Transferring Labels to Solve Annotation Mismatches Across Object Detection Datasets

Yuan-Hong Liao (Andrew), David Acuna, Rafid Mahmood, James Lucas, Viraj Prabhu, Sanja Fidler

Accepted in ICLR 2024





VECTOR INSTITUTE

UNIVERSITY OF TORONTO

Transferring Labels to Solve Annotation Mismatches Across Object Detection Datasets

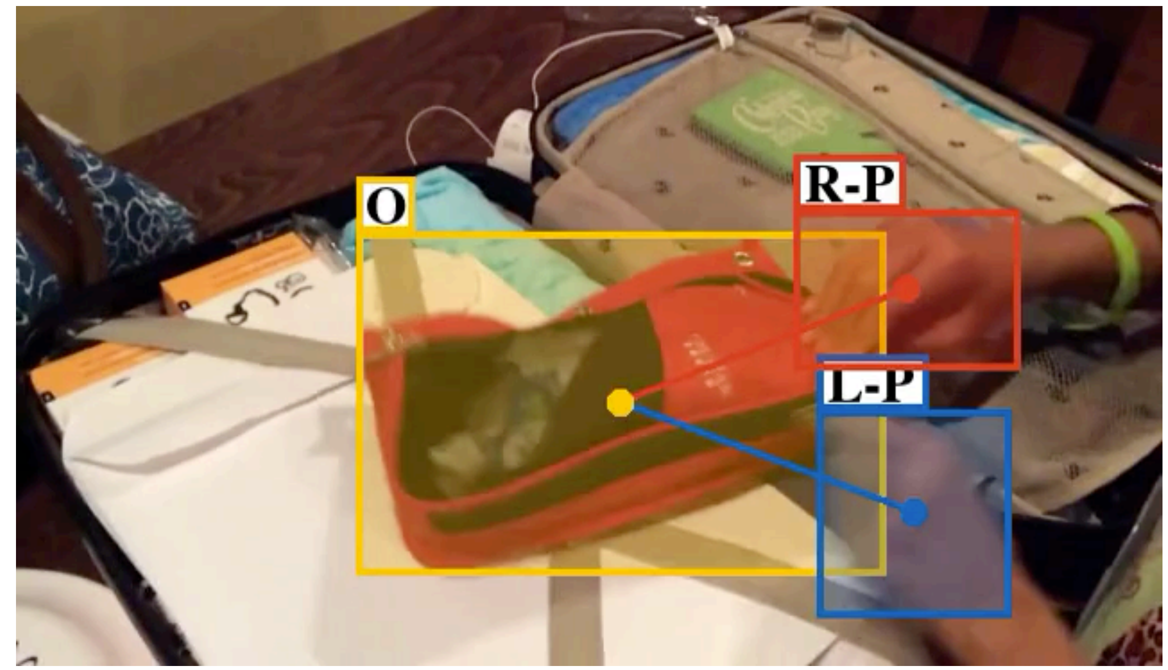
Key contributions:

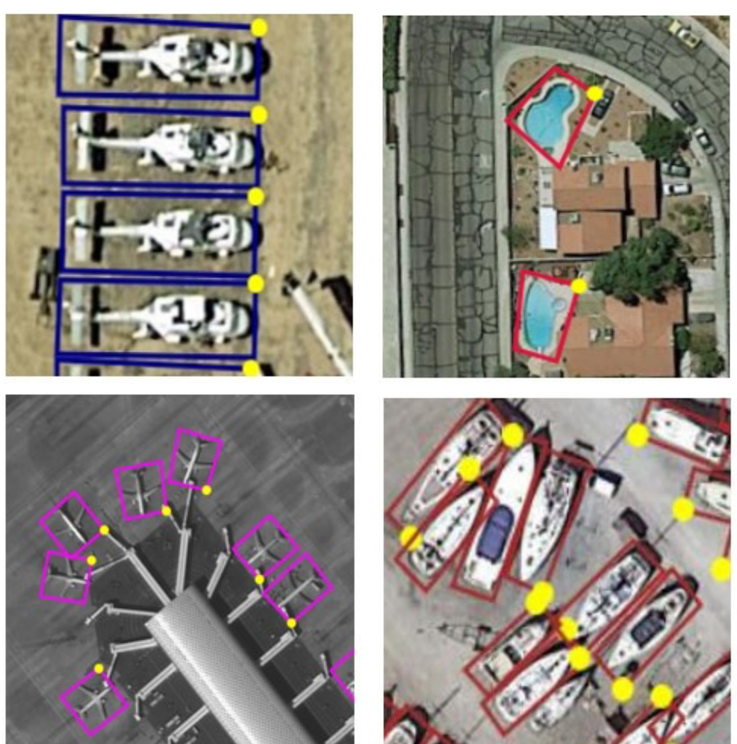
1. This paper characterizes a *prevalent but under-explored* label issue: **Annotation mismatch.**
2. To mitigate annotation mismatches, we propose a *data-centric* approach that **transfers the labels** in the datasets “before training any detector”.
3. The proposed approach is **agnostic** to detector learning algorithms & detector architectures.

Object detection datasets are everywhere

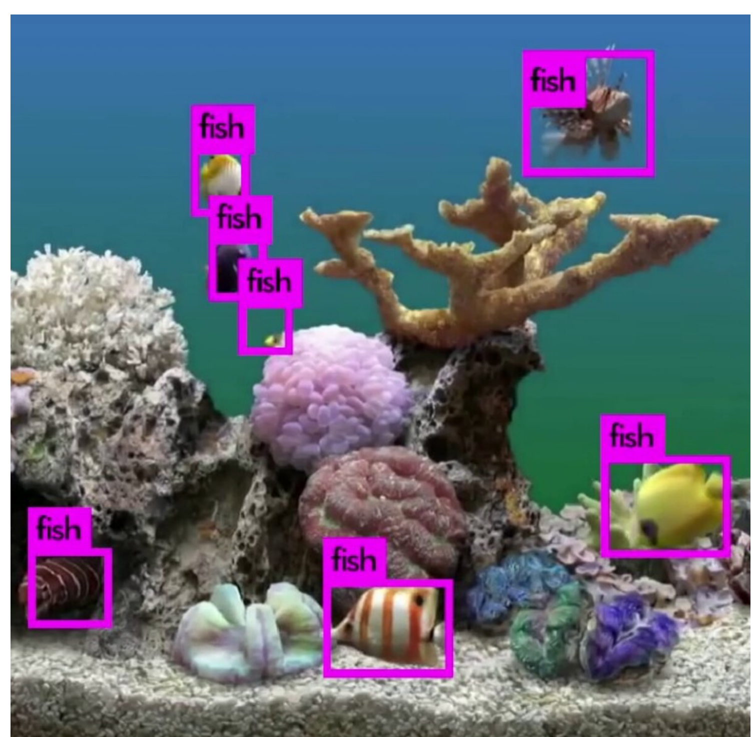
Autonomous Driving [1]

Ego Camera — Hand detection [2]

