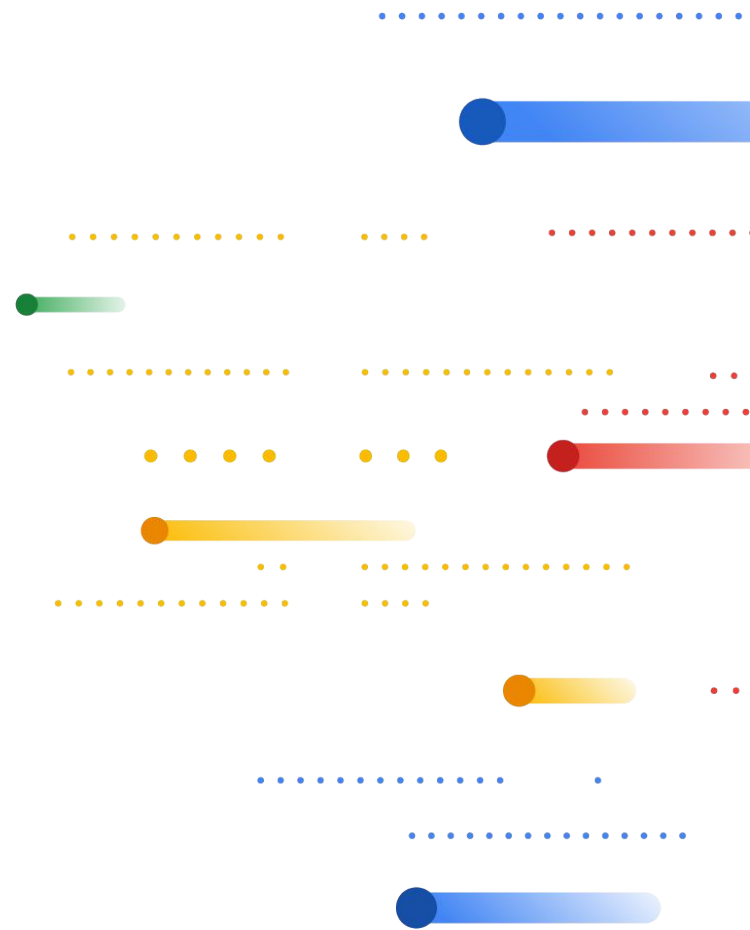


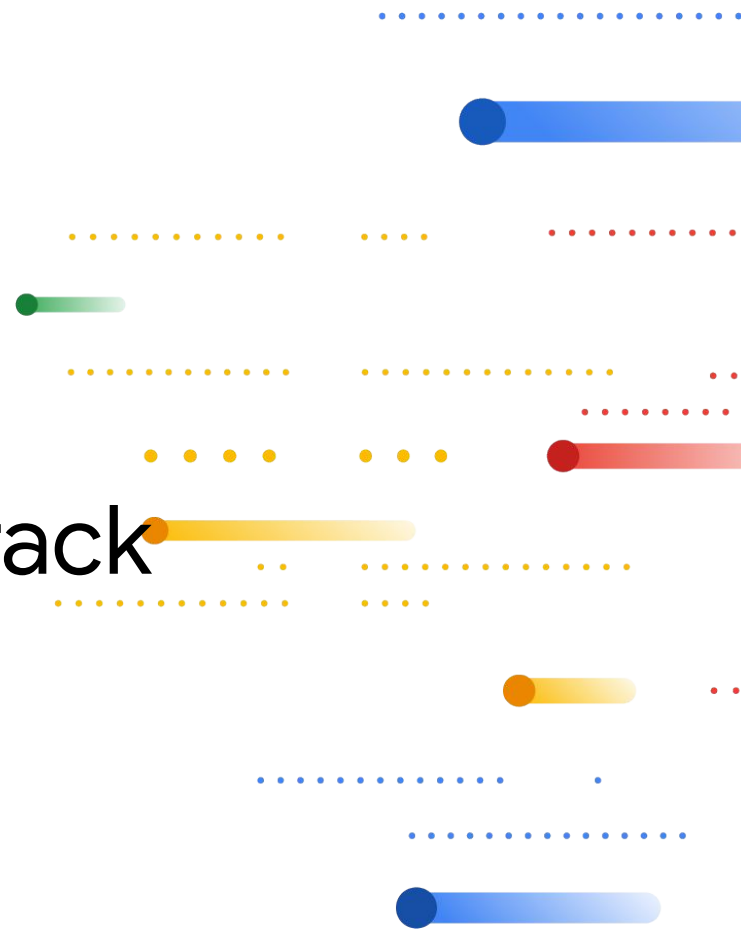
OPEN

MAGES

CHALLENGE 2019



Instance segmentation track



Participation and winning requirements

- Subset of Open Images V4/V5 used for training
- External data/pre-trained models are allowed but must be disclosed
- Evaluation server is hosted by Kaggle
- Full prize: 25K USD split between 5 winners
- Winner obligations:
 - Detailed, minimum 2-page description of method
- Winners encouraged:
 - Open-source their framework
 - Predictions for distillation

Why instance segmentation ?

- Datasets & Challenge help drive the field forward
 - Pushes model quality up
 - Increases usage of models in the wild
- Pascal, COCO, and Cityscapes had shown good success
 - However limited domains
- OpenImages boxes allowed us to create the newest largest dataset in the field
 - Compared to COCO:
 - 4x number of classes,
 - 3x number of instances,
 - 7x number of images
- We hope to see many applications spawn-off this new data



Dataset characteristics

Instance segmentation

Task: segment objects on an image

- object mask
- class label

Train set

- Built on top of box annotations
- 2,1M masks
- 848k images
- 300 classes

Validation set:

- 23k top quality masks



Training set:

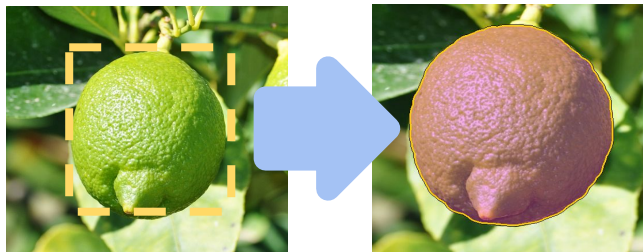
2.1M masks,

34s per mask

84% mIoU / 75% boundaries

(better than COCO polygons)

Interactive object segmentation



Training set:

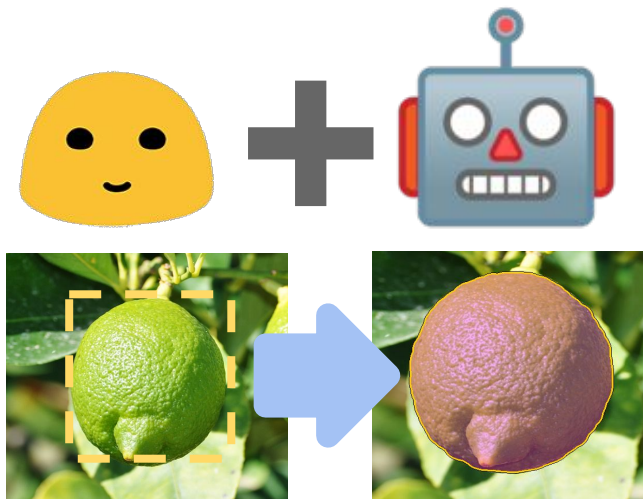
2.1M masks,

34s per mask

84% mIoU / 75% boundaries

(better than COCO polygons)

Interactive object segmentation



Challenge val & test set:

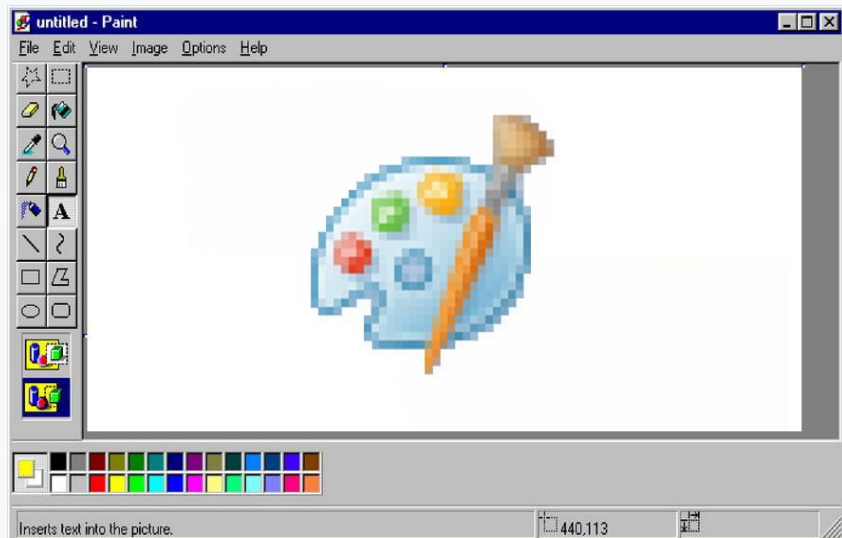
23k on val

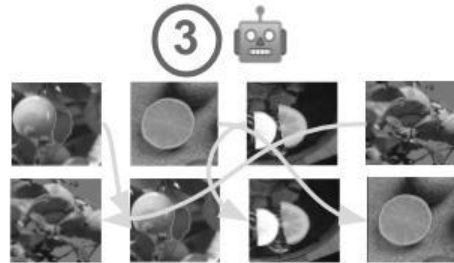
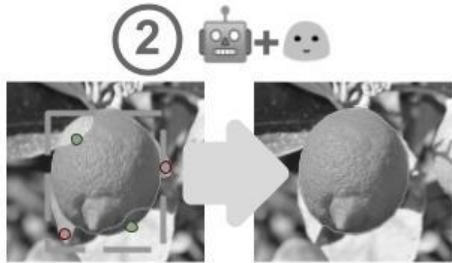
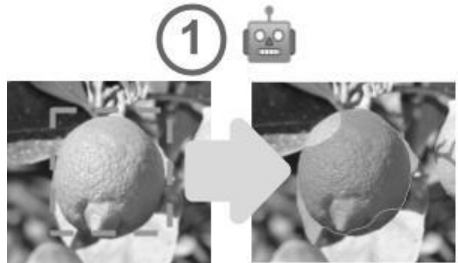
136s per mask

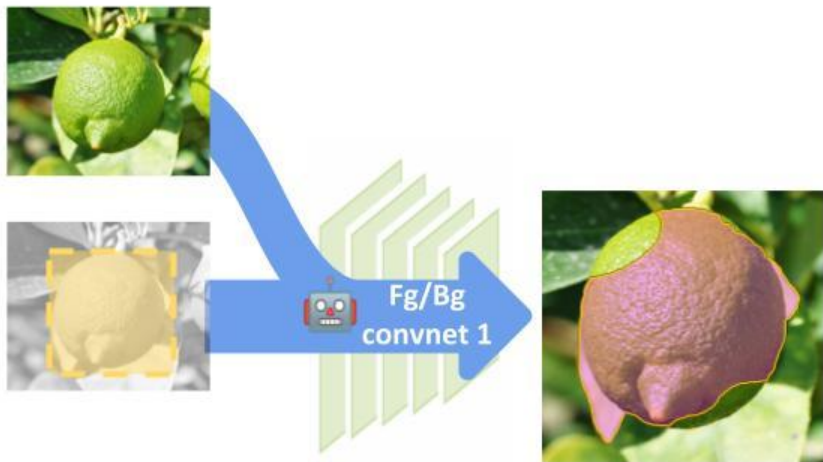
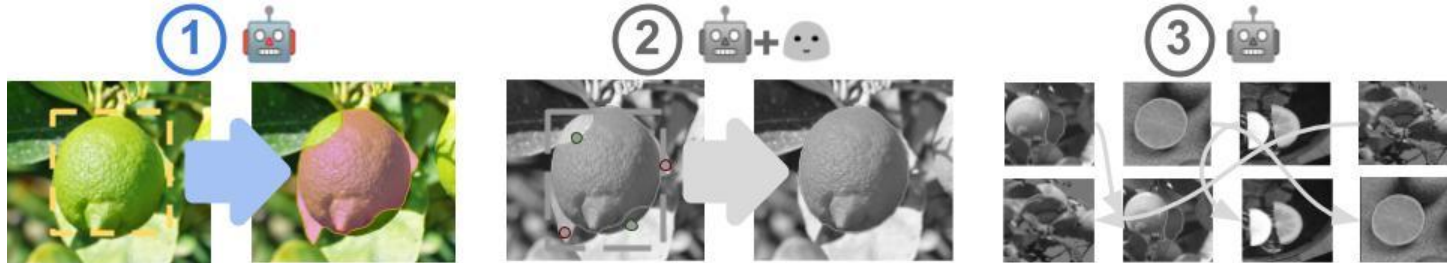
90% mIoU / 79% boundaries

