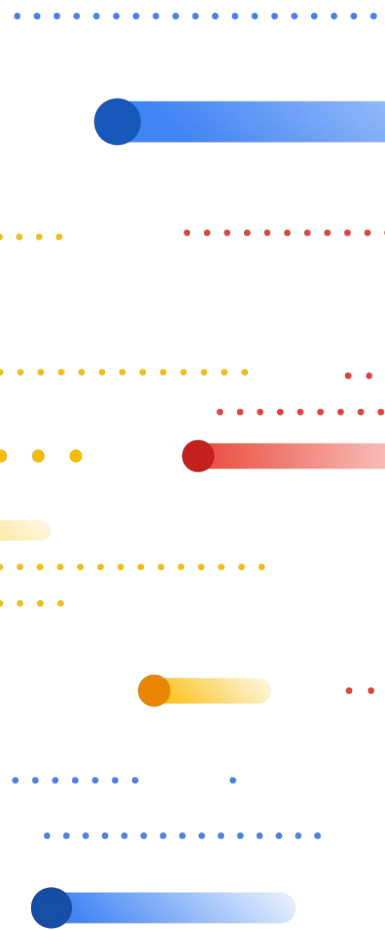


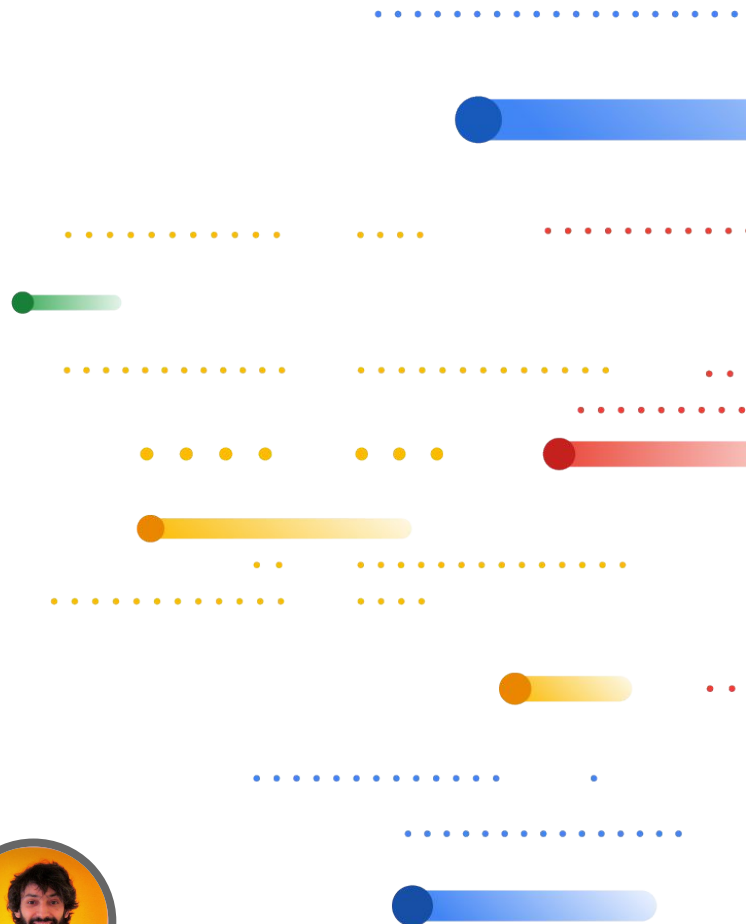
OPEN

MAGES

CHALLENGE 2018



Object Detection track



Outline

Object detection track overview

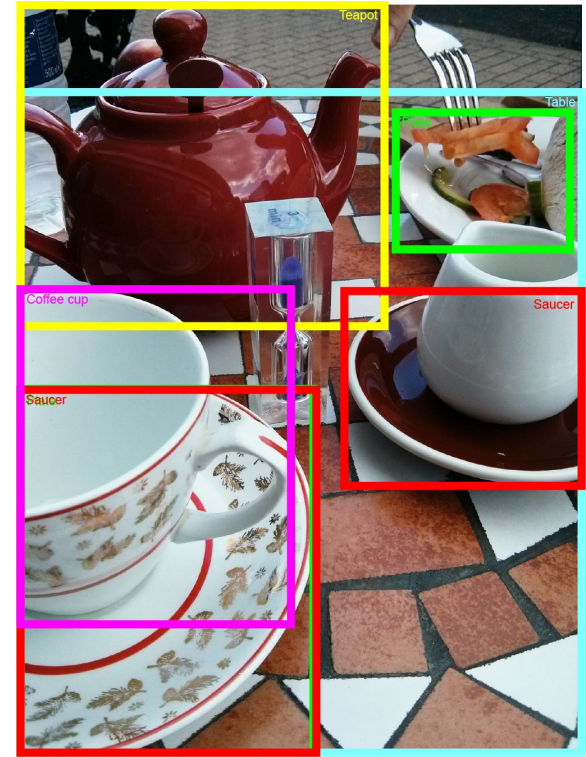
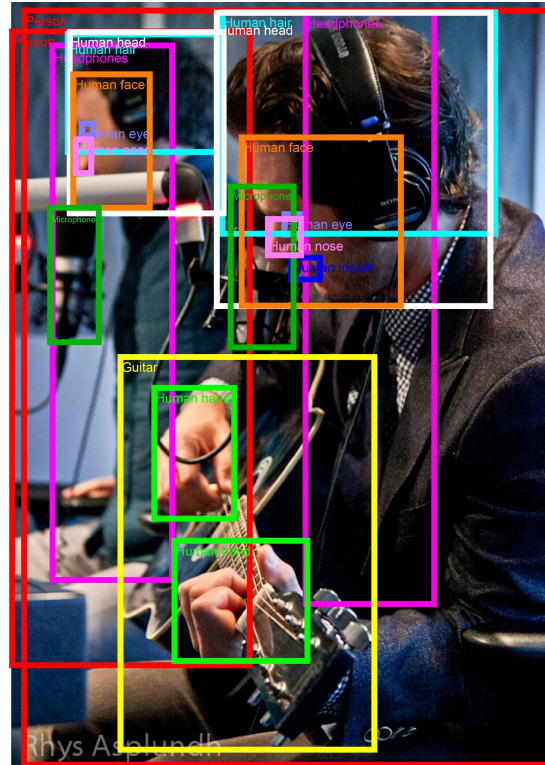
Dataset

Metrics

Result analysis

Object detection

New challenging
detection dataset
with bounding box
annotations of 500
classes



Participation and winning requirements

- Subset of Open Images V4 used for training
- External data/pre-trained models are allowed but must be disclosed
