

# Optimal Correction Cost for Object Detection Evaluation

## Supplementary Material

### A. Details of the Annotation Experiment

The three annotators in Sec. 4.2 are employed as our in-house annotation team. We explained the purpose of the project to the annotators in advance, and they were able to ask any questions during the work. Each annotator completed annotating 1057 samples in four days. They reviewed paired detection results for a subset of the COCO Detection dataset [14] and assigned a binary preference to each pair. We did not store any personal information for this project. We believe that this annotation task does not violate the ethical principles in the CVPR ethics guidelines. We do not show the annotation interface in this supplementary material because it may reveal the authors’ identity.

### B. Full Results of Consistency Analysis

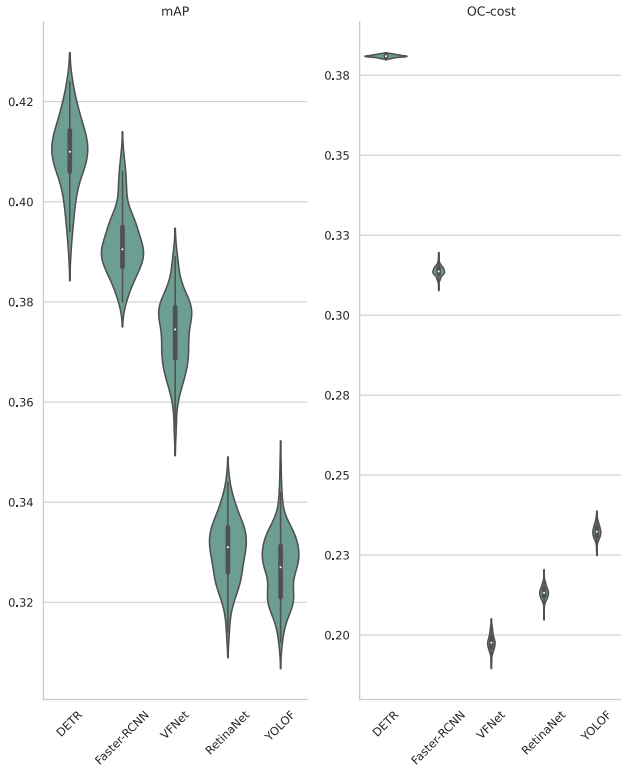


Figure 11. The full results of Fig. 8. Distributions of mAP and OC-cost are obtained by bootstrapping with 100 trials. The distributions’ overlaps across detectors imply that the detectors’ rankings likely to flip by chance.

We omit the DETR’s result in Fig. 8 for visibility. We

show full results in Fig. 11. The DETR’s result does not change our conclusion that OC-cost’s detectors’ rankings are more stable than mAP.