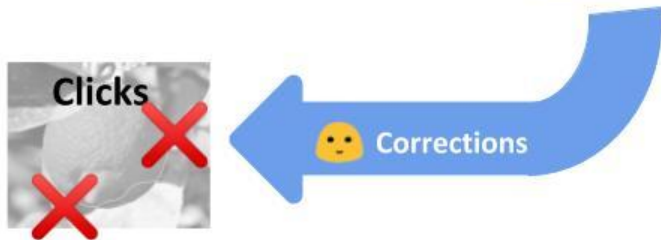
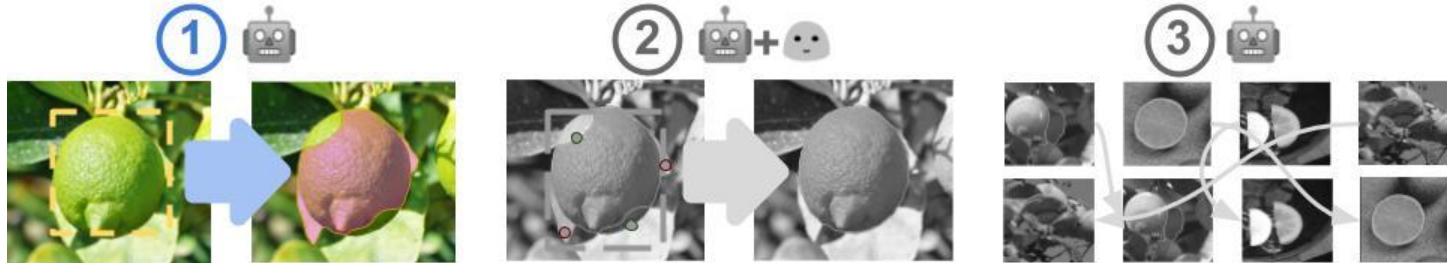
(self-agreement upper-bound)

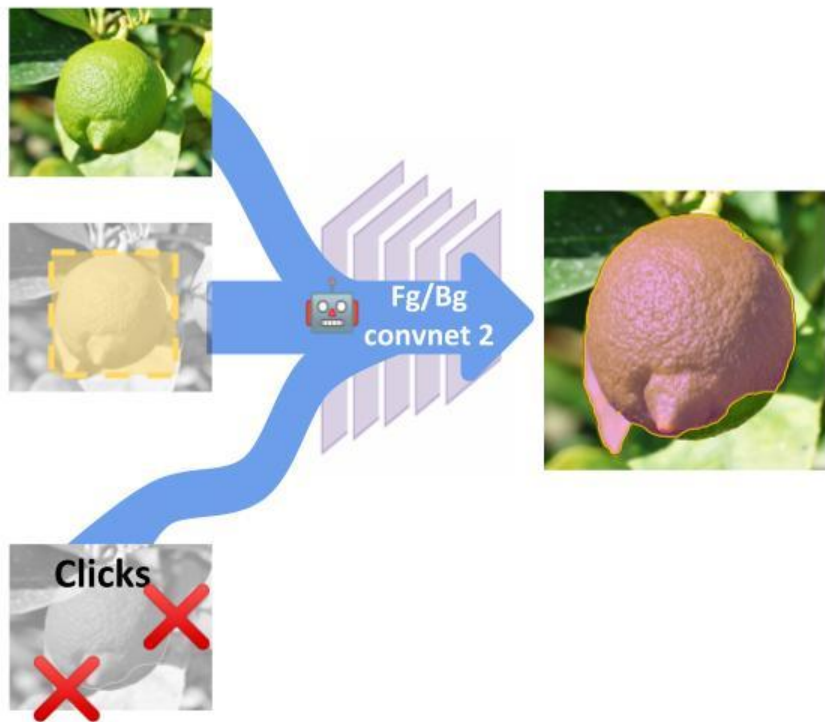
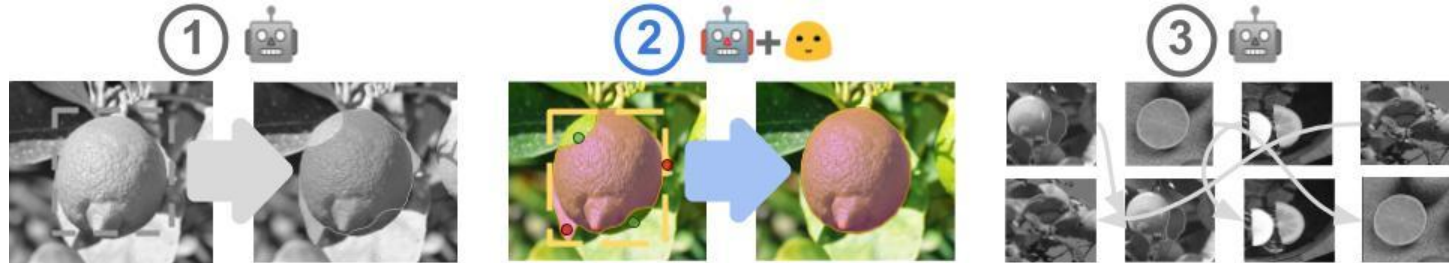
Free painting annotation

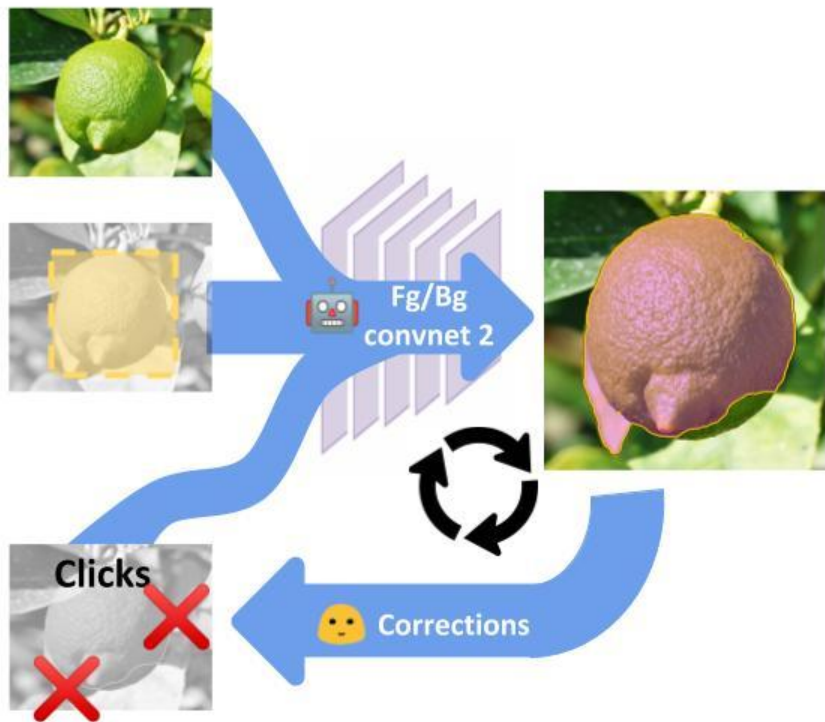
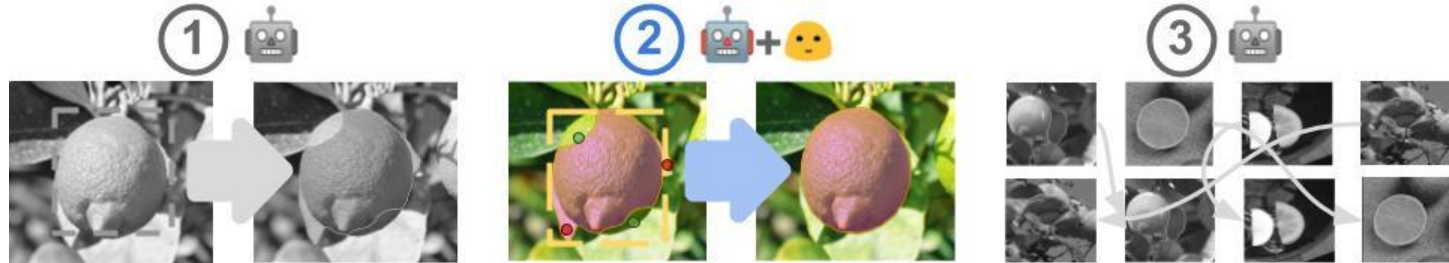