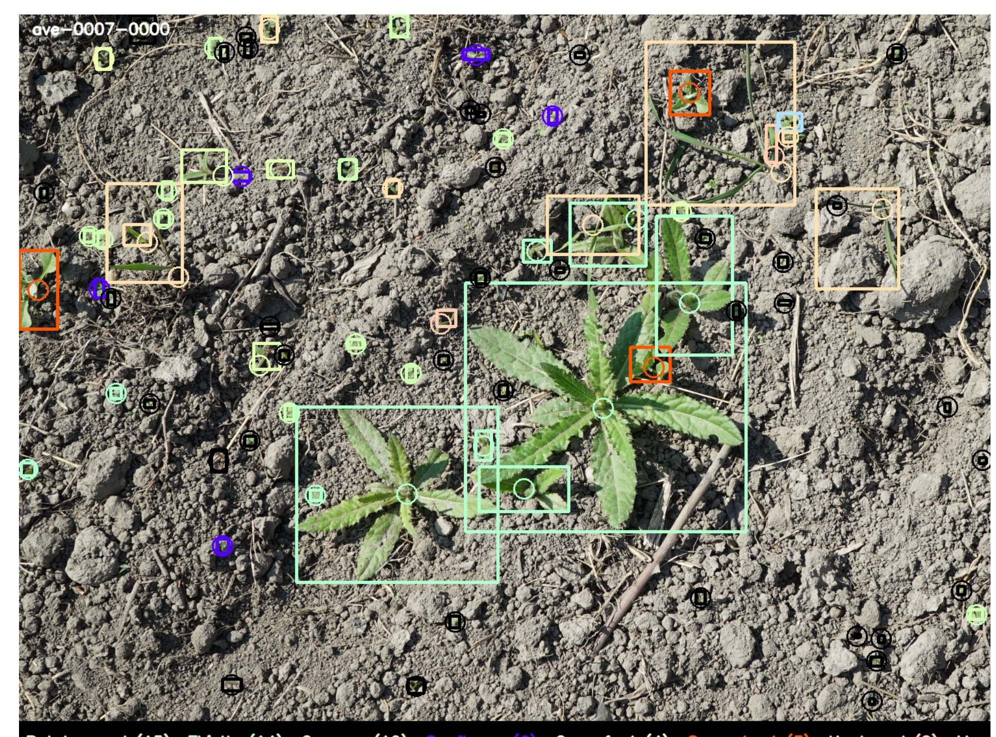




Aerial Images [3]



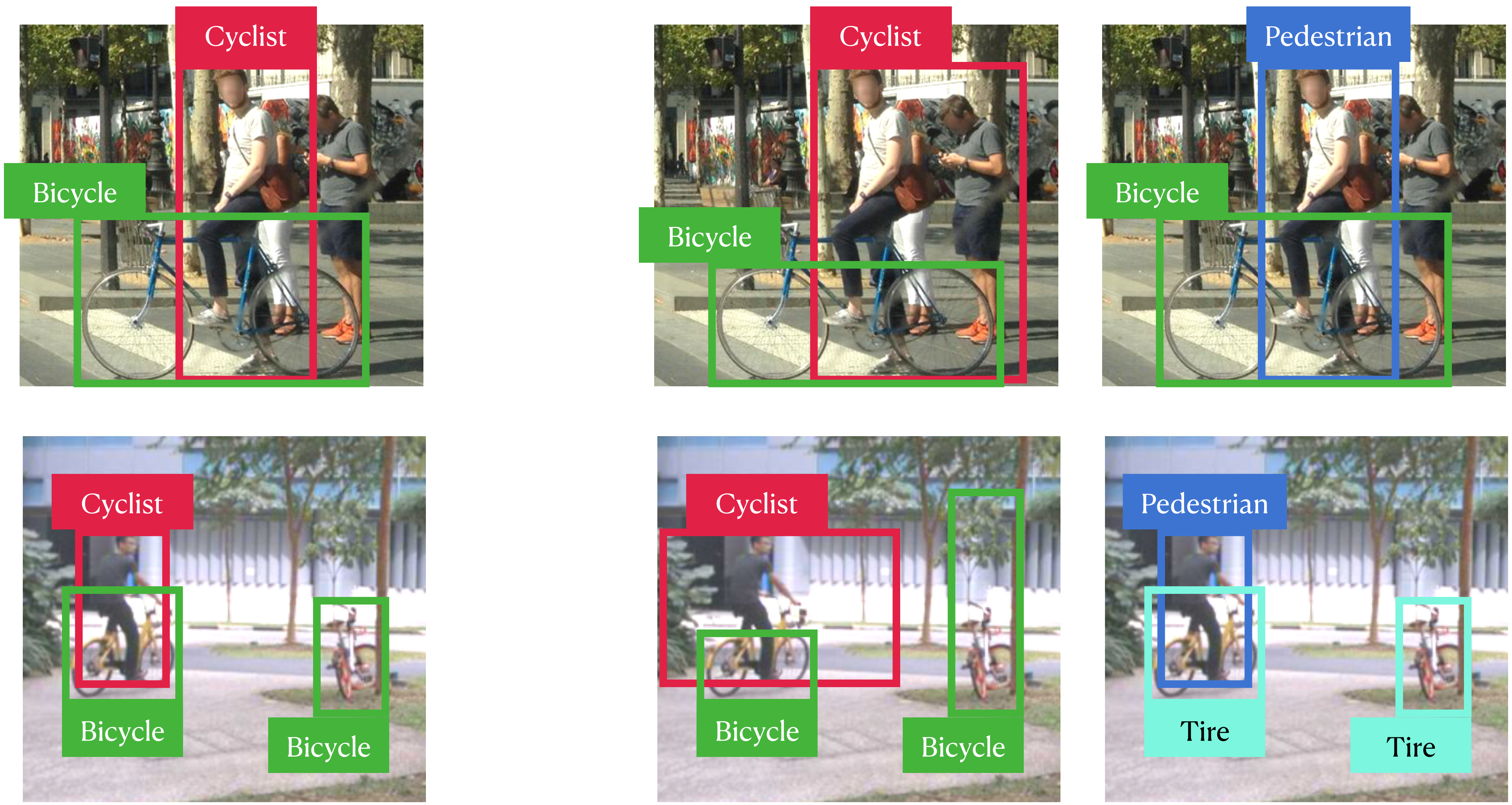
Underwater [4]



Agriculture [5]

Label quality matters

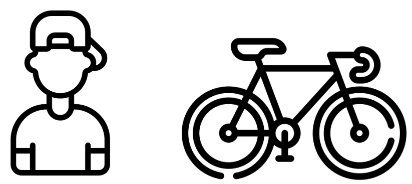
Lower label quality leads to worse detector performances [6,7]

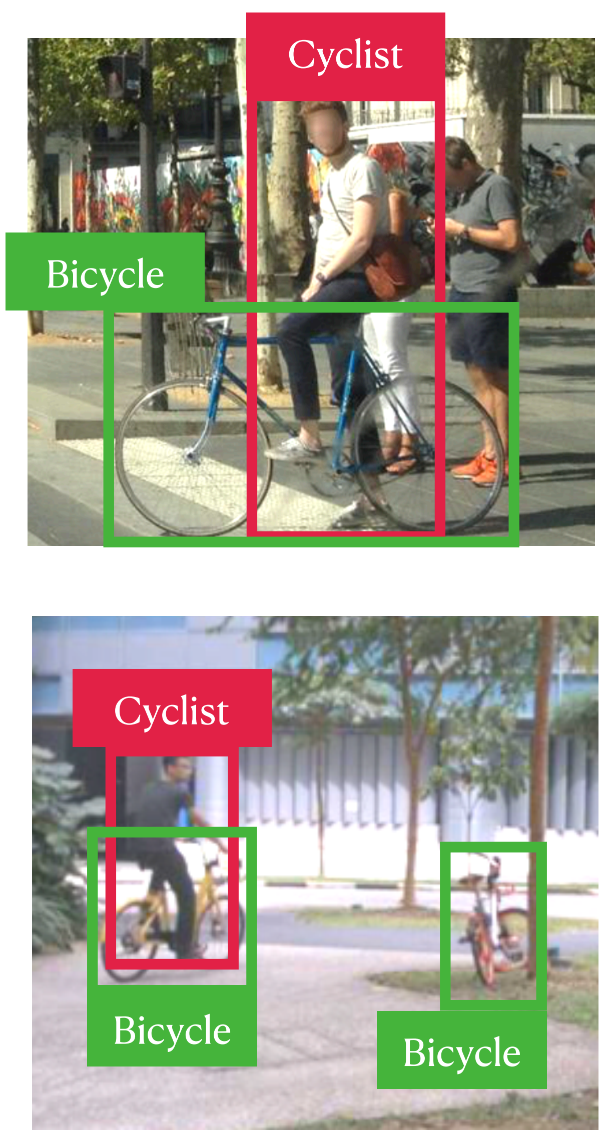


Incorrect Bounding Boxes

Incorrect Class Labels

Are they truly “correct” correct?

