

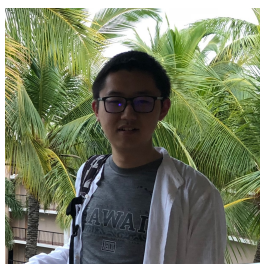
# Solution for Large-Scale Hierarchical Object Detection Datasets with Incomplete Annotation and Data Imbalance

Video Plus Team

Yuan Gao, Xingyuan Bu, Yang Hu, Hui Shen, Ti Bai, Xubin Li, Shilei Wen

# Team Members

## Video Plus Team



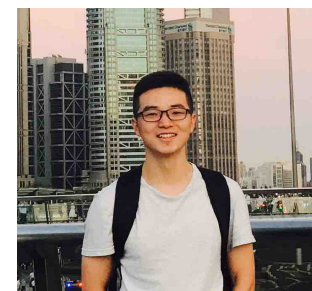
Yuan Gao



Xingyuan Bu



Yang Hu



Hui shen



Ti Bai



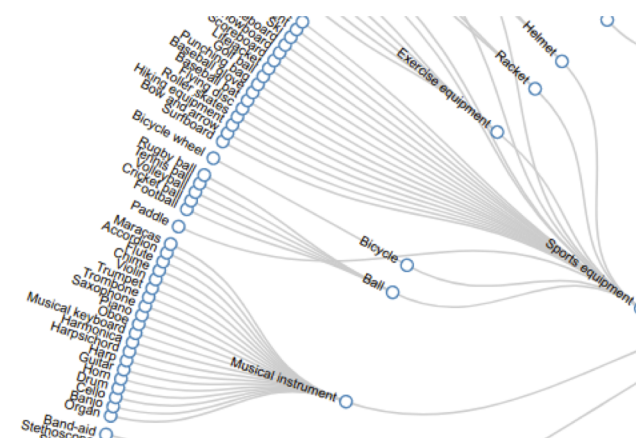
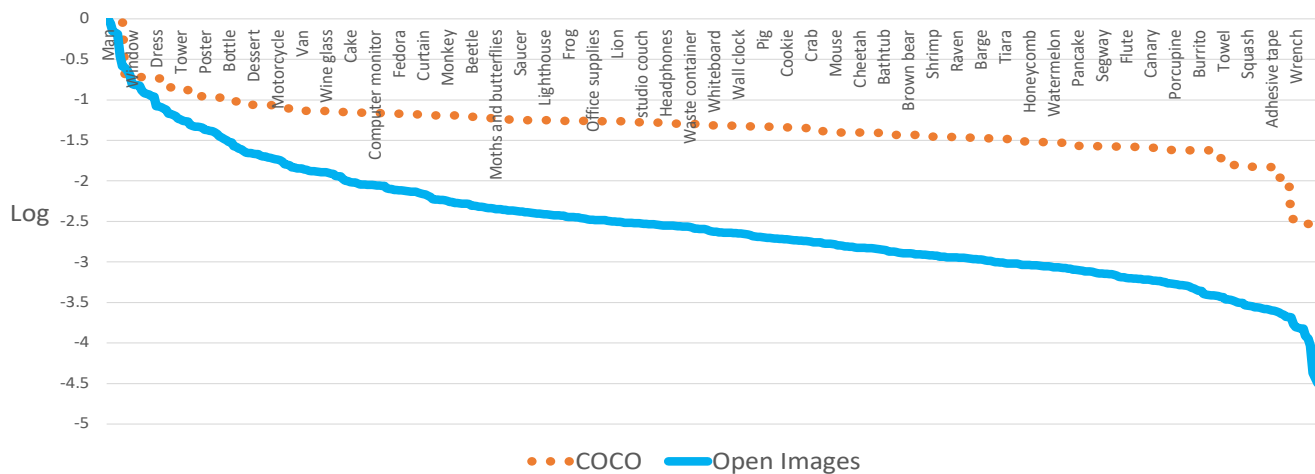
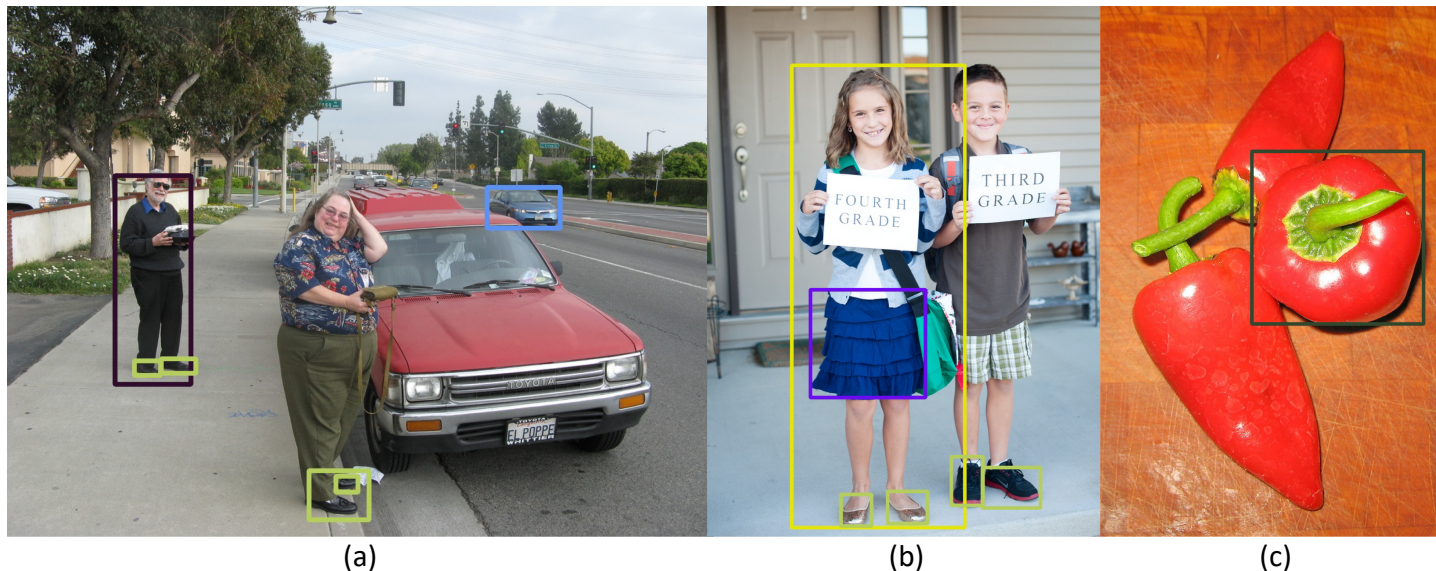
Xubin



