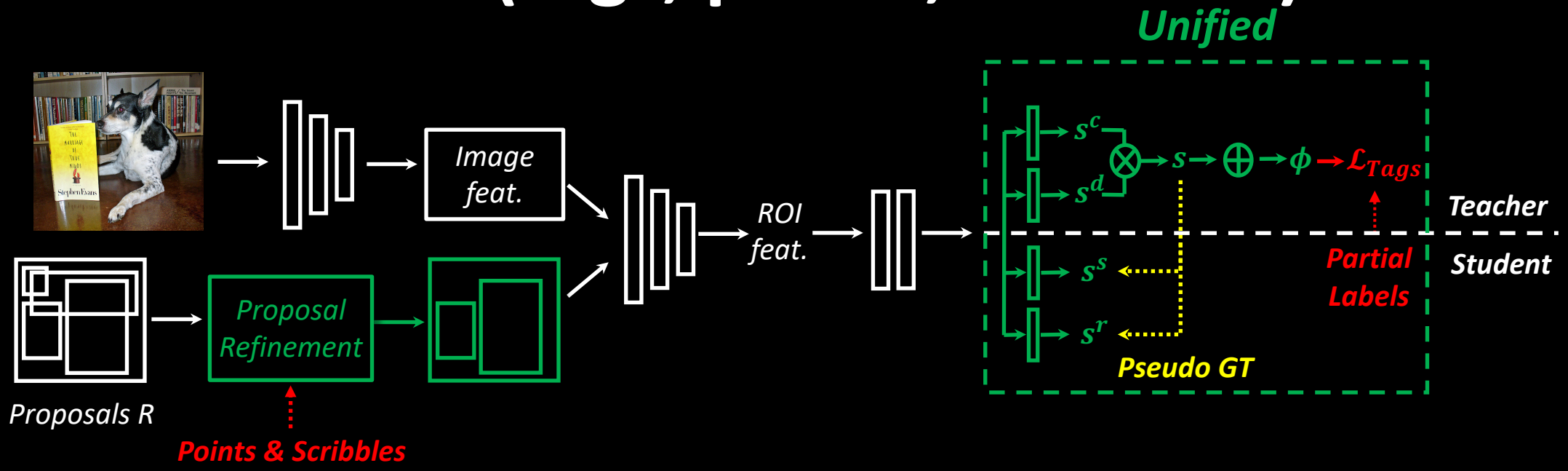
Ours

- **Unified**
- **Strong labels are not necessary**
- **More labels supported**

UFO²: a Unified Framework



Partial Labels (tags, points, scribbles)



No GT boxes available

Teacher heads:

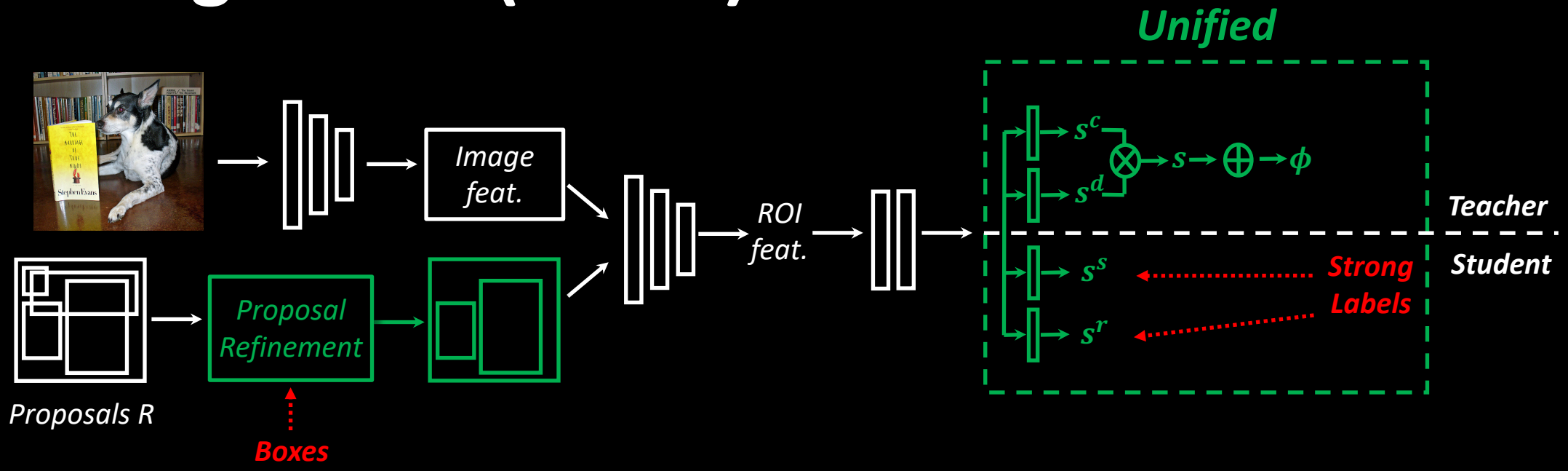
- image-level multi-label classification ($\mathcal{L}_{Tags}$)