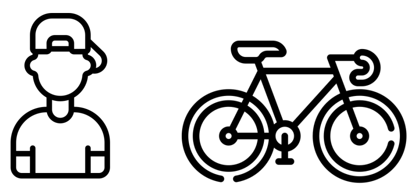


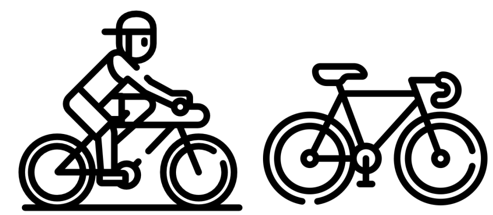


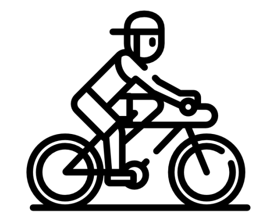
Incorrect Bounding Boxes

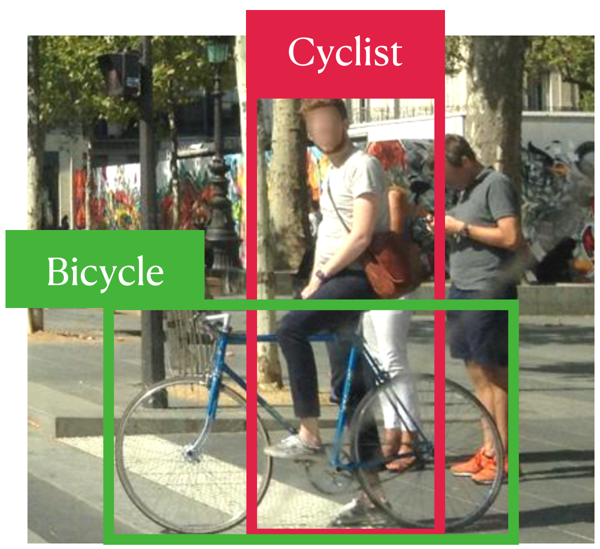
Incorrect Class Labels

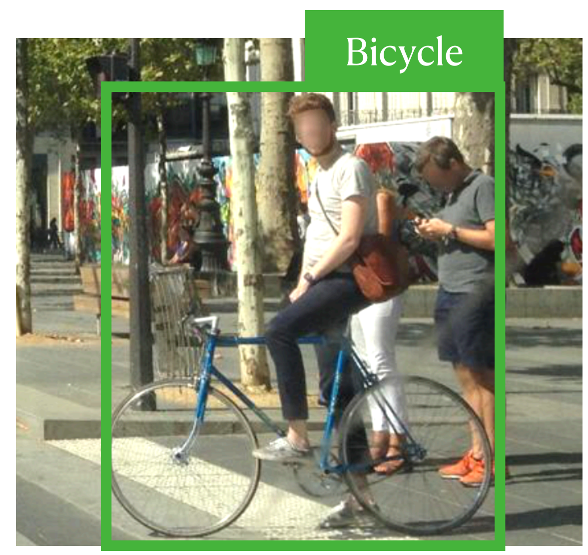
Which labels are more correct?