Shilei Wen

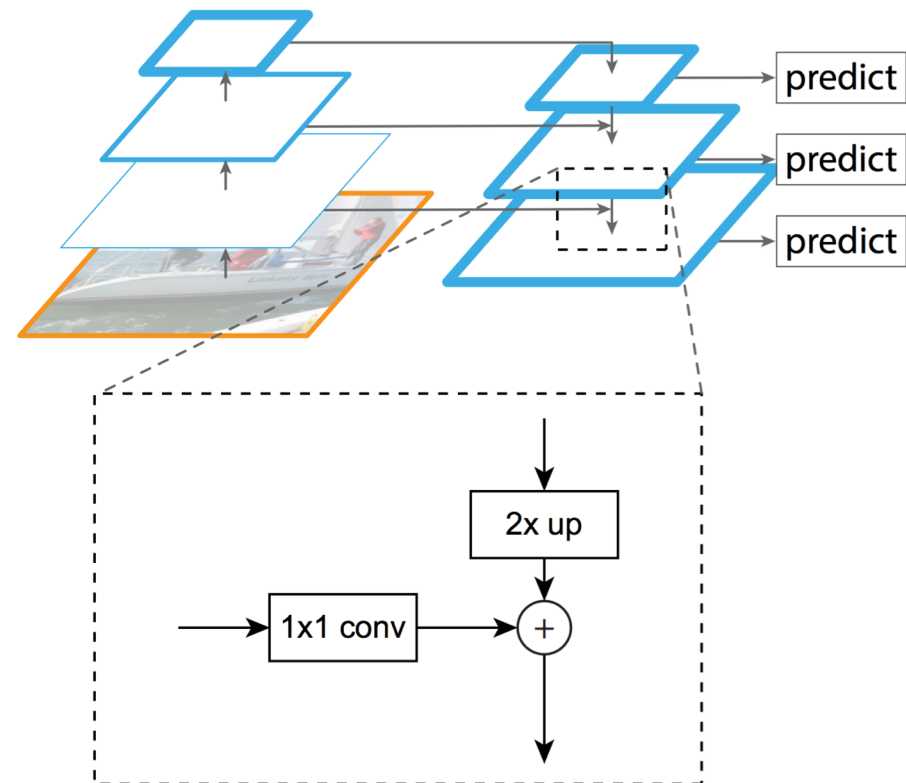
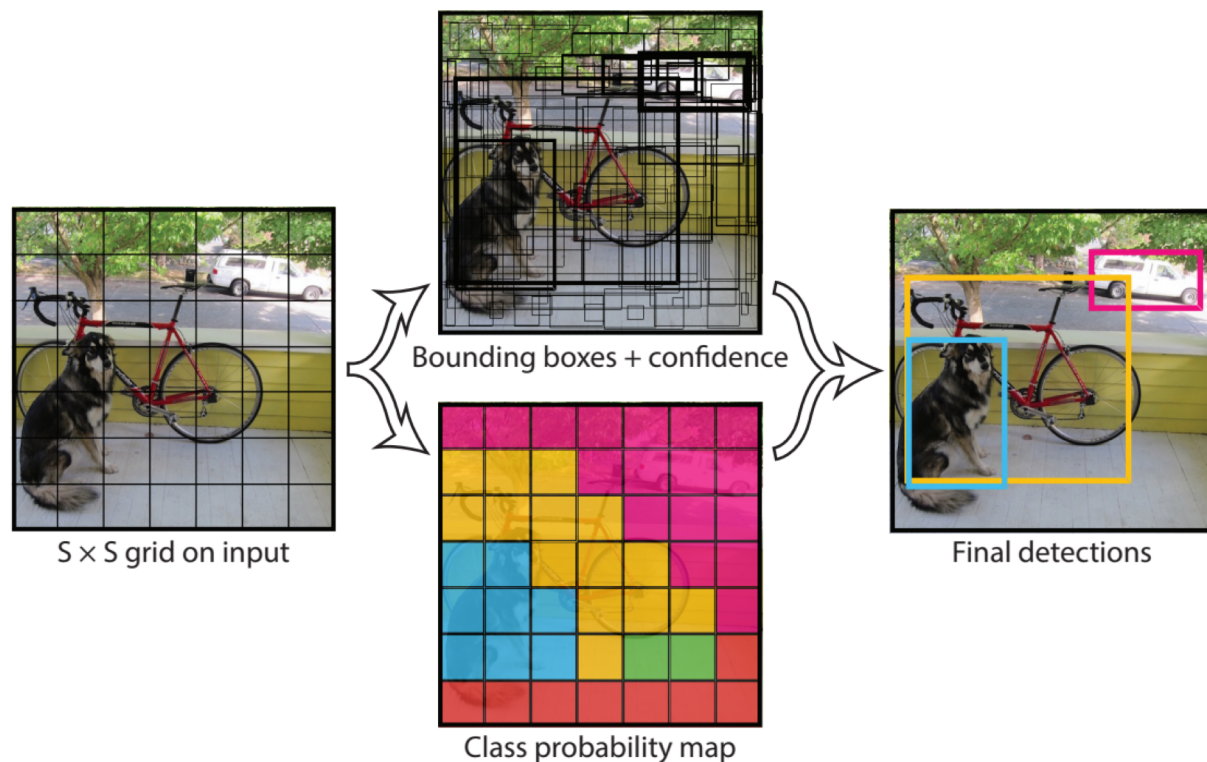
# Open Images Dataset

- Large-scale
- Annotation incompleteness
- Data imbalance
- Hierarchical tag system



Krasin, Ivan, et al. Openimages: A public dataset for large-scale multi-label and multi-class image classification. Dataset available from <https://github.com/openimages>