- Evaluation server is hosted by Kaggle
- Full prize: 30K USD split between 3 winners
- Winner obligations:
 - Detailed, minimum 2-page description of method
- Winners encouraged:
 - Open-source their framework
 - Predictions for distillation

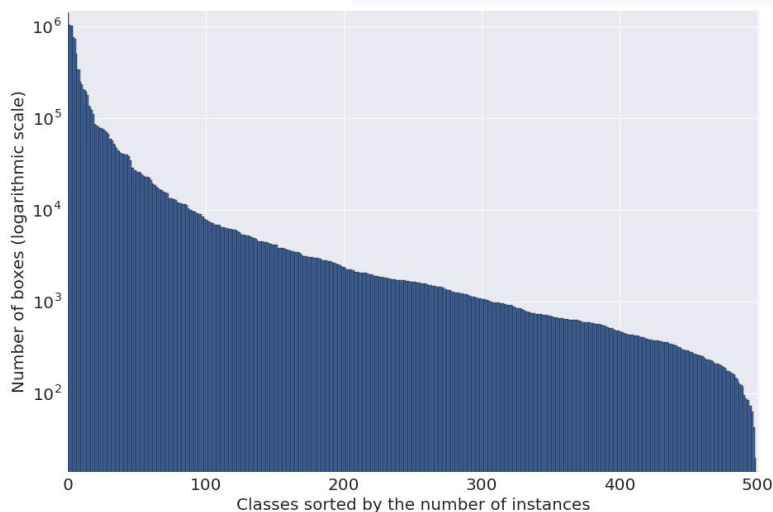
Dataset: statistics

Train set:

- 1,743,042 images
- 1,913,455 negative image-level labels
- 3,830,005 positive image-level labels
- 12,195,144 boxes
- 100k image subset recommended for validation

Test set:

- 100K images
- 20% in public split
- 80% in private split



Total: 500 classes

Leaf classes: 442 classes

Non-leaf classes: 58 classes



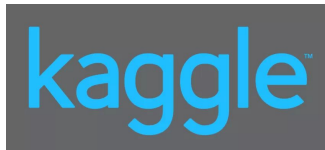
Evaluation

Properties of annotation process:

- Non-exhaustive image-level labeling
- Semantic hierarchy
- Group-of boxes

A modification of Mean Average Precision (mAP) takes those properties into account

Evaluation server hosted by [Kaggle](#)



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



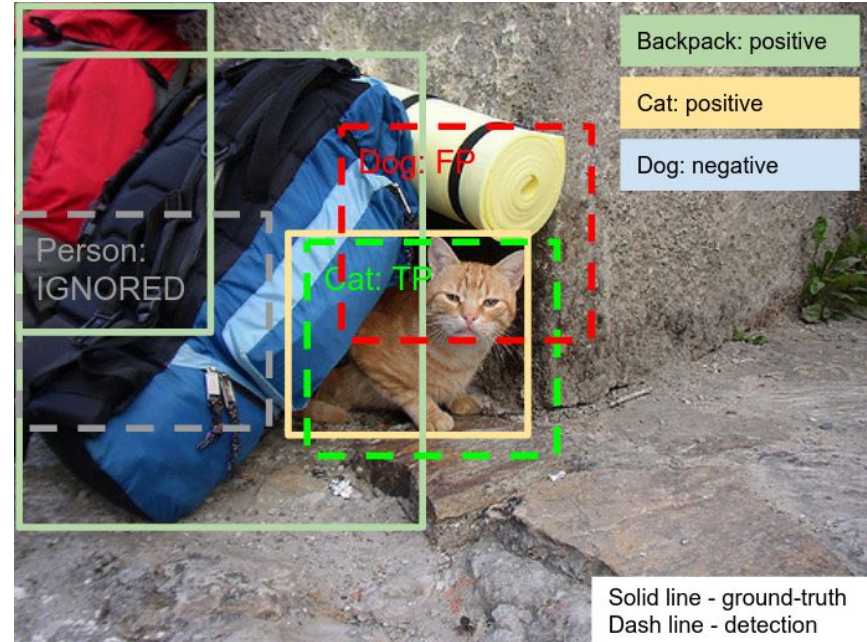
Evaluation metrics: Non-exhaustive image-level labeling

Ground-truth image labels have 3 cases:

- Positive: class is present
- Negative: class is absent
- Unannotated: we do not know

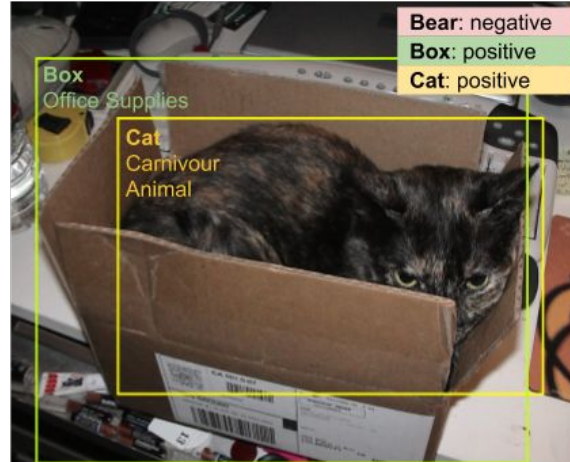
Ignore detections of unannotated classes

Rest as in PASCAL VOC Challenge



Evaluation metrics: Semantic hierarchy

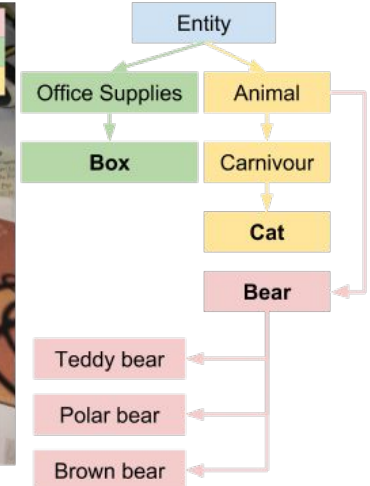
- Ground truth replicates boxes and image labels following the hierarchy
- AP is computed for both leaf and non-leaf classes.
- AP for non-leaf classes is evaluated on both boxes of this class and all descendant class boxes
- Participants required to output multiple boxes on same object