Generate pseudo-GT online

Student heads:

- RoI classification
- RoI regression

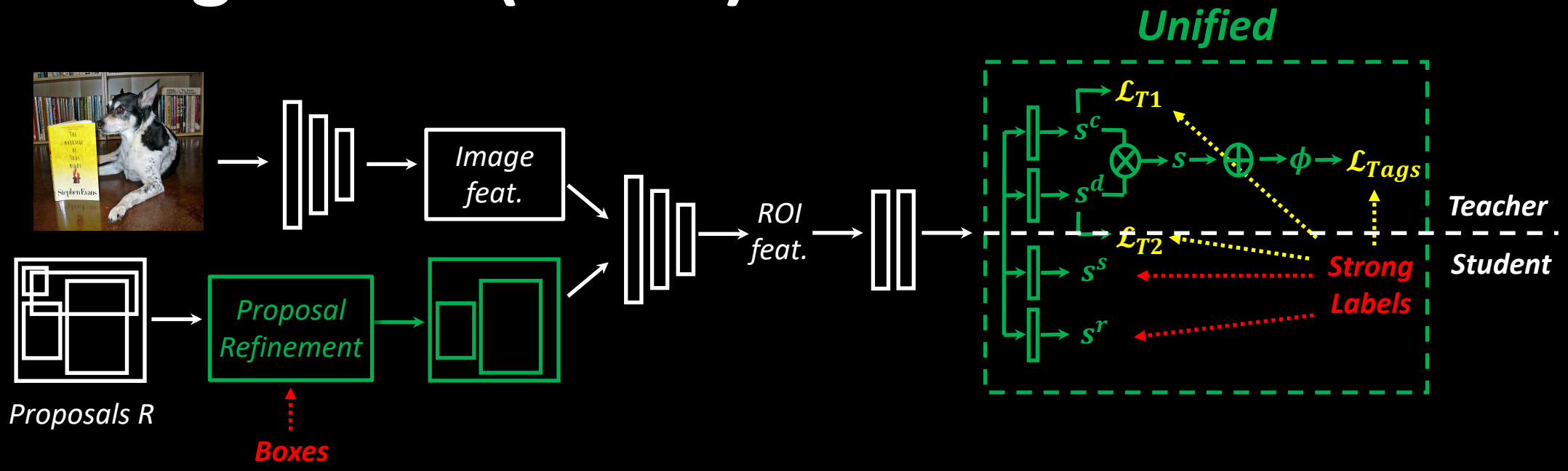
Strong Labels (boxes)