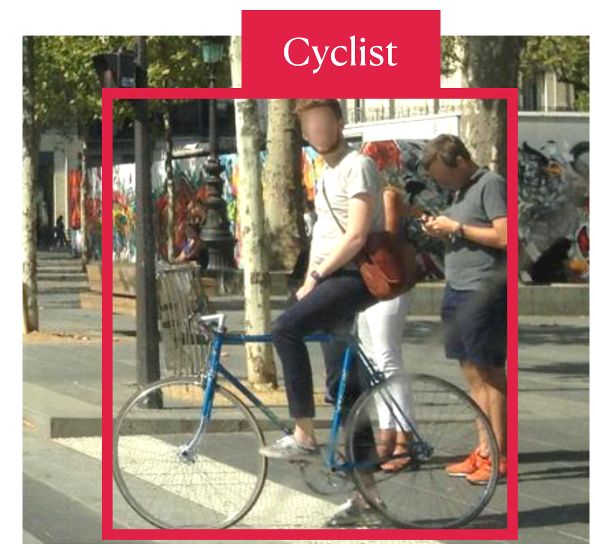


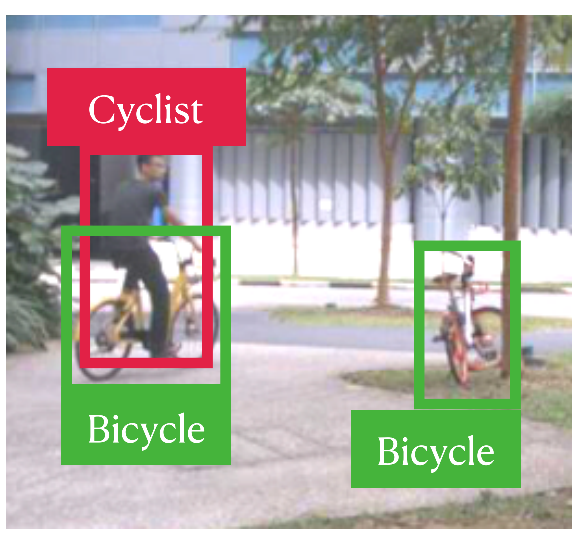


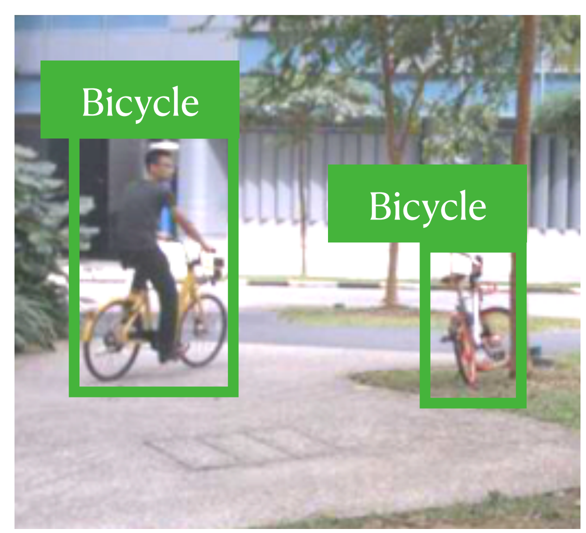






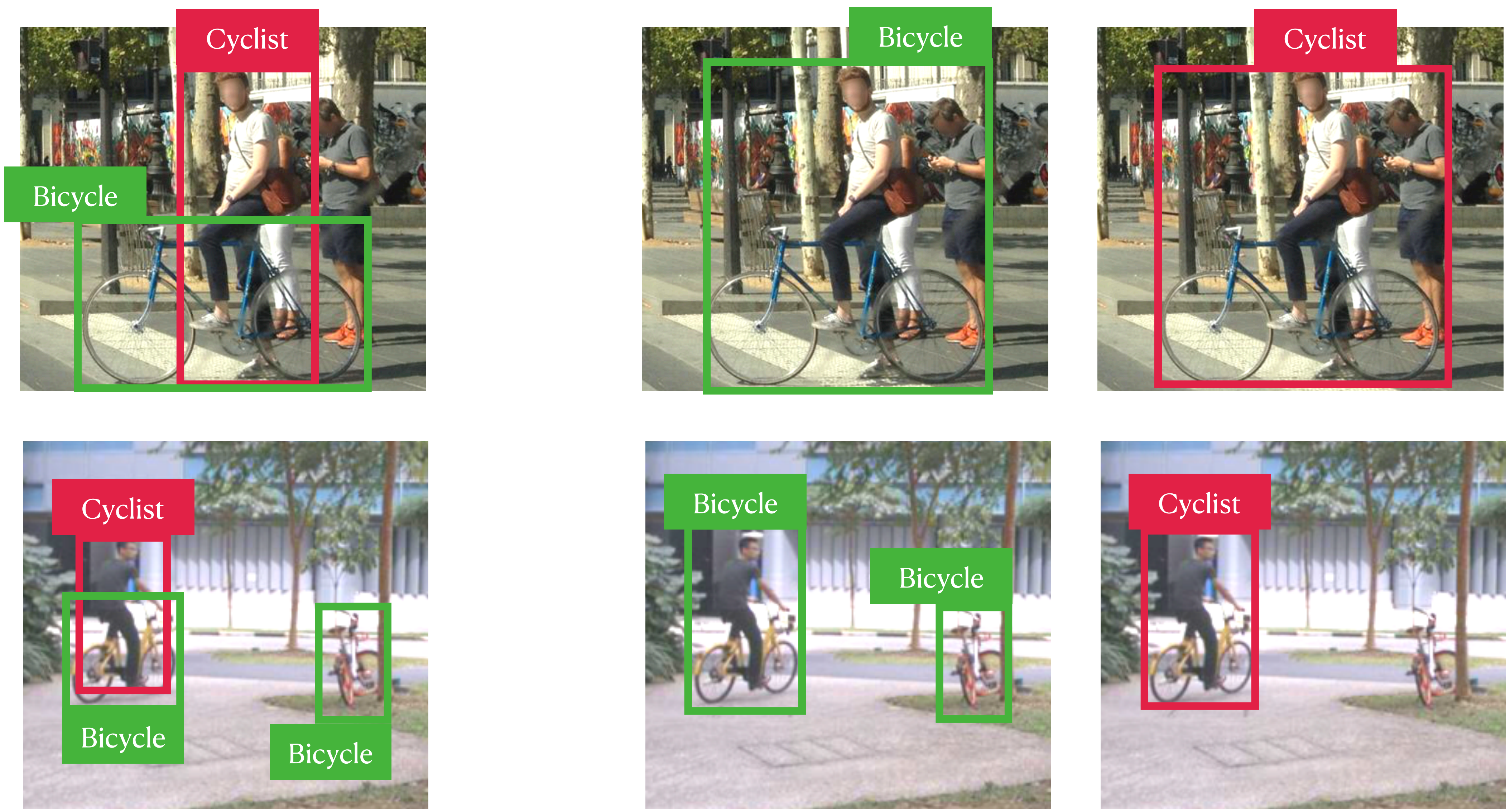






They are all “correct”

but with annotation mismatches



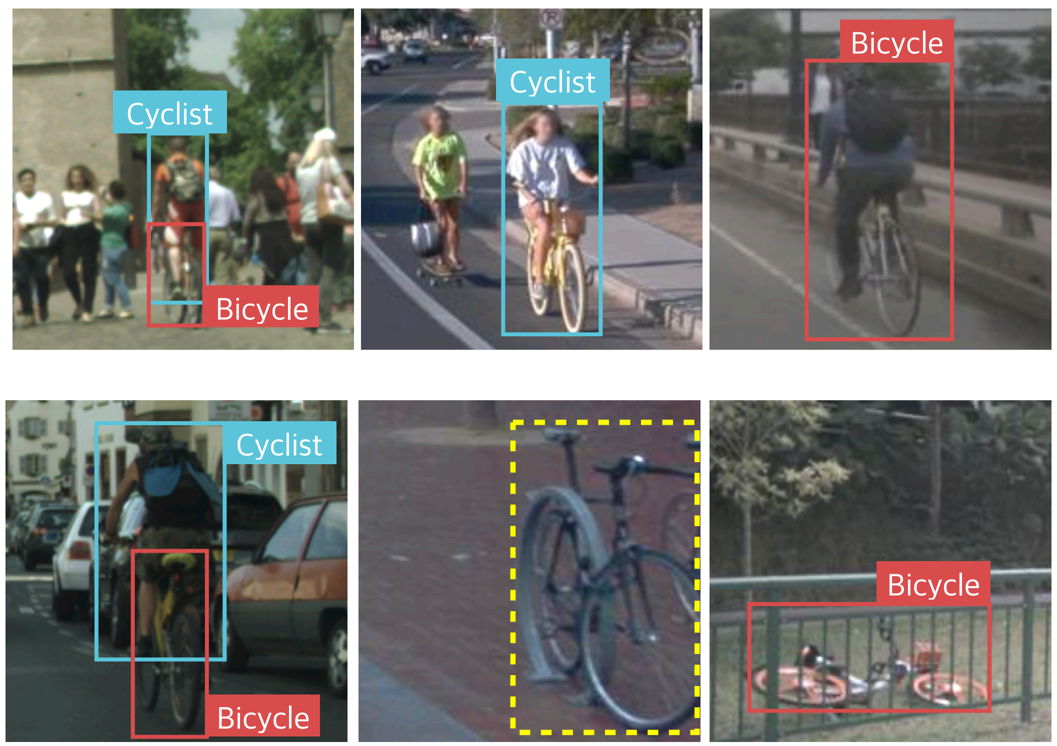
Cityscapes Labels

nuImages Labels

Waymo Labels

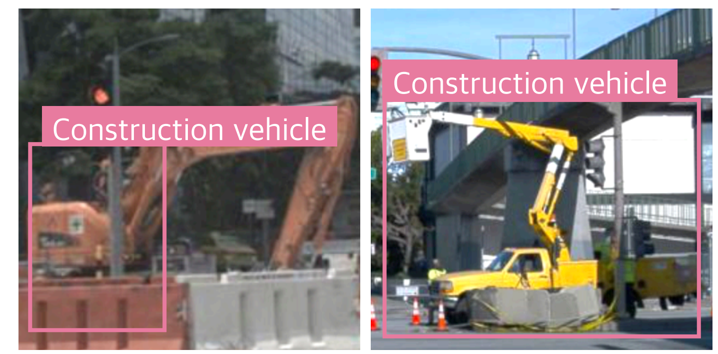
Prevalent annotation mismatches

Class semantics

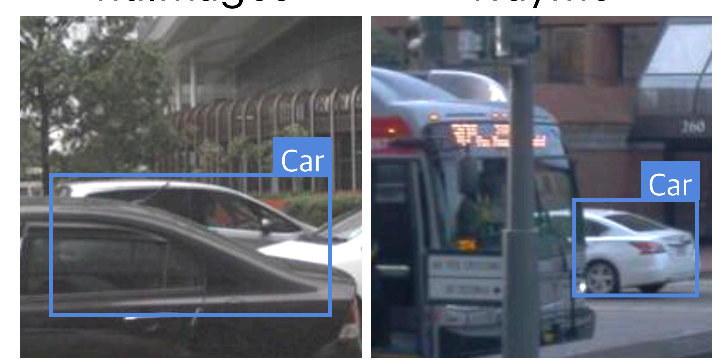


Cityscapes Waymo nuImages

Annotation instructions



nuImages Waymo



nuImages
(amodal)

Waymo
(pixel-based)