# Object Detection

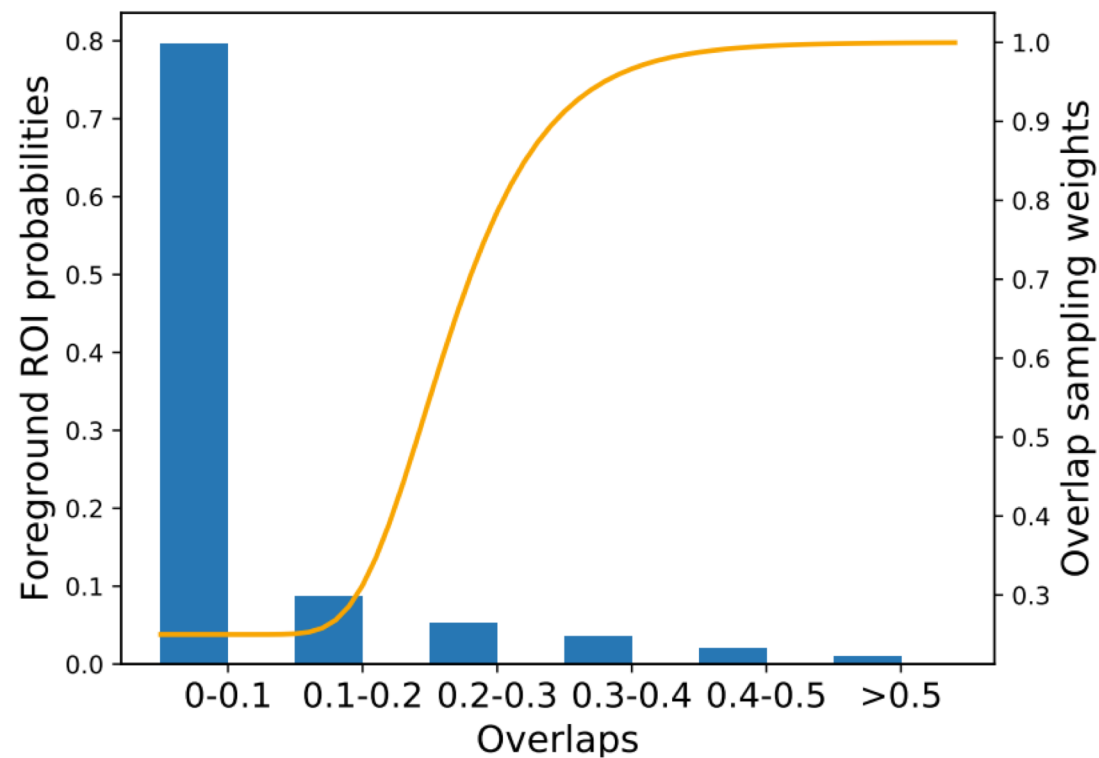


[1] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. CVPR 2016

[2] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan. Feature Pyramid Networks for Object Detection. CVPR 2017

# SNIPER & Soft Sampling

Annotation  
Incompleteness



[1] B. Singh, M. Najibi, and L.S. Davis. Sniper: Efficient multi-scale training. arXiv preprint arXiv:1805.09300, 2018

[2] Z.e.a. Wu. Soft sampling for robust object detection. arXiv preprint arXiv:1806.06986, 2018