Naïve solution: directly supervise student heads using GT boxes

Issue: Weak Teacher & Strong Students

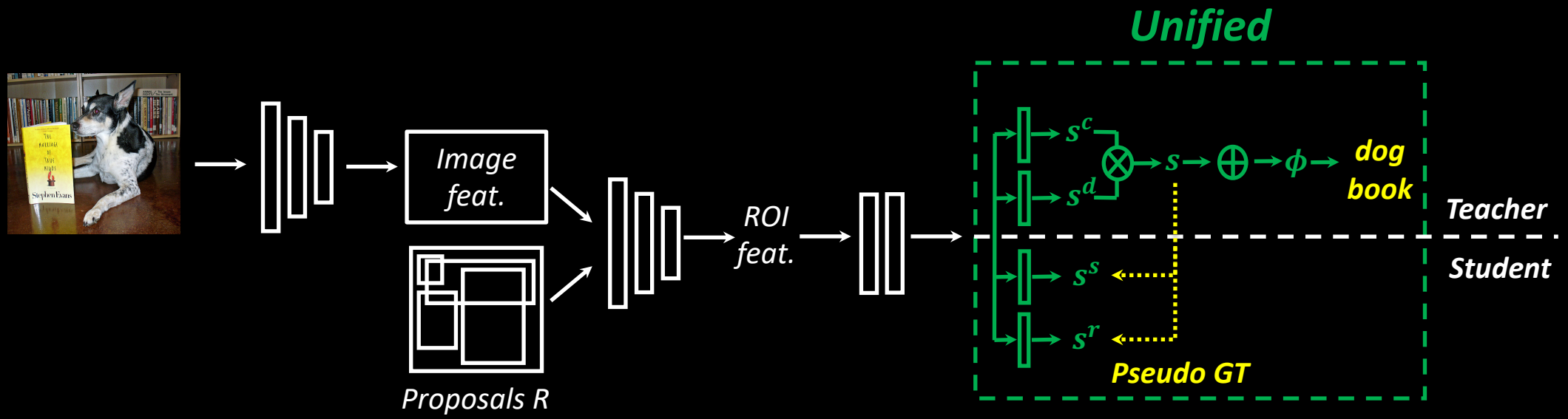
Strong Labels (boxes)



Make Teacher Great Again!

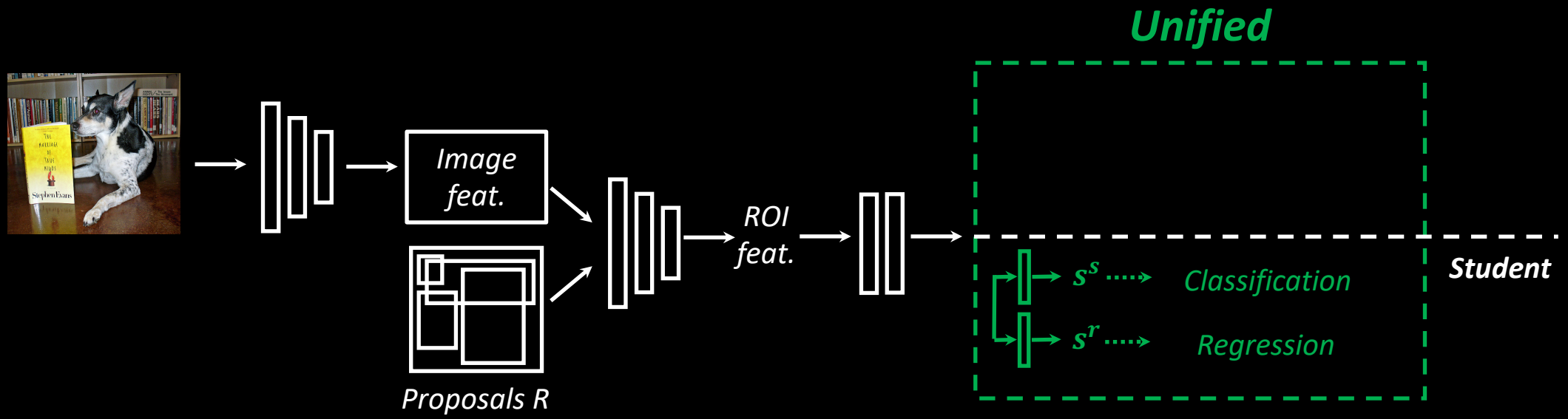
- **Image-level multi-label classification** $(\mathcal{L}_{Tags})$
- **RoI classification** $(\mathcal{L}_{T1})$
- **RoI objectness regularization** $(\mathcal{L}_{T2})$