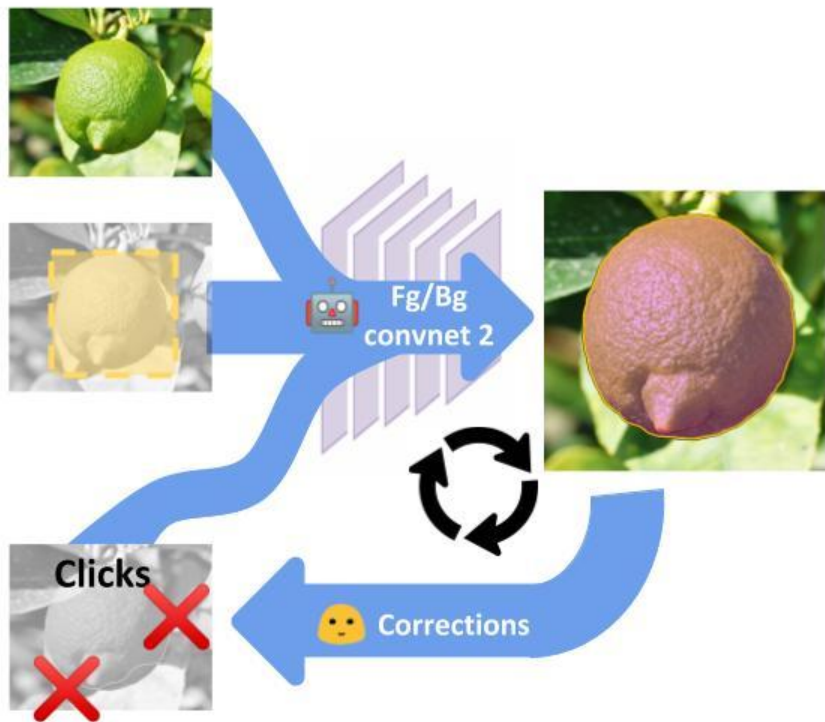
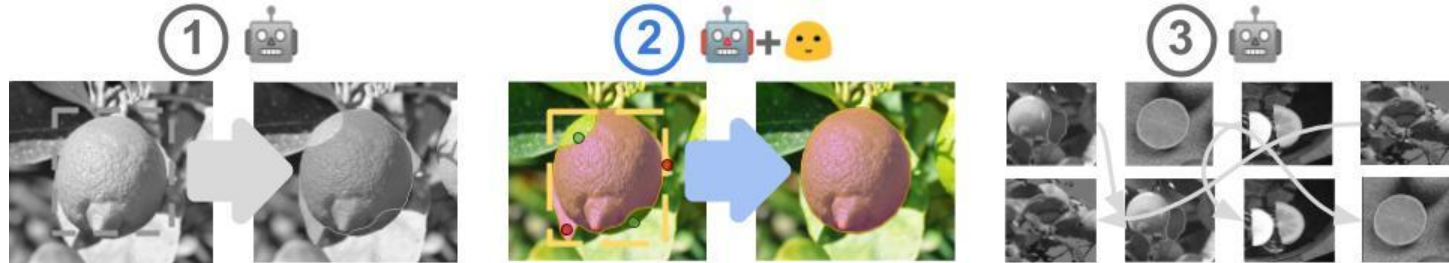


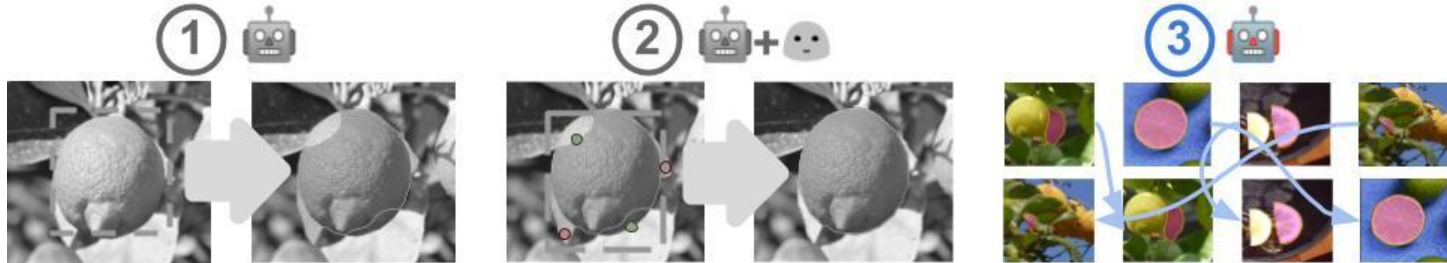












IoU score
RF regressor

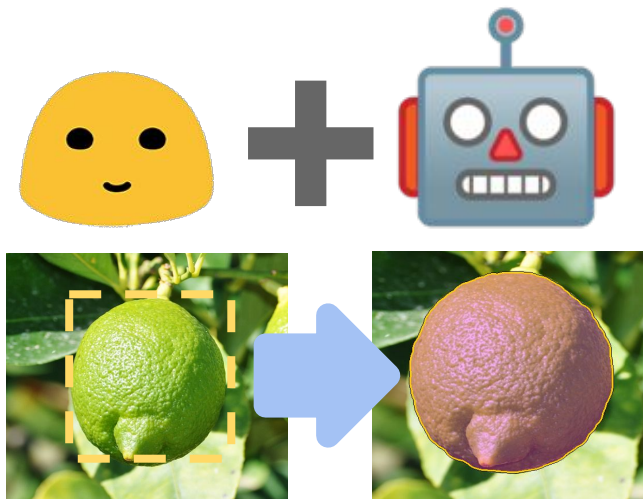


See our [CVPR19 paper for details](#).