# Class Aware Sampling

Data Imbalance

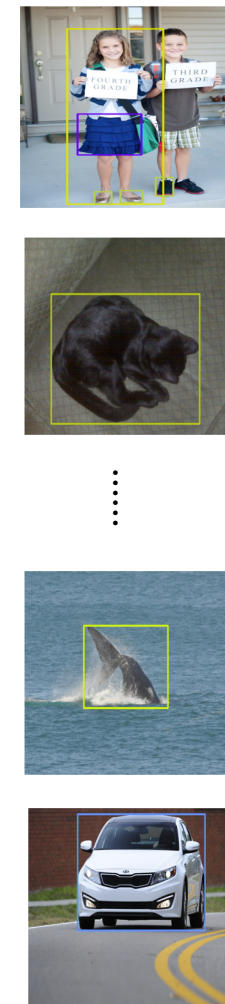
500 Classes

Person
Tree
Face
Car
Boat
.....
.....
Screwdriver
Flashlight
cooker
Torch

Category  
Sampling

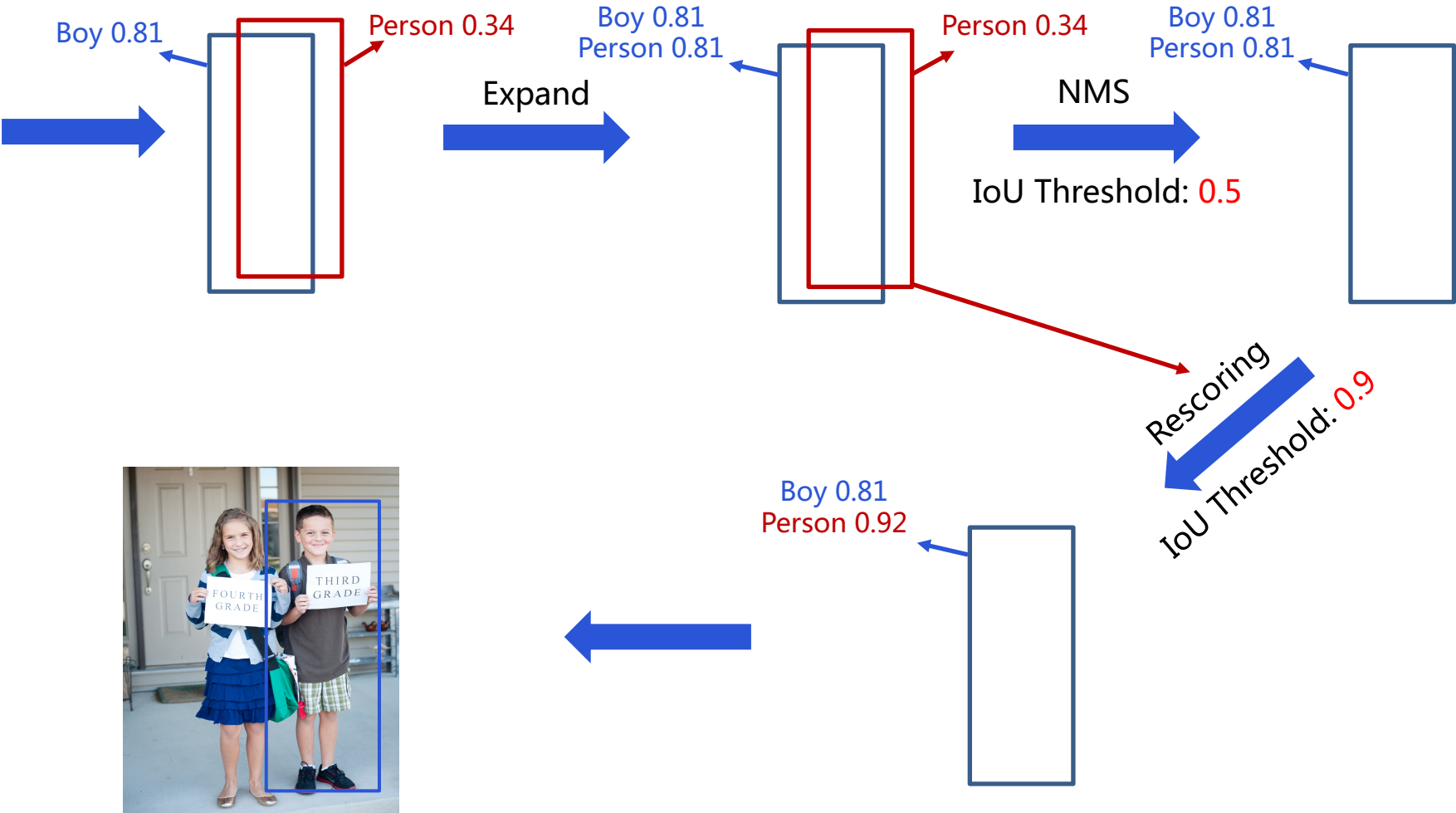
Person
Cat
.....
.....
Whale
Car