Unlabeled Data



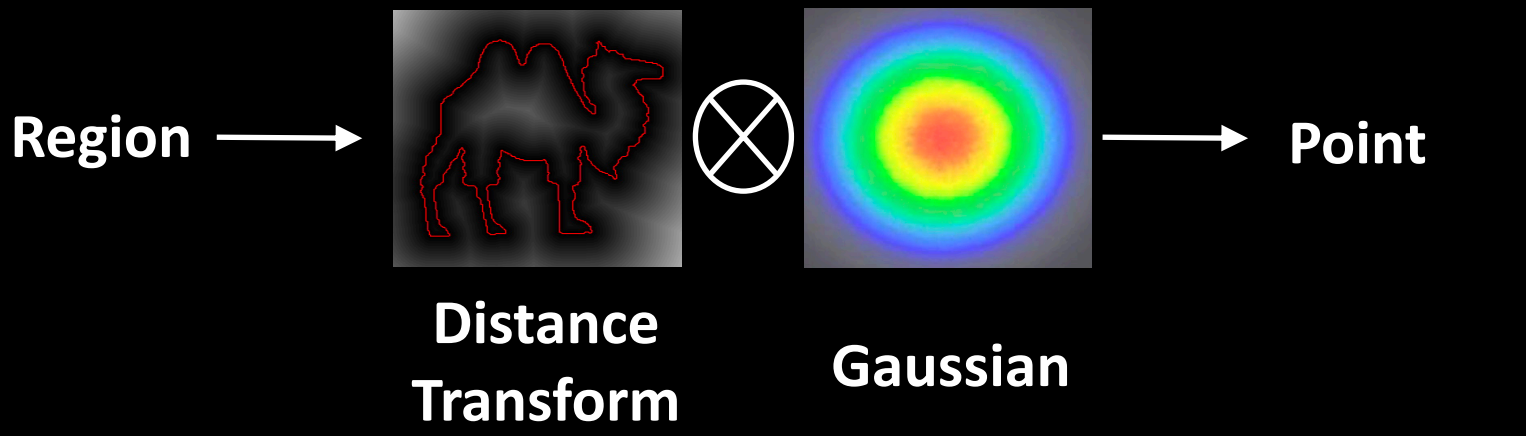
- Take confident classes from teacher's prediction
- Then follow “Tags” setting