Training set:

2.1M masks,
34s per mask
84% mIoU / 75% boundaries
(better than COCO polygons)

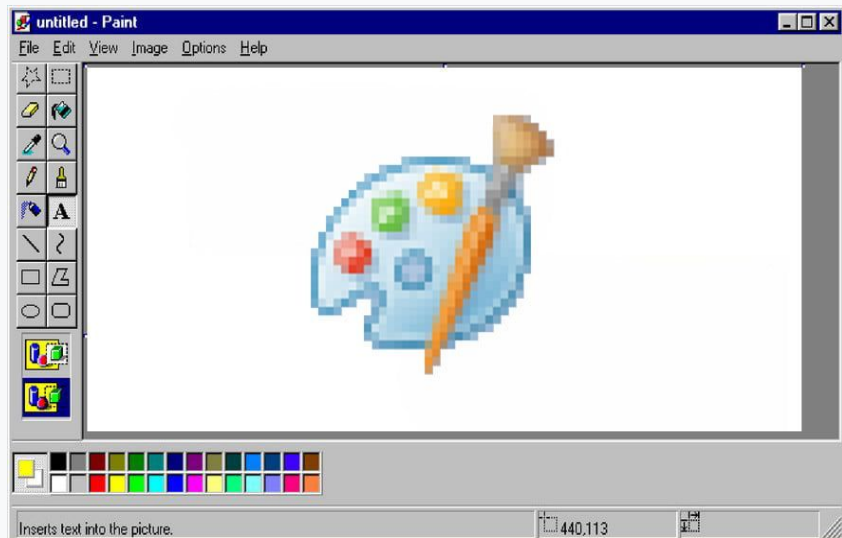
Interactive object segmentation



Challenge val & test set:

23k on val
136s per mask
90% mIoU / 79% boundaries
(self-agreement upper-bound)

Free painting annotation



Training set:

2.1M masks

34s per mask

84% mIoU / 75% boundaries

(better than COCO polygons)

Interactive object segmentation**Challenge val & test set:**

23k on val

136s per mask

90% mIoU / 79% boundaries

(self-agreement upper-bound)

Free painting annotation



Training set is data rich

Each image comes with:

- **Positive** and **negative** image-level labels (as well as machine class-scores)
- Each positive label of covered classes has bounding boxes
- Most bounding boxes of covered classes have a mask, and its corrective clicks, and predicted_iou



Positive: Carnivore, Cat, Food, Table, Animal

Negative: Sink, Human face



Evaluation protocol

Evaluation protocol

- Same approach as Object detection, we replace box IoU with masks IoU
- On COCO leaderboard we observed strong correlation between mAP@0.5, mAP@0.75 and mAP@0.5-0.95
→ we picked mAP@0.5 as suitable evaluation threshold
- Additional complications for boxes without masks (too small, or annotators omitted), and group-of boxes:
→ detection masks touching these boxes are ignored in evaluation

Public metric implementation is available as a part of [Tensorflow Object Detection API](#)

A decorative graphic in the top-left corner consisting of a yellow circle, a green arc, a red line, and a green line, all connected by dotted lines.

Results analysis

Overview

Number of teams with at least one submission: **193**

Number of teams with private mAP ≥ 0.1 : 45

External datasets:

Objects365, LVIS, OIDv4,
COCO, ImageNet

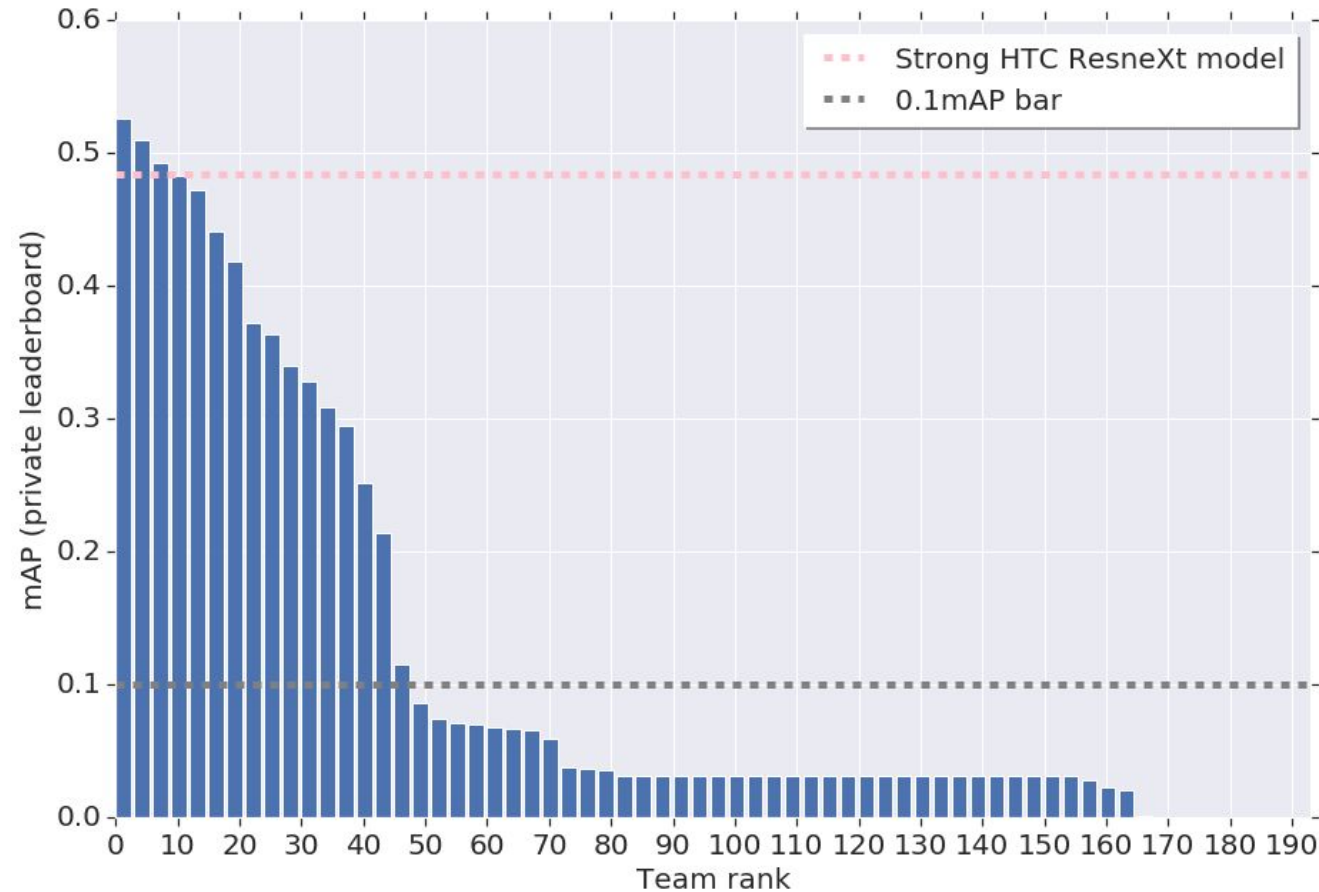
Deep learning frameworks:

PyTorch, Tensorflow, Keras, Chainer
Often via MMDetection, TensorPack

Base model architectures:

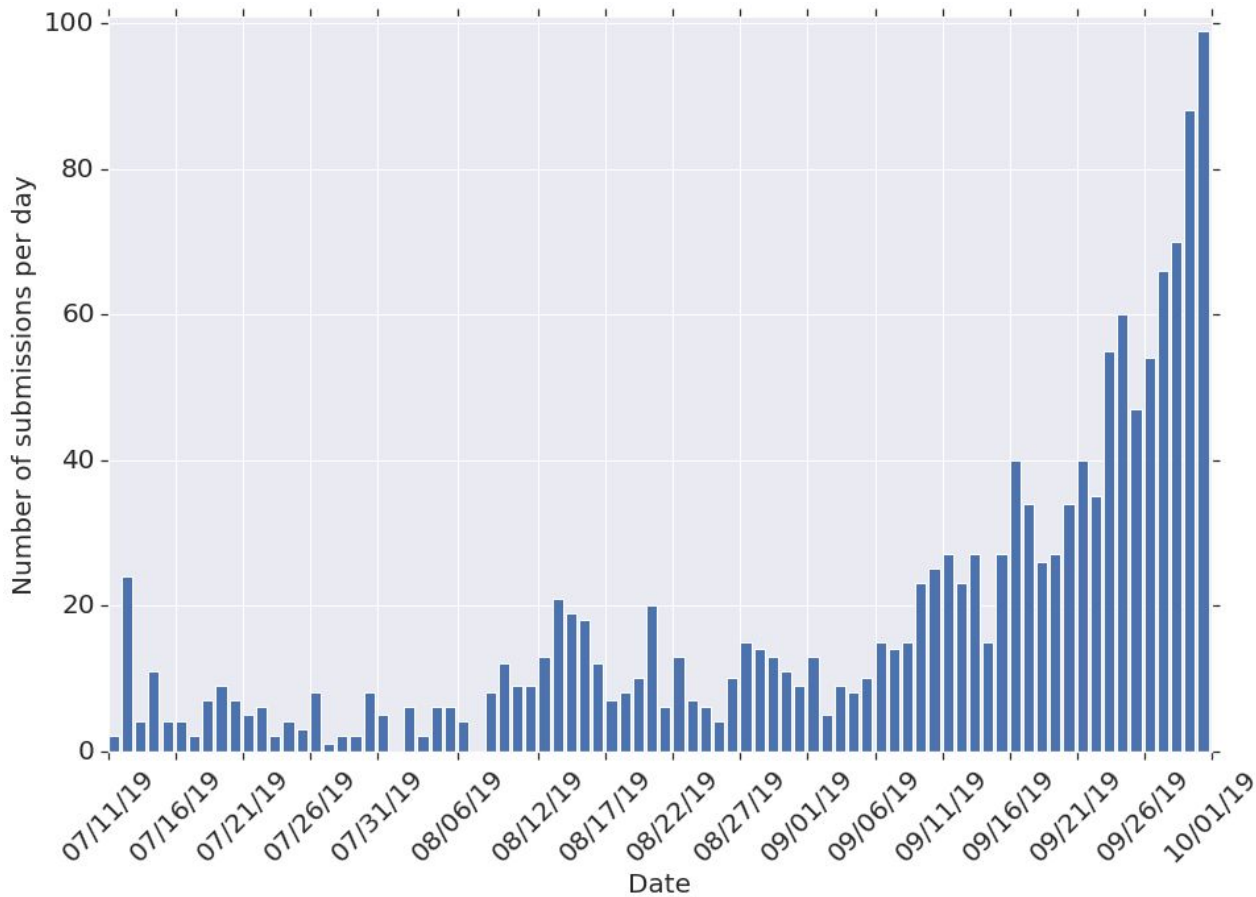
HTC, MaskRCNN, FasterRCNN, FastRCNN, Yolo, Retinanet, FPN, UNet, HRNet
RestNet, ResneXt, InceptionResnet, Darknet, Nas, SENet, EfficientNet