Bear: negative
 Teddy bear: negative
 Polar bear: negative
 Brown bear: negative

Car: positive
 Carnivour: positive
 Animal: positive

Box: positive
 Office supplies: positive



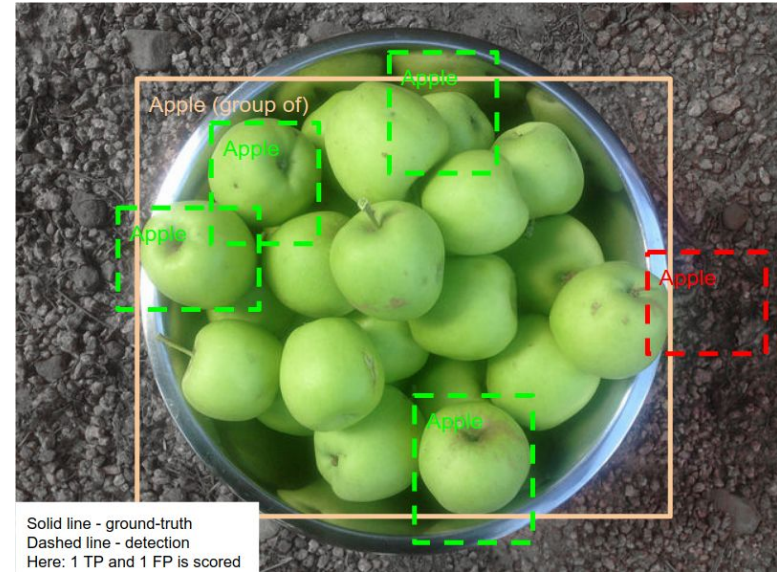
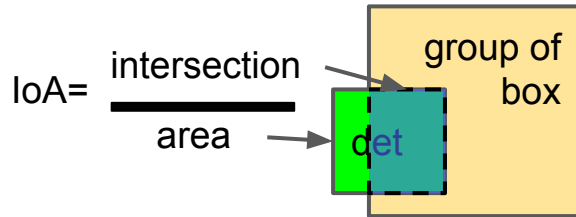
Evaluation metrics: Group-of boxes

A group-of box:

- contains >5 instances
- Instances occlude each other

Matched box:

$\text{IoA}(\text{group of box, detection}) > 0.5$



The highest-scoring detection is a TP. Rest is ignored

Results analysis: overview

Number of teams with at least one submission: **454**

Evaluation server days: **59**

External datasets/pre-trained models used:

- OpenImagesV4
- ImageNet
- COCO

Base model architectures:

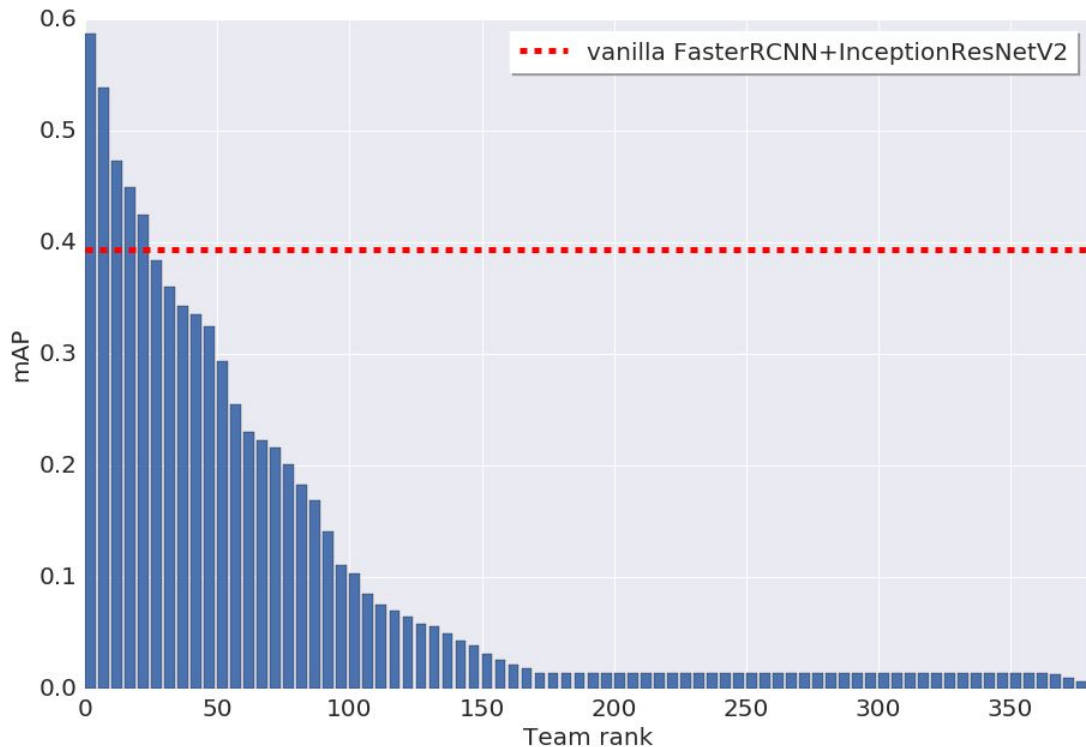
- ResNets, YOLO, Darknet, SNet, Retinanet ...