Inference

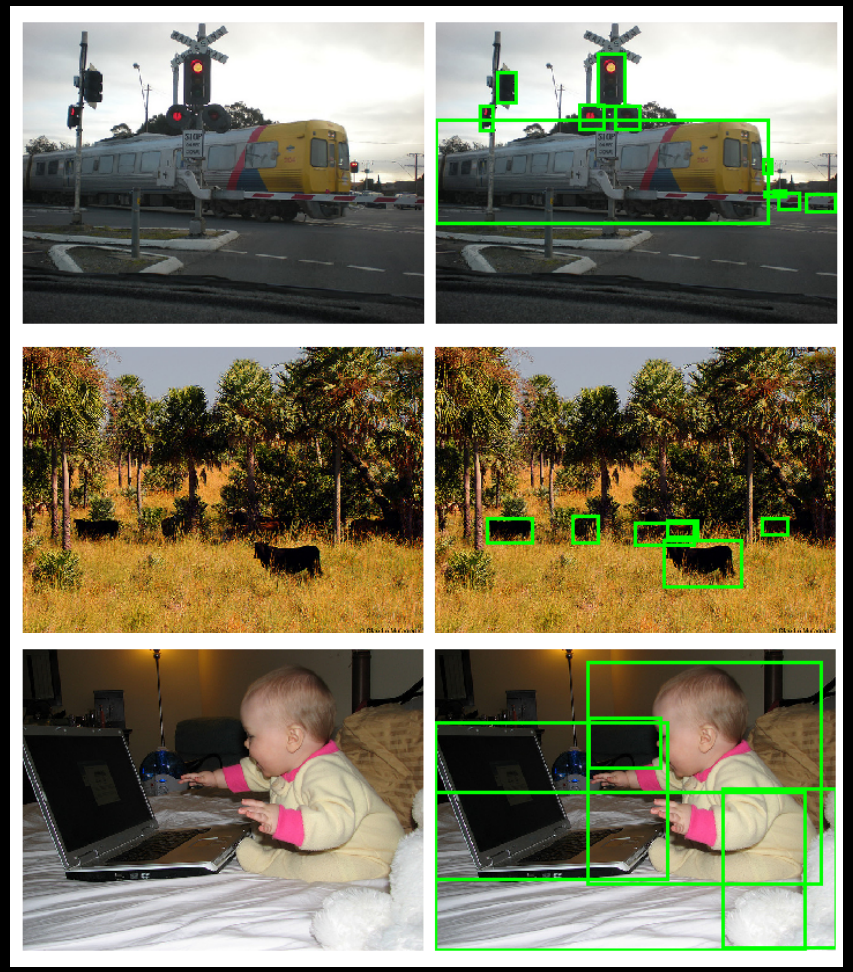


- Use only the student heads
- As efficient as standard supervised detectors

Dataset: Partial Labels Simulation

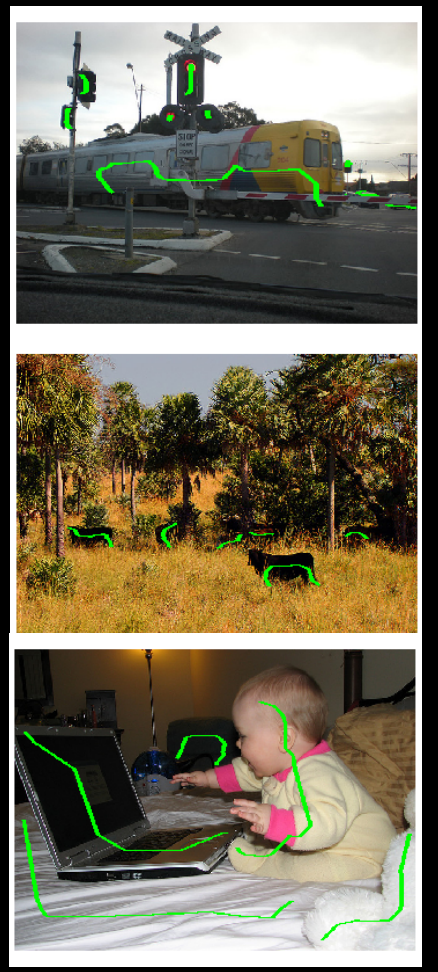