### C. Interactive Demo

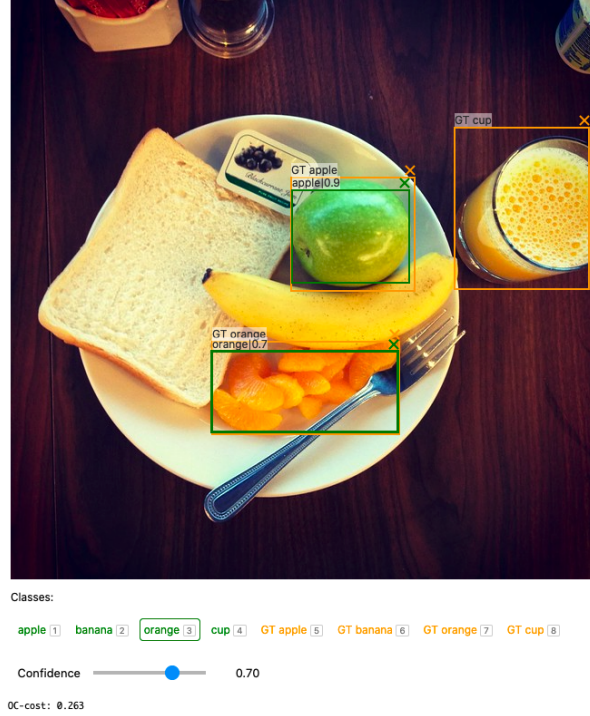
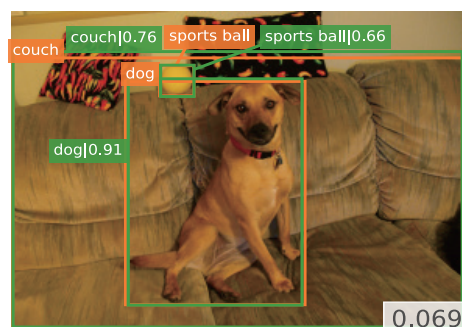
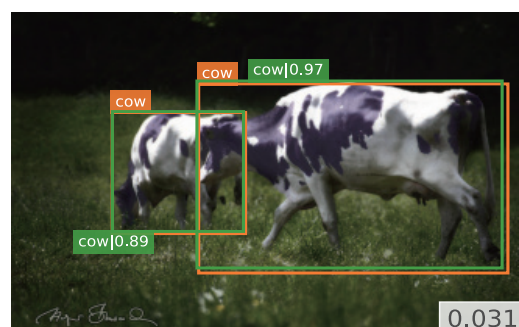
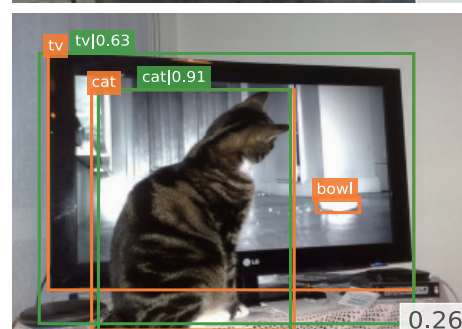
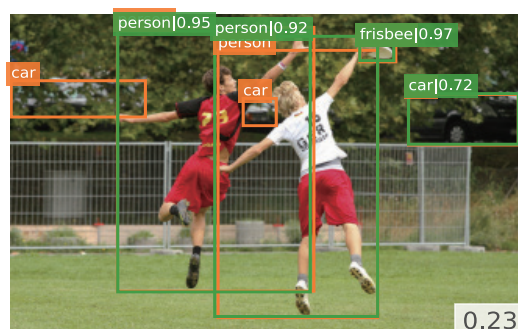
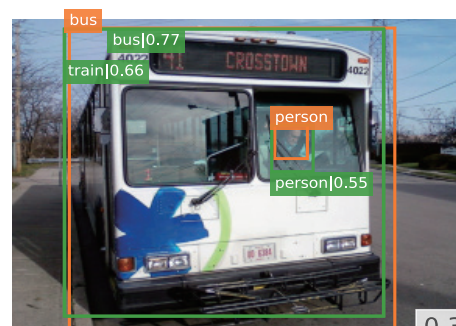
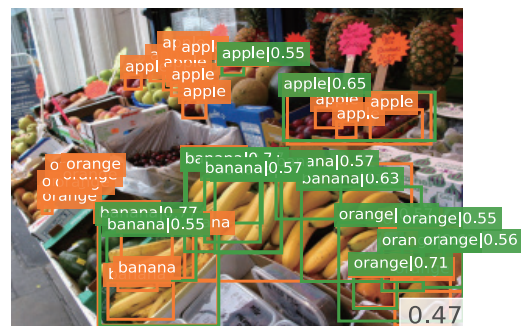
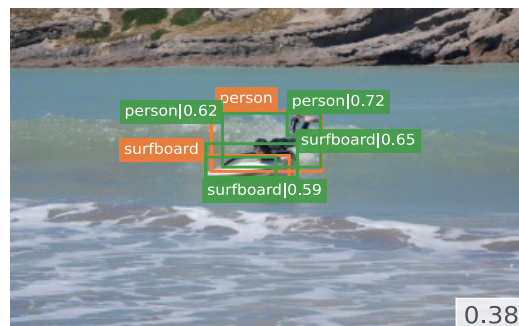
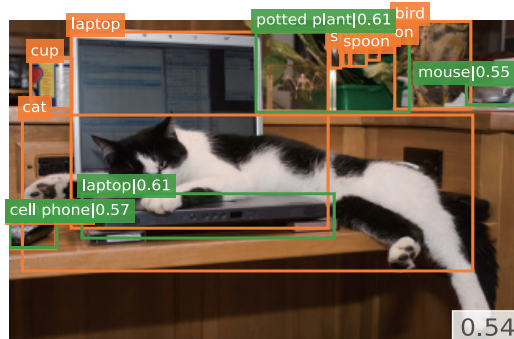
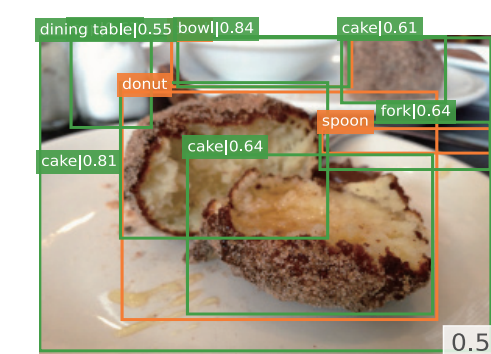


Figure 12. We can interactively give ground truths with orange boxes and detections with green ones. Once the ground truths and the detections are modified, corresponding OC-cost is displayed below the image.

We attach to this supplementary material a python notebook for an interactive demo. The screenshot of the demo is in Fig. 12. In the demo, OC-costs are computed for different detections and ground truths.

### D. OC-cost Examples

We showcase detection examples on MS-COCO dataset and corresponding OC-costs in Fig. 13. From top to bottom, the examples are displayed in the order of OC-cost. The parameters  $\lambda$  is 0.5, and  $\beta$  is 0.6. The detections (green) are produced by VNet, and NMS is tuned on OC-cost. Ground truths are represented by orange bounding boxes.



OC-COST

High

Low

(better)

Figure 13. OC-cost examples. The parameters  $\lambda$  is 0.5, and  $\beta$  is 0.6. The detections (green) are produced by VFNet, and NMS is tuned on OC-cost. Ground truths are represented by orange bounding boxes. OC-cost is displayed on the right bottom of each image.