Submissions overview

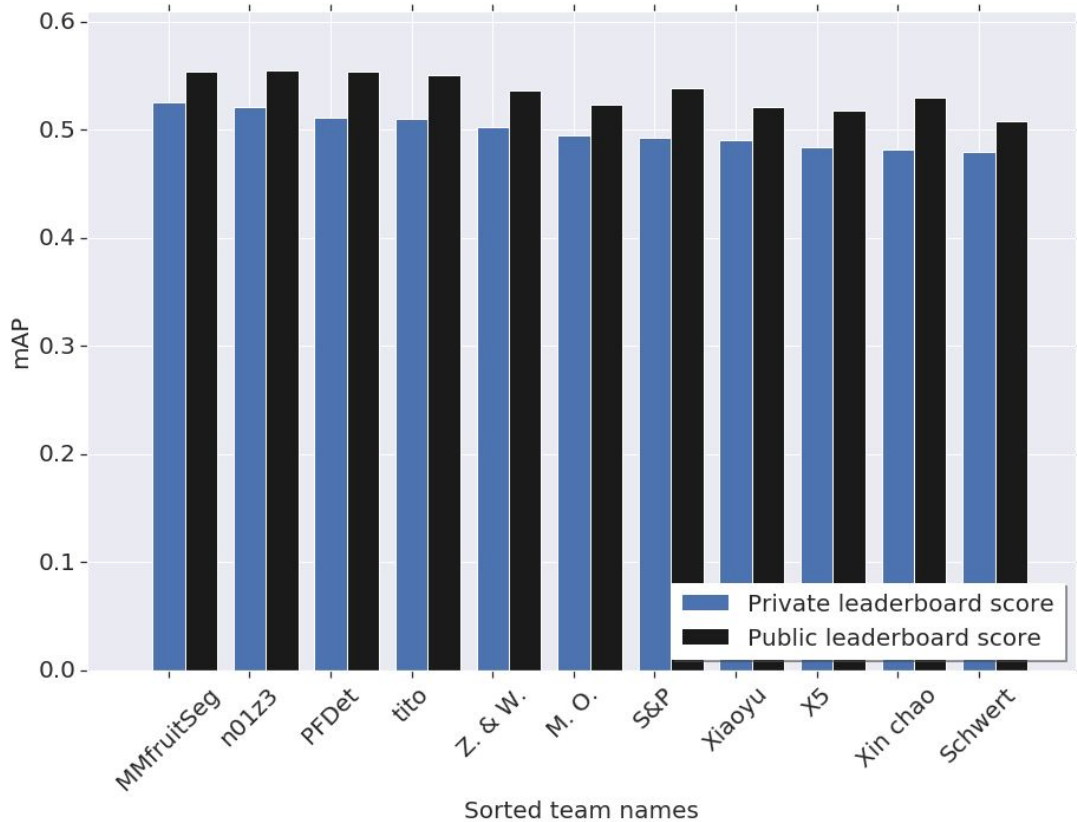


- Single strong model, without ensemble, reaches high performance.
- Top result is ~10 mAP points below object detection track. Indicates difficulty in segmentation.

Results analysis: number of submissions per day

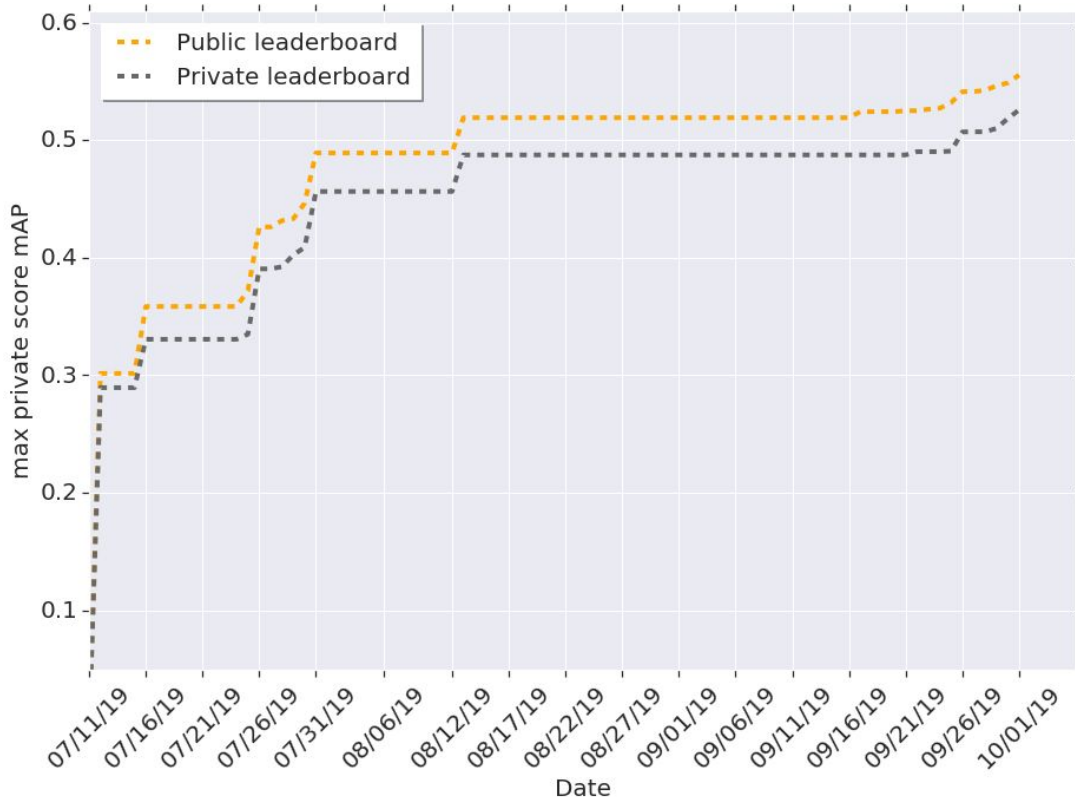


Results analysis: public vs private leaderboards



Public leaderboard: 20% of Challenge test set, Private leaderboard: 80%

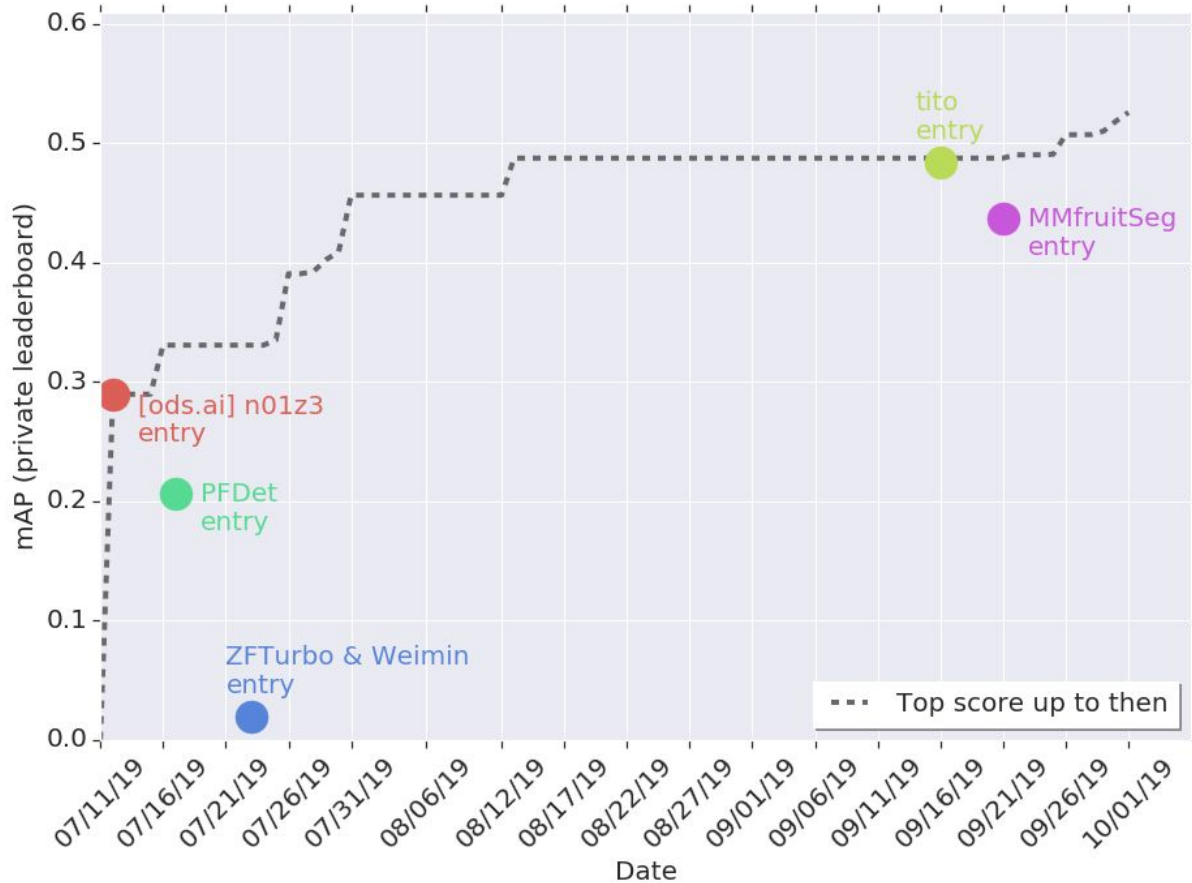
Results analysis: public vs private leaderboards



Public and private leaderboard are perfectly correlated.