Dataset: Partial Labels Simulation



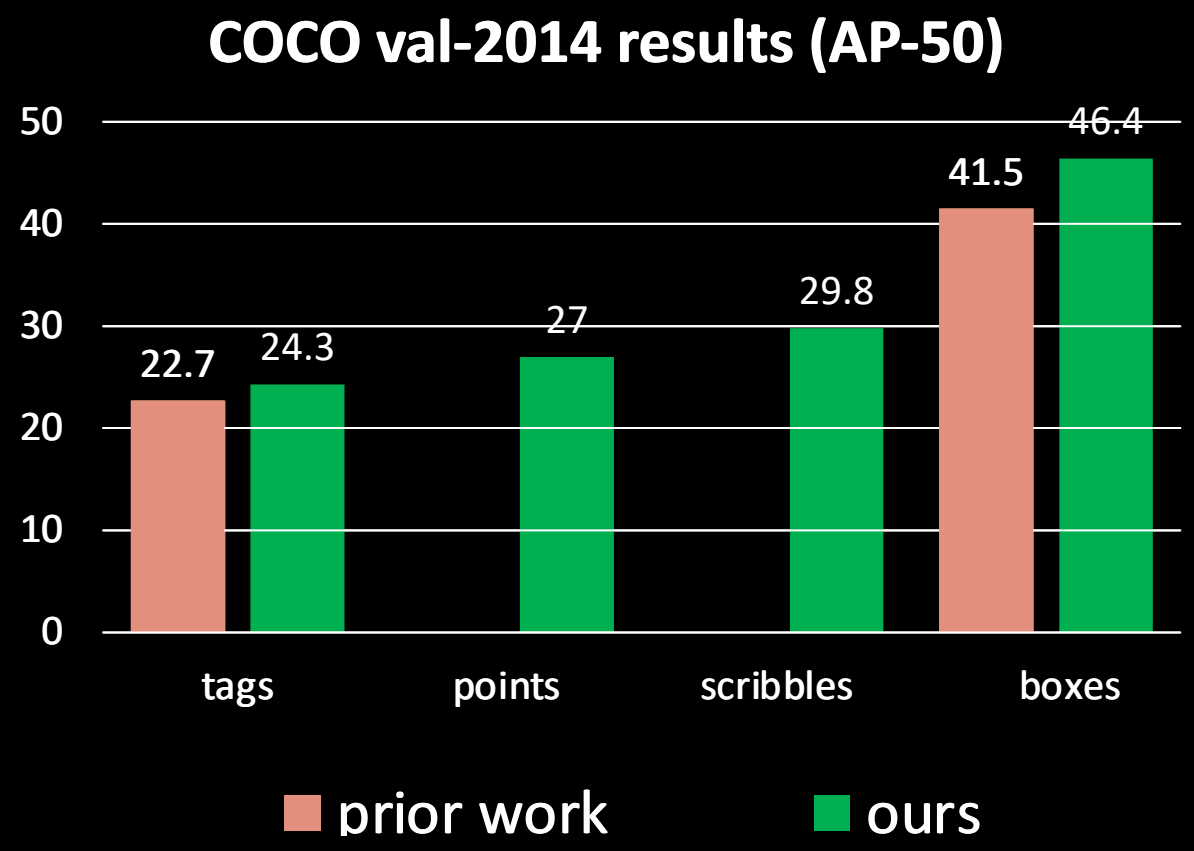
COCO images & boxes

points



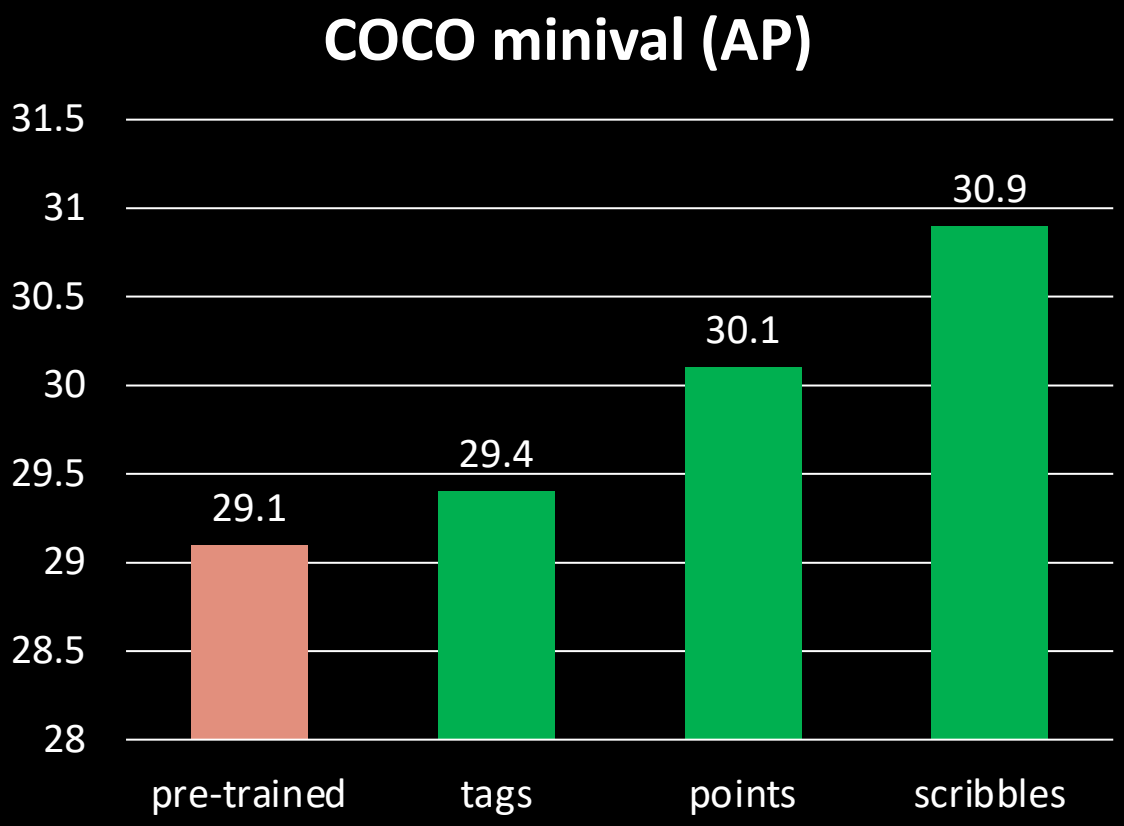
scribbles

Experiments: Train from Scratch

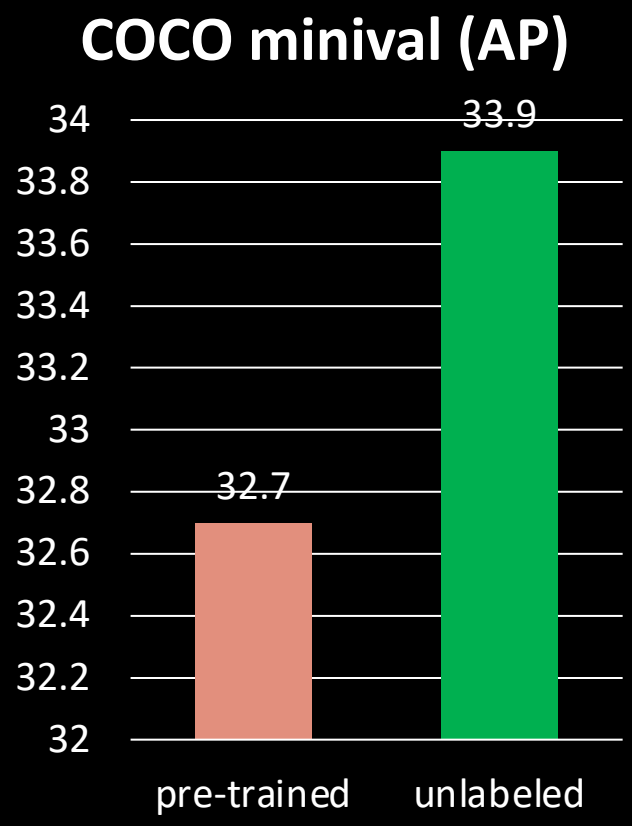