Human-machine misalignment

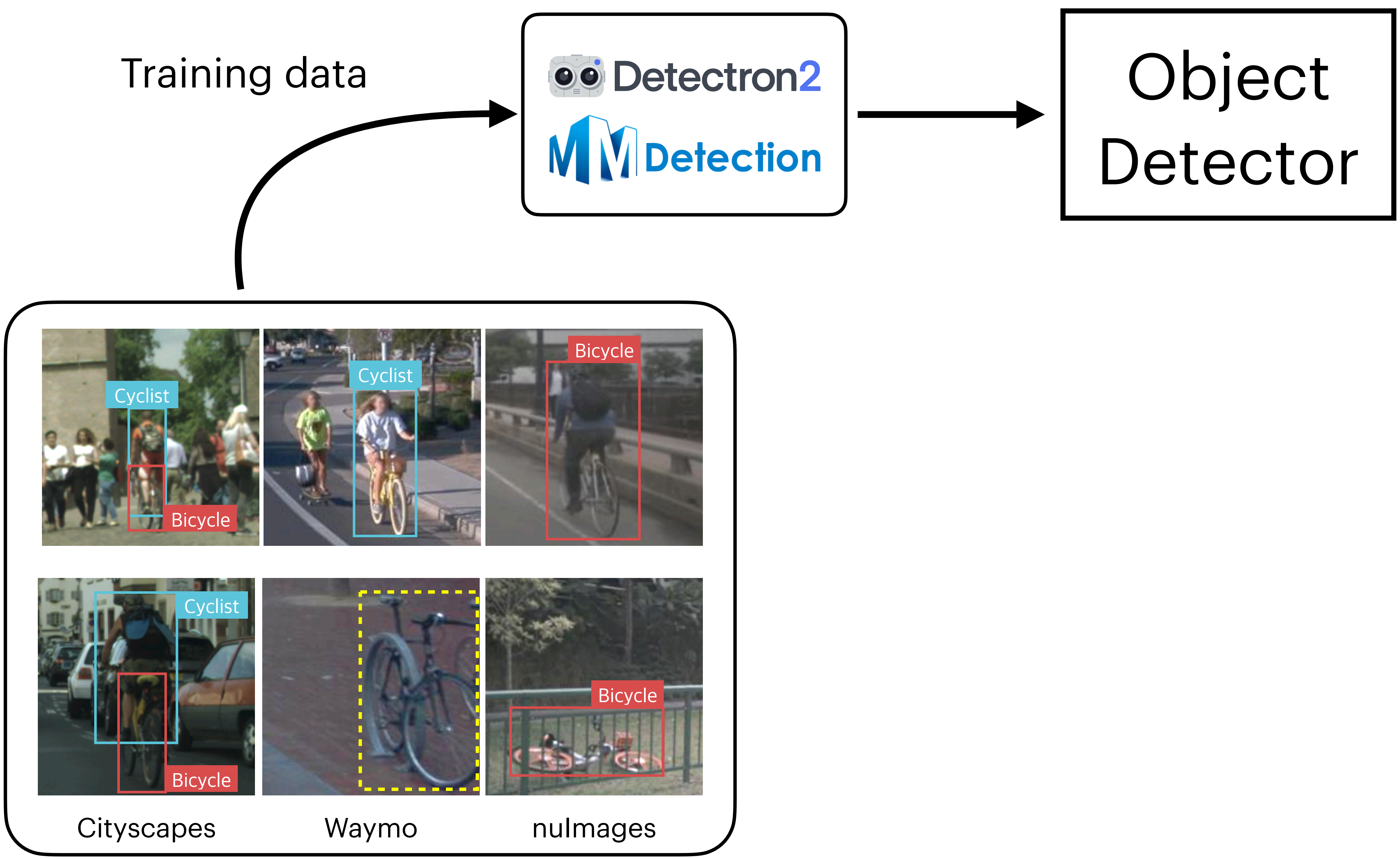
Synscapes

Cross-modality labels

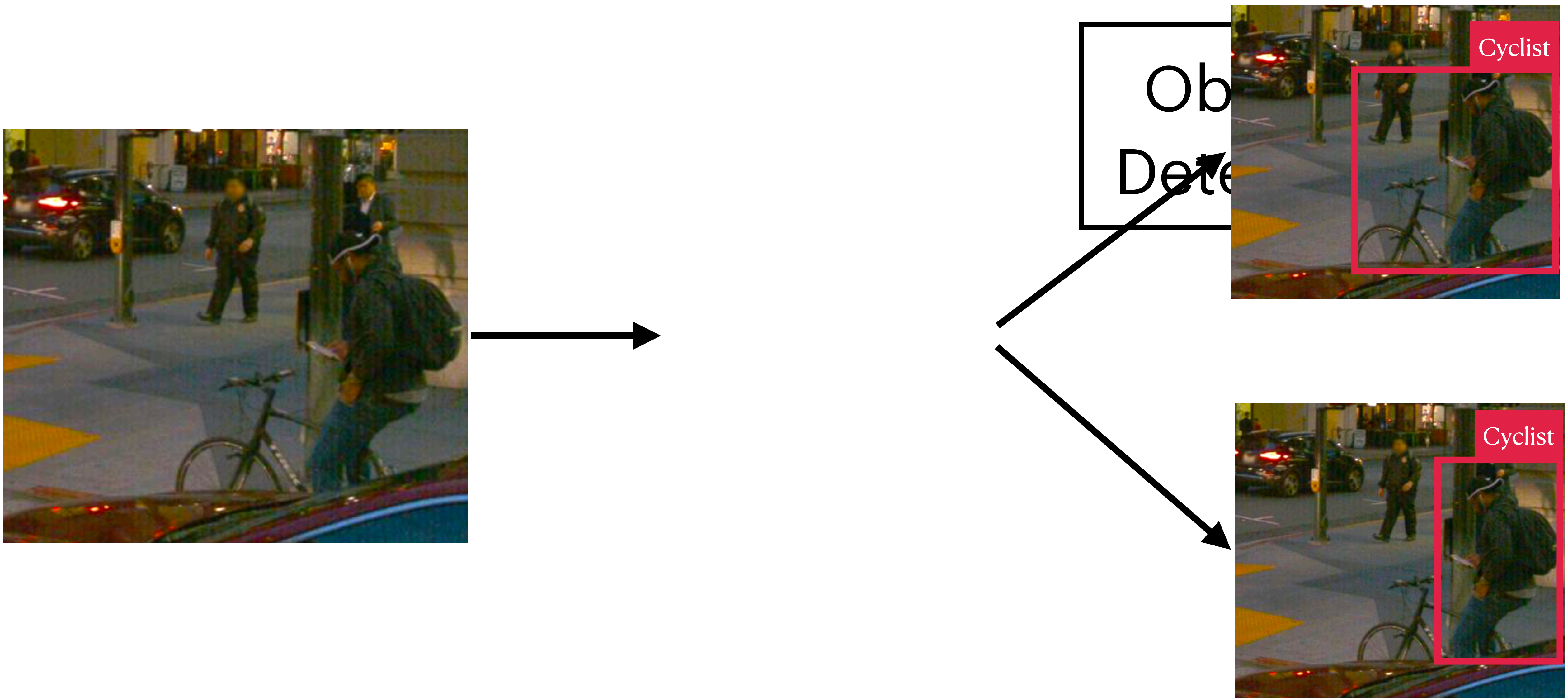
nuScenes (from 3-D bounding boxes)

MVD (from segmentation maps)

If Annotation Mismatches are the Answers, Then What is the Question?