Deep learning frameworks:

- Tensorflow Object Detection API, Detectron, Cadene (pyTorch), fastai library, ImageAI, ChainerCV, TensorFlow-Slim, Keras, MXNet

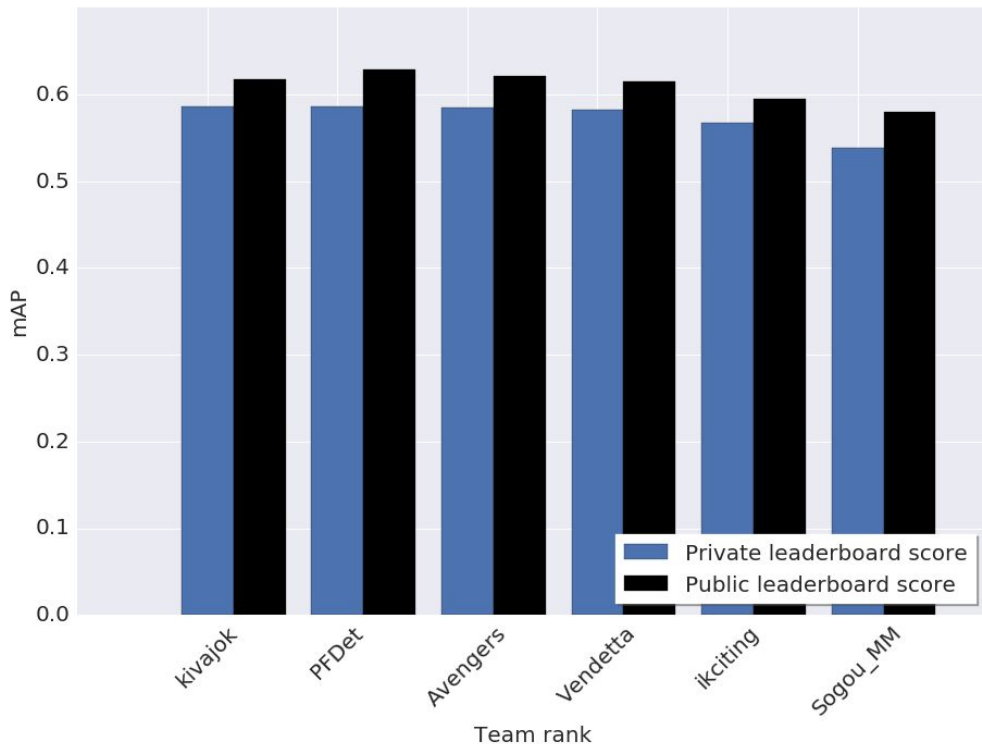
Results analysis: teams

Number of teams: **454**
Number of teams above
baseline model: **23**

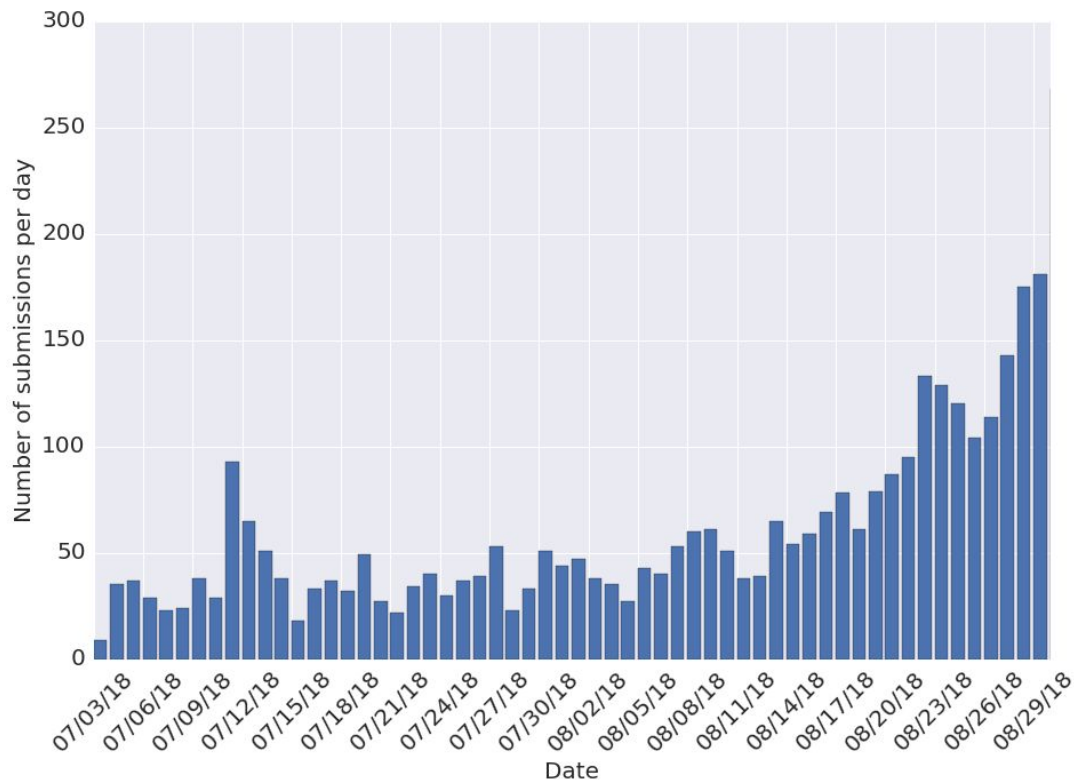


Results analysis: public vs private leaderboards