Training from scratch (single label)

Experiments: Improve Pre-trained Models



COCO-35

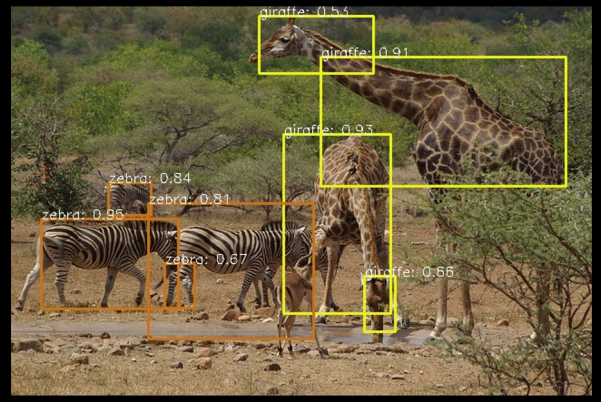


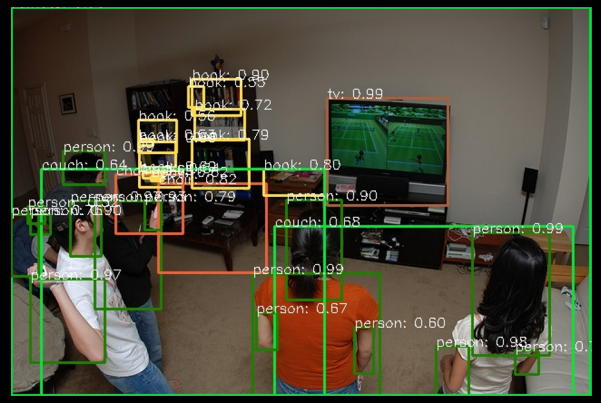
COCO-115

Improving pre-trained models (Mixed labels)