Image  
Sampling



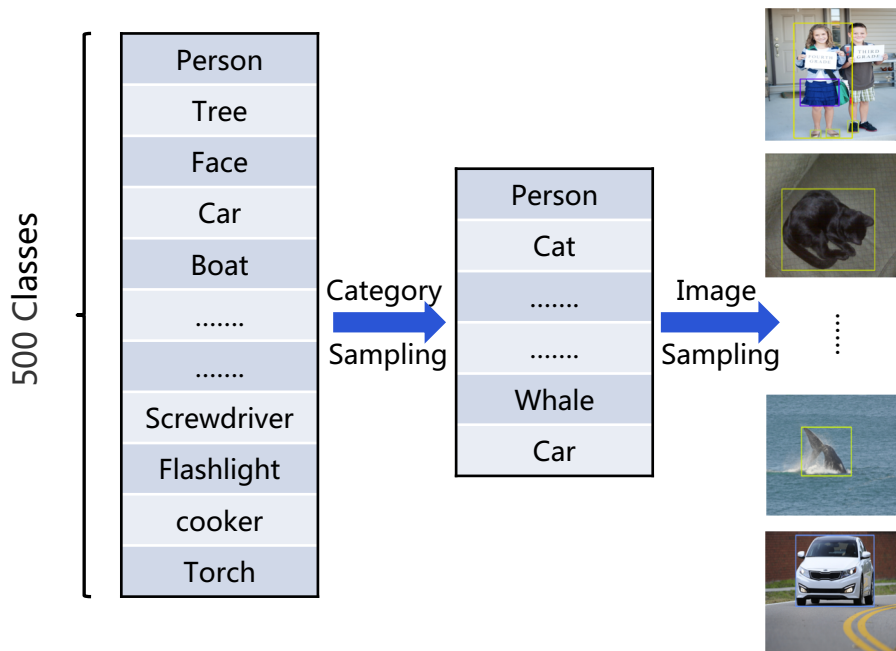
# Hierarchical NMS

Hierarchical Tag System

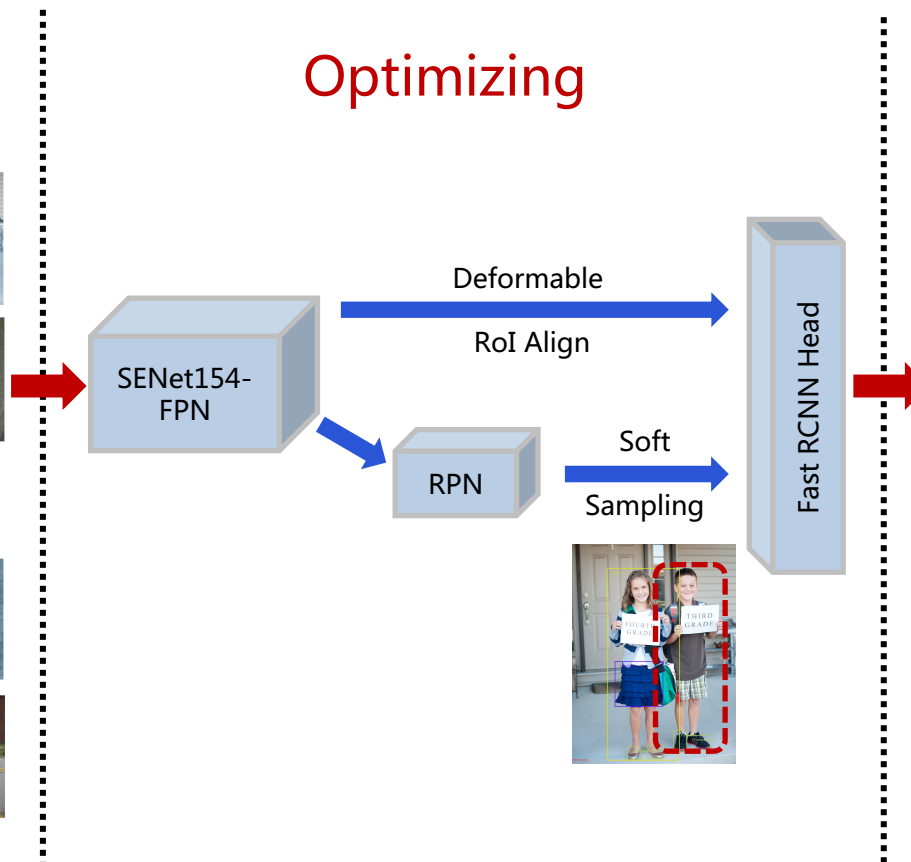


# Workflow

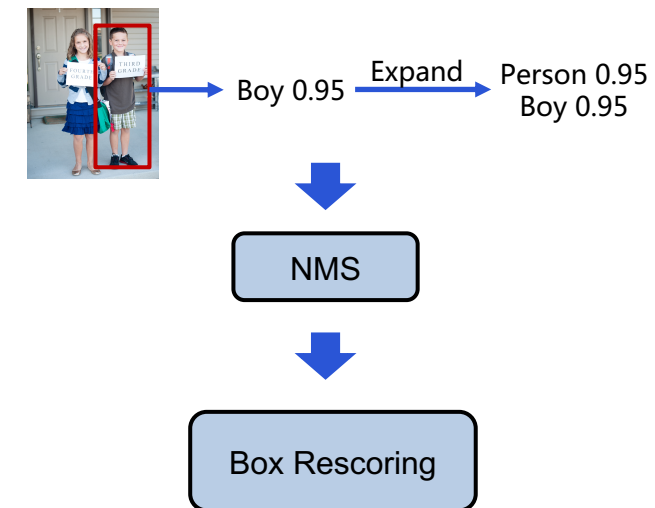
## Class Aware Sampling



## Optimizing



## Hierarchical