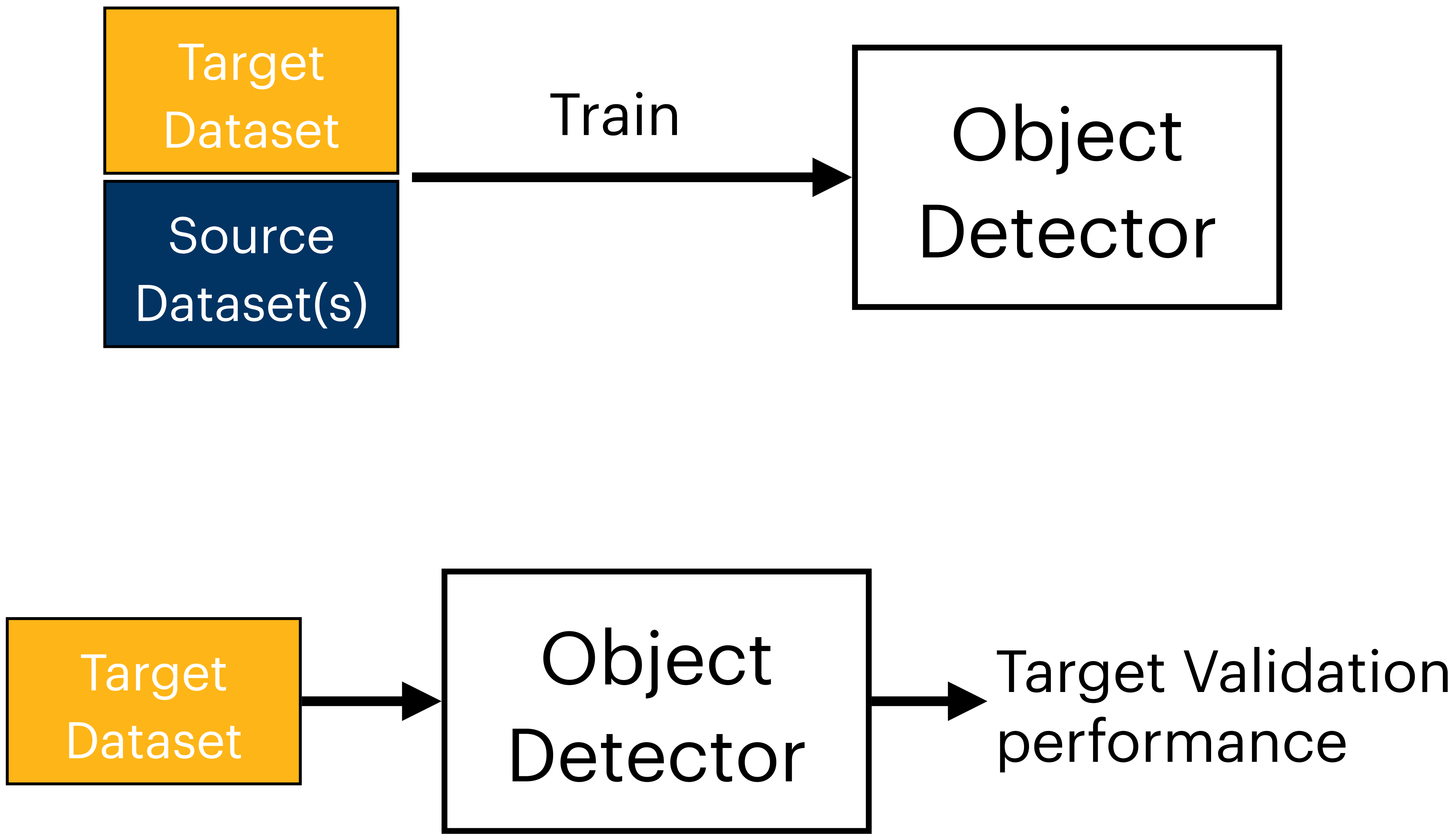


If Annotation Mismatches are the Answers, Then What is the Question?



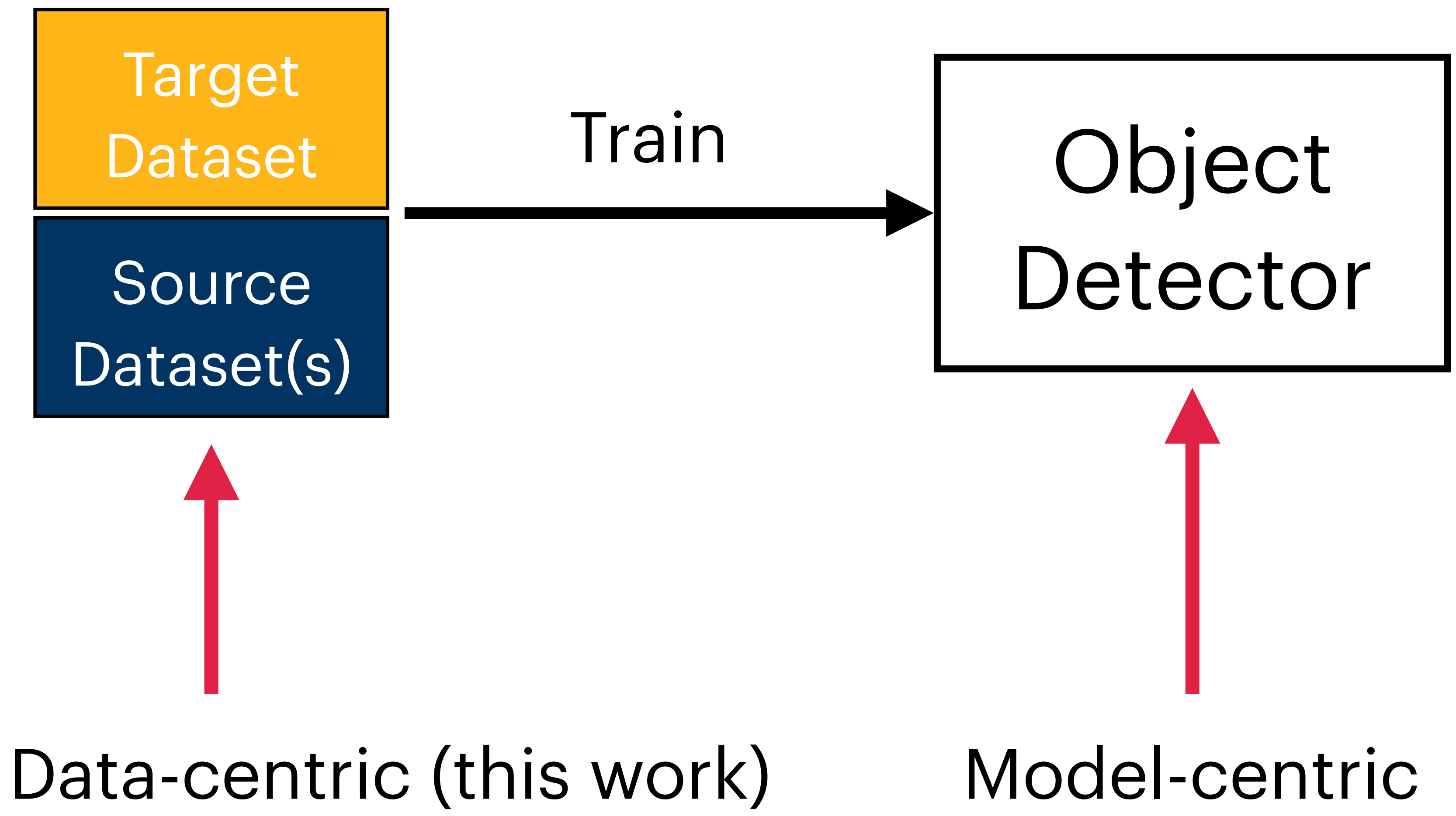
Application: Supervised Domain Adaptation