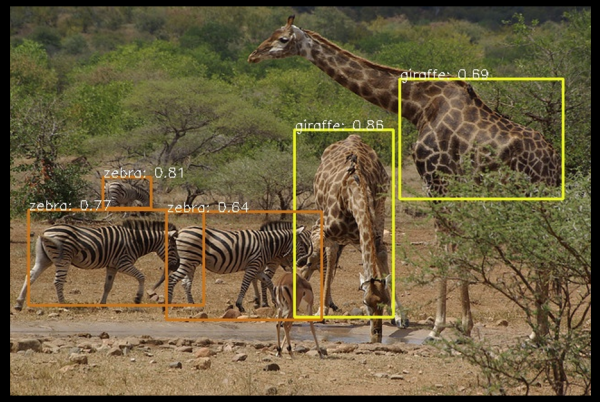
Experiments

Tags

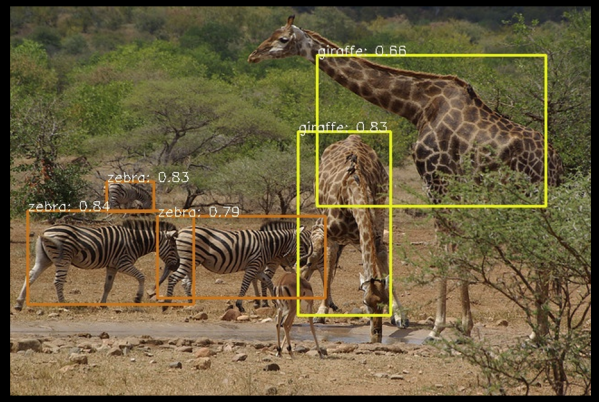




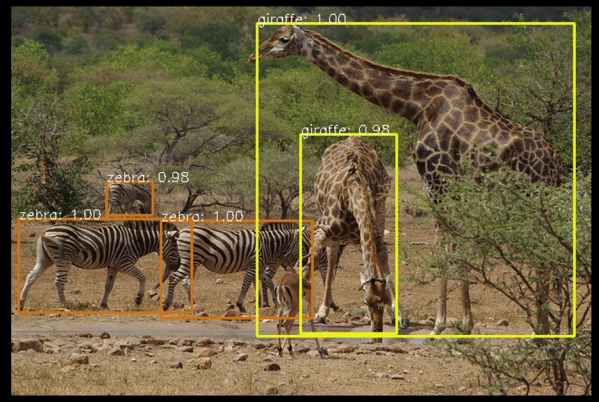
Points

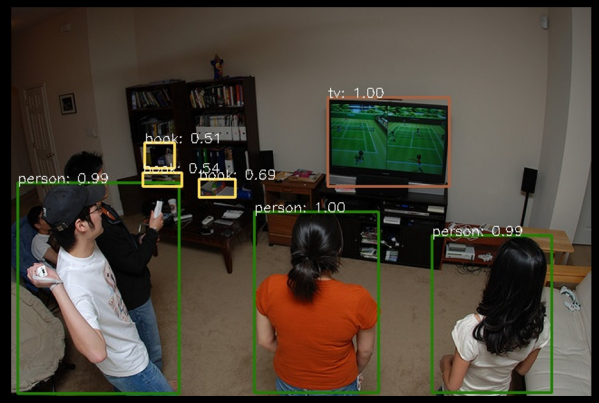


Scribbles



Boxes





Budget-aware Omni-supervised Detection

Approx. per-img budget on COCO:

- tags: 80s, points: 88.7s, scribbles: 160.4s, boxes: 346s

Given a fixed annotation budget (time), the most common strategy:

- **STRONG**: annotate using only boxes

UFO^2 allows a promising new policy:

- **N%B:** use N% budget for boxes and (1-N%) for points

Policy	Labels	# Labeled Images	AP
STRONG	B + U	2312 + 7688	13.97 ± 0.98
80%B	P + B + U	1804 + 1850 + 6346	14.11 ± 1.01

Thanks for watching!