NMS





# Ensemble

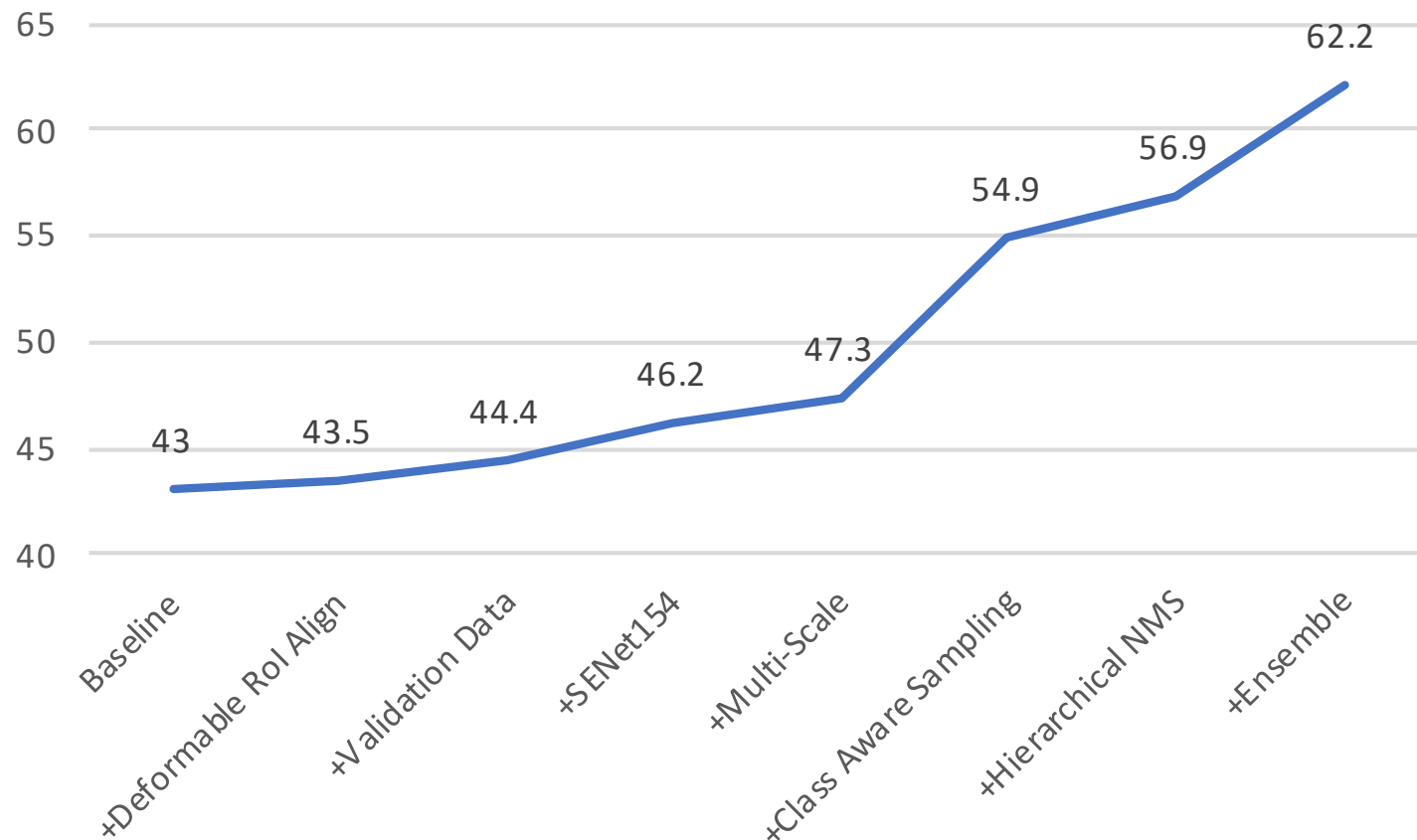
- Faster-RCNN FPN
  - 4 SENet154, 2 ResNeXt152, 3 ResNet101
- Diversity
  - w/ and w/o class aware sampling
  - w/ and w/o soft sampling
  - w/ and w/o SNIPER
  - Different pretrained model (ImageNet, COCO)