Fully labeled, Annotation mismatches

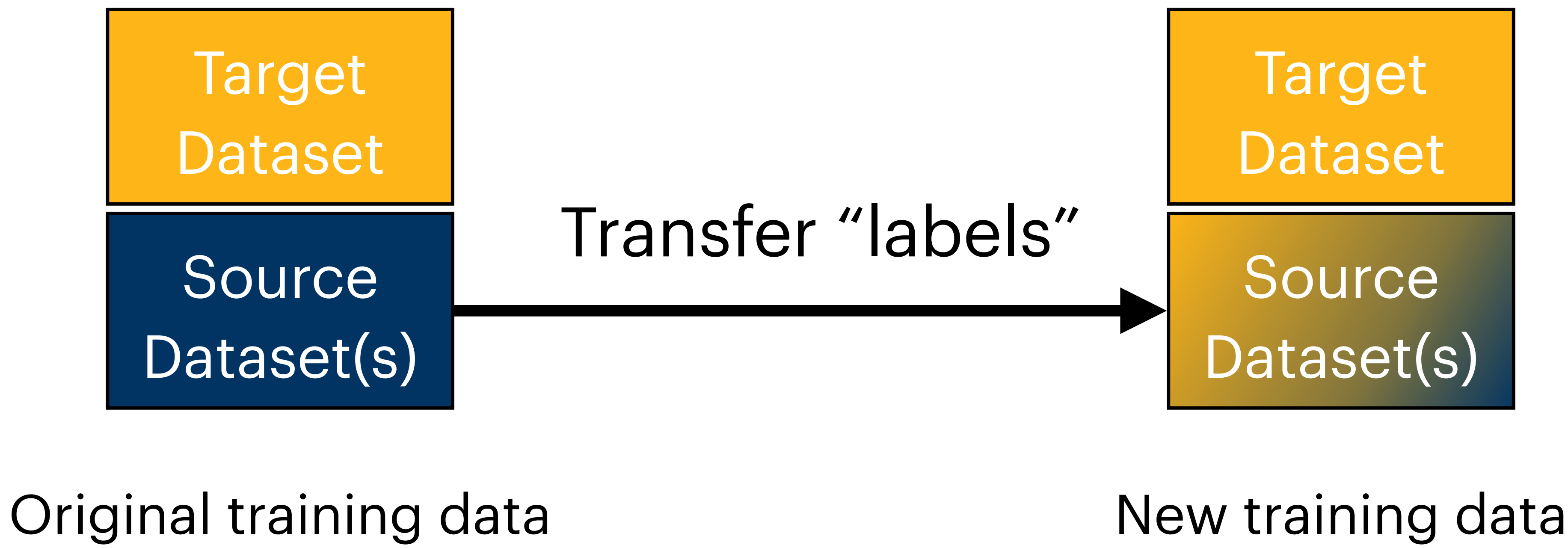


Prabhu et al., Bridging the Sim2Real gap with CARE: Supervised Detection Adaptation with Conditional Alignment and Reweighing

Application: Supervised Domain Adaptation



Problem Formulation — Label Transfer

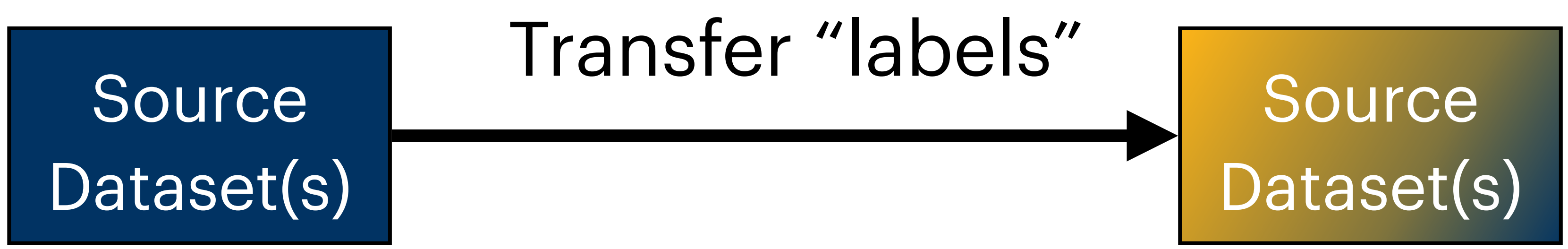


Data-centric key benefits: Agnostic to downstream detectors

*Please refer to our paper for a more formal problem formulation in our paper.

Challenges

1. No paired labels

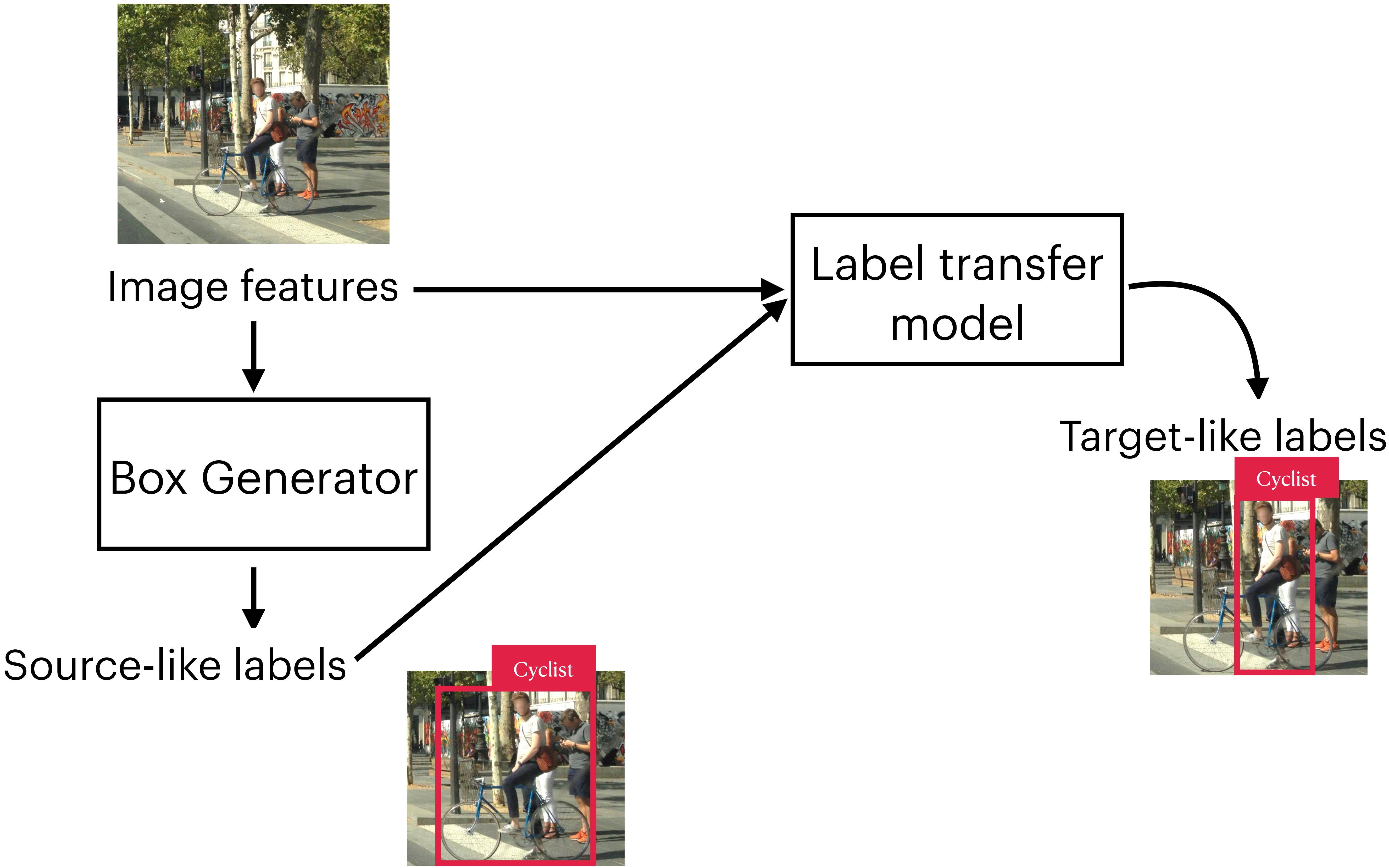


2. Complex task

Task requirements

1. Translate bounding boxes
2. Adjust class labels
3. Remove detection labels
4. Add detection labels

Label Guided Pseudo-Labeling (LGPL)



Label Guided Pseudo-Labeling (LGPL)

Supervision: Source Dataset(s)