Public leaderboard: 20% of Challenge test set, Private leaderboard: 80%

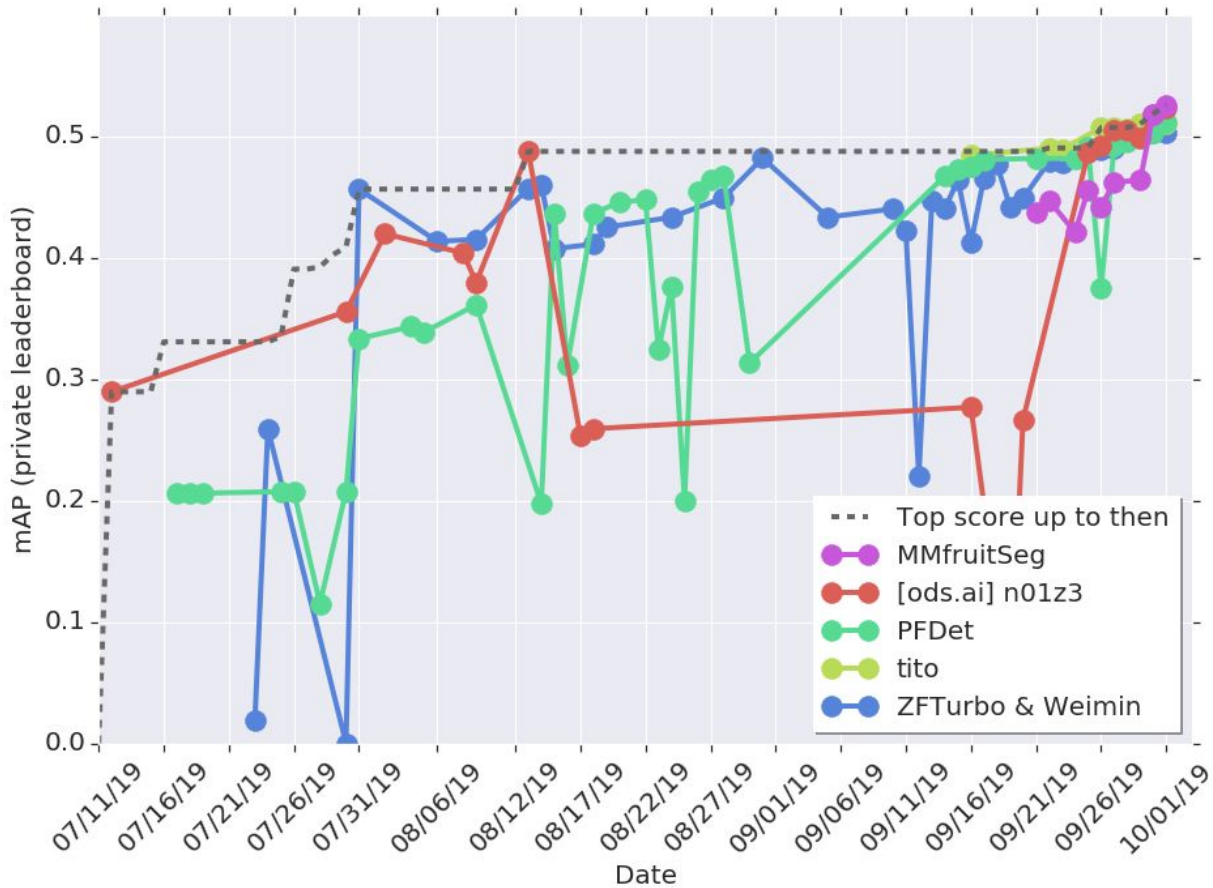
Results analysis: evolution of maximal leaderboard score



Dots: winners entering the competition.

- Long plateau
- Late entry winners

Results analysis: evolution of scores (winning teams)



Non-monotonic progression within most teams.

Winning teams: final results

| Team | Public score | Private score | | Num. entries | In OD track last year | In OD track this year |
|------------------|--------------|---------------|--|--------------|--------------------------|--------------------------|
| MMfruitSeg | 0.5539 | 0.5257 | | 19 | <input type="checkbox"/> | ✓ |
| [ods.ai] n01z3 | 0.5552 | 0.5213 | | 56 | ✓ | ✓ |
| PFDet | 0.5533 | 0.5110 | | 95 | ✓ | ✓ |
| tito | 0.5500 | 0.5098 | | 37 | ✓ | ✓ |
| ZFTurbo & Weimin | 0.5368 | 0.5022 | | 62 | ✓ | <input type="checkbox"/> |

- Only minuscule relative score changes between public and private leaderboard
- More entries does not lead to better results
- Most participants are well experienced on these problems

General trends from the methods descriptions

- Large ensembles (15+ models), all kinds of backbone networks
- Methods addressed class imbalance
- Most team conscious of computing time for fast iterations
- Use of high-level detection libraries: MMDetection, TensorPack

Questions?

**Next - presentations by winning teams
[PFDet, n01z3, MMfruitSeg]**

Today's program

| Time | Section |
|---------------|--|
| 13:30 - 13:40 | Overview of the Open Images Challenge |
| 13:40 - 14:00 | Object detection track - settings, metrics, winners, analysis, comparison to the previous year |
| 14:00 - 14:45 | Presentations by three winners of the Object detection track |
| 14:50 - 15:05 | Instance segmentation track - settings, metrics, winners, analysis |
| 15:05 - 15:50 | Presentations by three winners of the Instance Segmentation track |
| 15:50 - 16:30 | Break and Poster session |
| 16:30 - 16:50 | Visual Relationship Detection track - settings, metrics, analysis, comparison to the previous year |
| 16:50 - 17:20 | Presentations by two winners of Visual Relationship Detection track |
| 17:25 - 17:30 | Concluding remarks |