# Others

- Multi-scale testing
  - Shorter edge: 400, 500, 600, 700, 900, 1000, 1100, 1200
  - Longer edge: 1800
- Horizontal flip
- Performance drop - OHEM
- Batch size 48
- Learning rate 0.06 Momentum SGD
- 48 Tesla-V100 (16GB) GPUs in 2 days

# Result

## Open Images Score on Public Leader Board

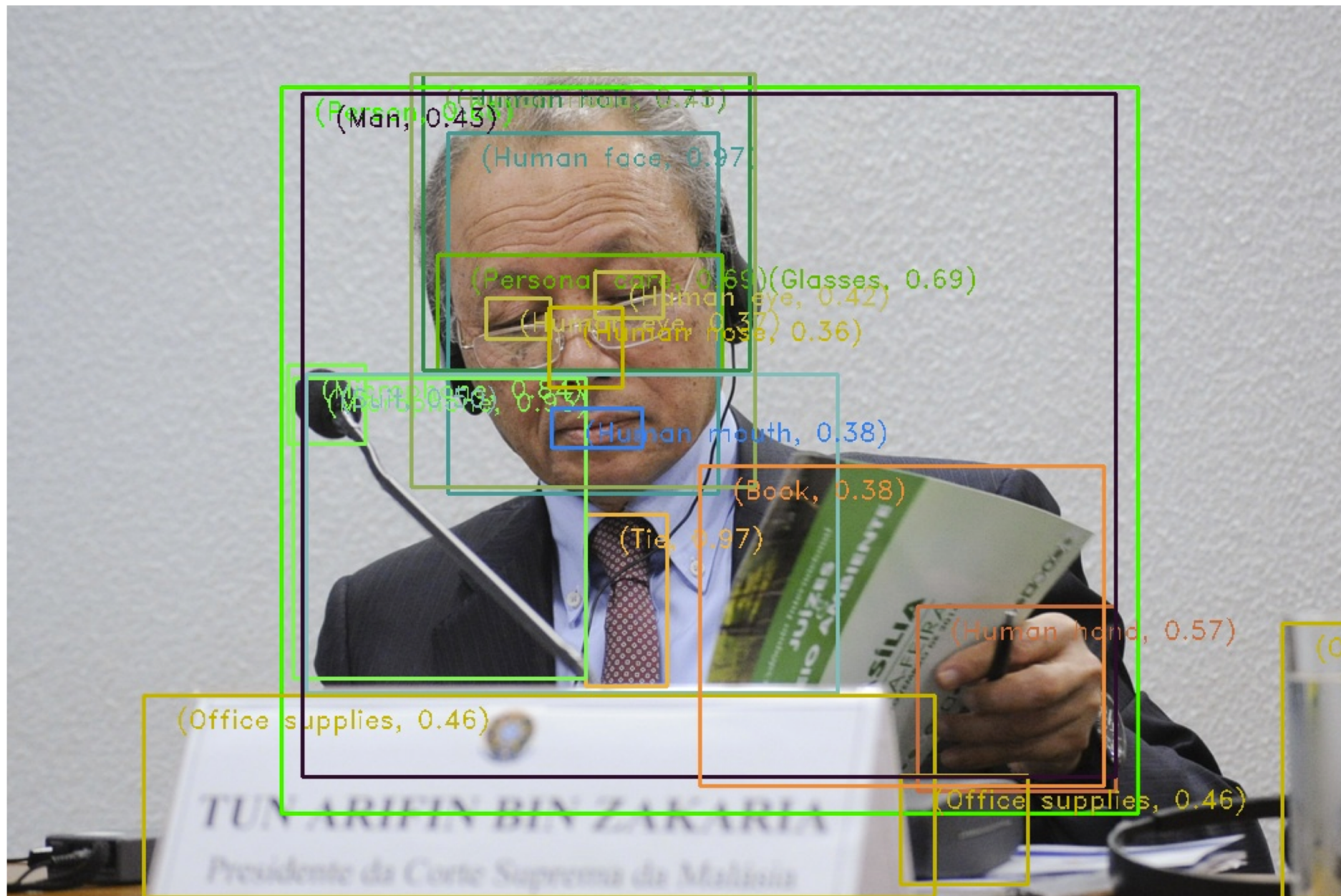


2<sup>nd</sup> on Public LB

3<sup>rd</sup> on Private LB  
(0.05% to the 1st)

Best Single Model on Public / Private LB: 56.9 / 53.5

# Visualization



# Visualization



Thank You !