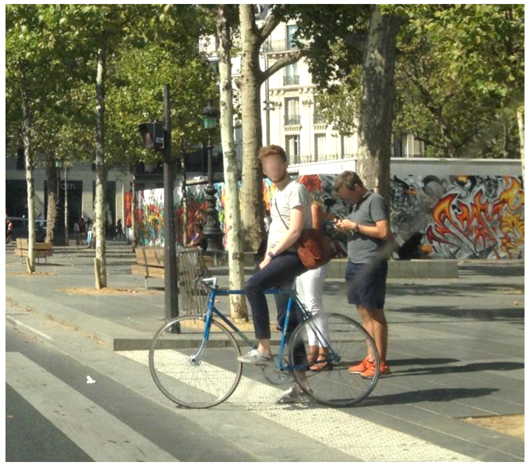


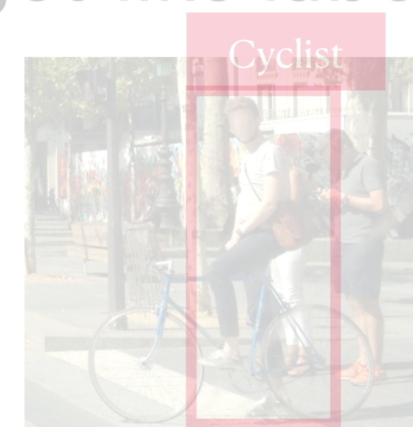
Image features

Box Generator

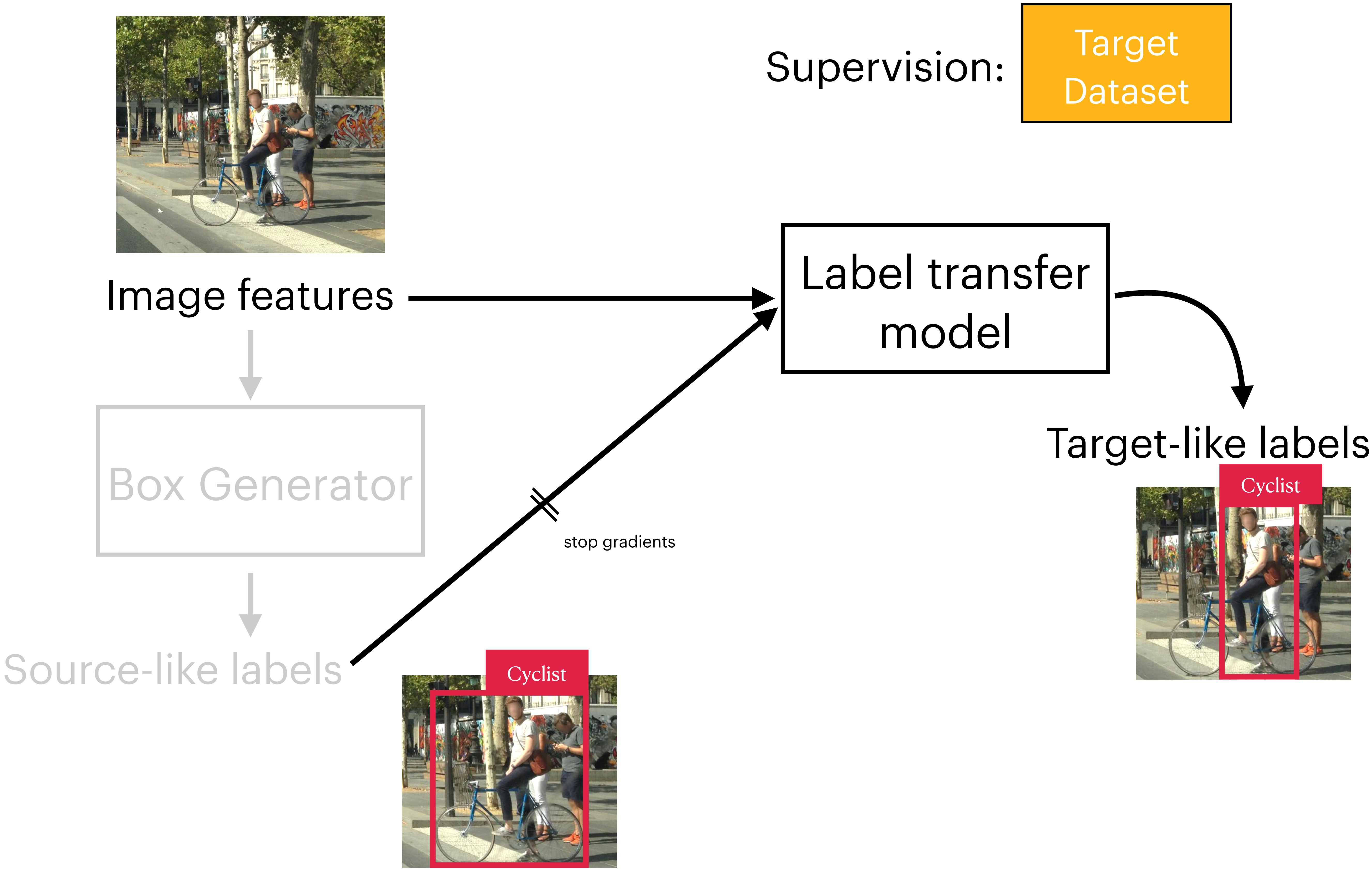
Source-like labels

Label transfer model

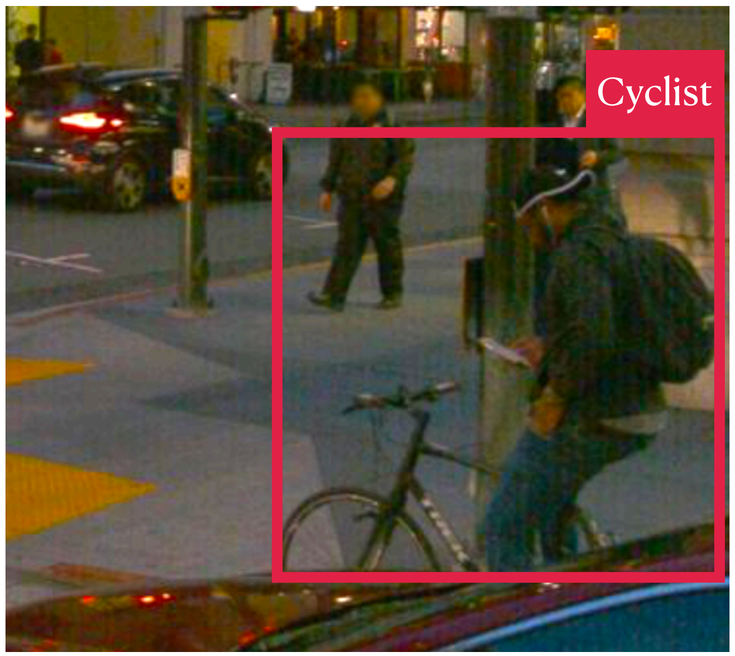
Target-like labels



Label Guided Pseudo-Labeling (LGPL)



Label Guided Pseudo-Labeling (LGPL)



Label
m

Source labels

Label transfer model

Target labels

Experimental Results

Evaluation:

- Evaluate on 4 transferring scenarios involving 5 real-world and 2 synthetic datasets
- Evaluate 3 downstream detectors

Results:

By pre-processing the source labels with LGPL, detectors ***always*** improve, increasing 2.7, 1.68, 1.2 mAP in YOLOv3, Deformable DETR and Faster-RCNN, respectively

References

[1] nuScenes: A multimodal dataset for autonomous driving

[2] Understanding Human Hands in Contact at Internet Scale

[3] Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges

[4] https://www.kaggle.com/datasets/slavkoprytula/aquarium-data-cots

[5] The CropAndWeed Dataset: a Multi-Modal Learning Approach for Efficient Crop and Weed Manipulation

[6] Evaluating the Effect of Common Annotation Faults on Object Detection Techniques

[7] Quantifying the Effects of Ground Truth Annotation Quality on Object Detection and Instance Segmentation Performance

[8] Lessons Learned From Cloudsen12 Dataset: Identifying Incorrect Annotations in Cloud Semantic Segmentation Datasets