Public leaderboard: 20%
Private leaderboard: 80%

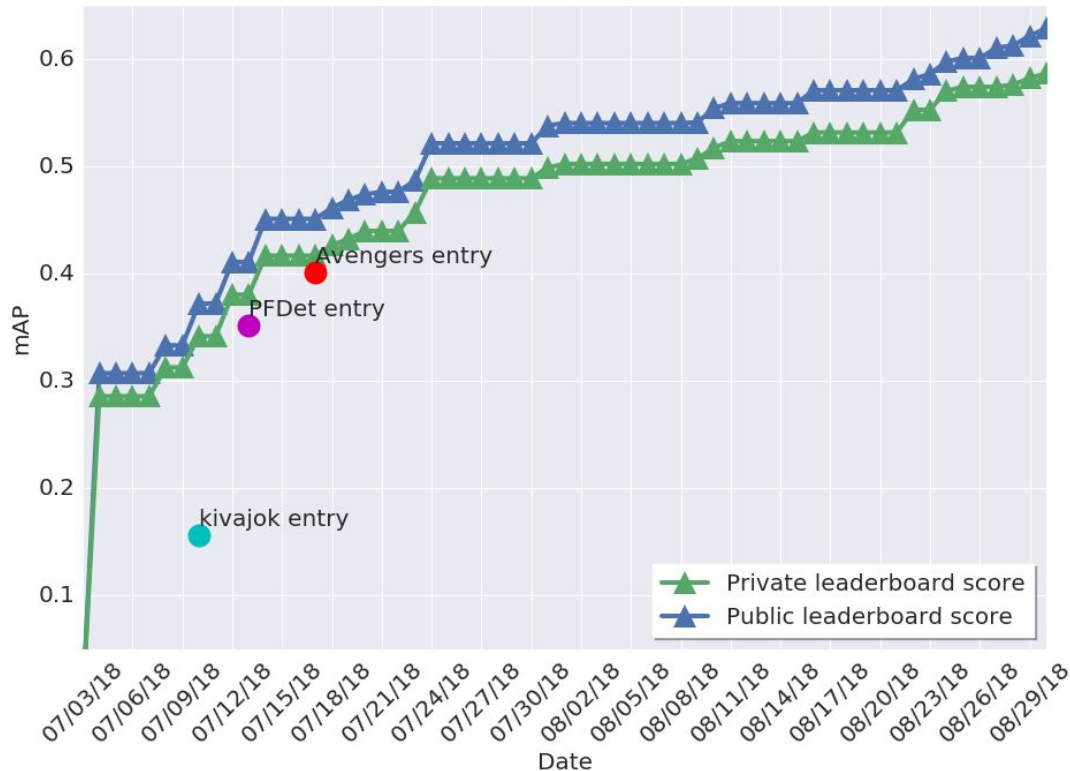


Results analysis: number of submissions per day



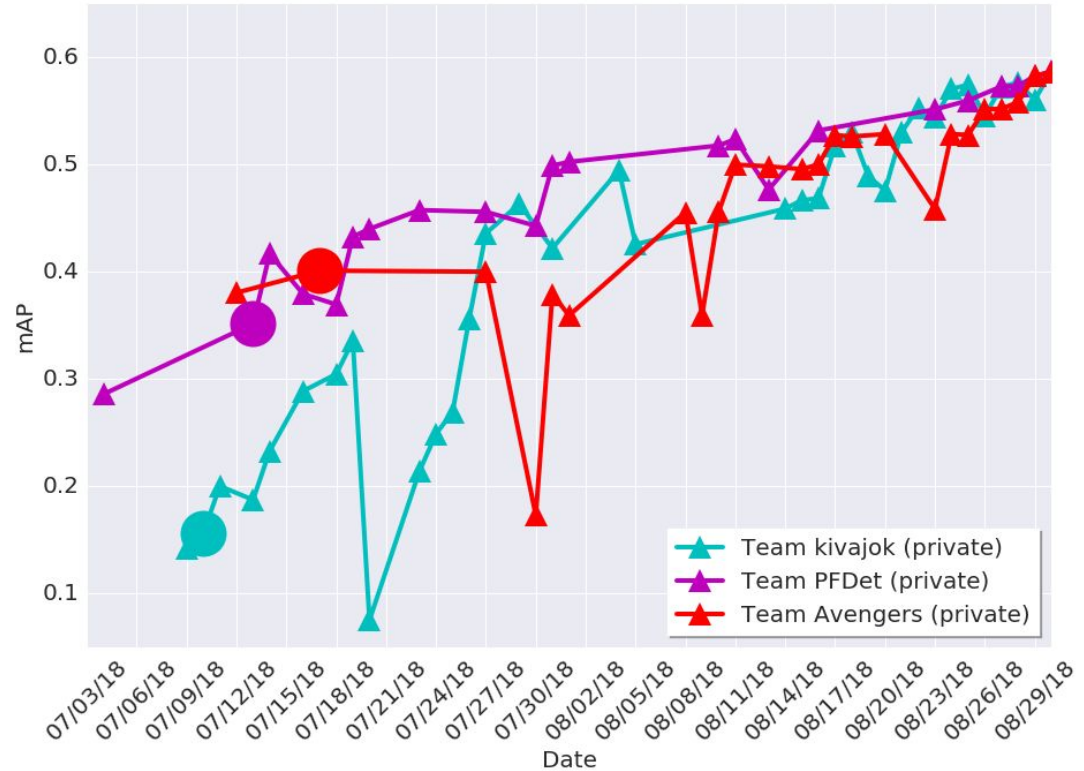
Results analysis: evolution of maximal leaderboard score

Dots: winners entering the competition



Results analysis: evolution of scores (winning teams)

Evolution of the private
leaderboard score per day



Result analysis: which classes work and which do not

Winning teams: final results

Team	Place	Public leaderboard	Private leaderboard
kivajok	1	0.61707	0.58657
PFDet	2	0.62882	0.58634
Avengers	3	0.62161	0.58616

Winning models

Commonalities:

- Used ensemble of various size
- Addressed class imbalance
- Architecture: FasterRCNN with additions

Questions?

Next - presentations by winning teams