Google AI

OPEN IMAGES CHALLENGE 2019

Open Images dataset

~9M images with:

- image-level labels,
- object bounding boxes,
- object segmentation masks,
- and visual relationships

(not all annotations types on all images)

Drawer
Drawer
Shelf
Shelf
Bookcase
Bookcase
Chest of drawers
Cabinetry
Filing cabinet
Shelf
Desk
Table
Office building
Desk
Desk
Chest of drawers

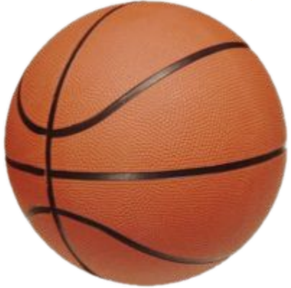
Why Open Images?

- It’s open!
- Composed only of images with CC_BY license
  → no copyright problems
- Can even use it commercially
- Enables legally safe crowdsourcing

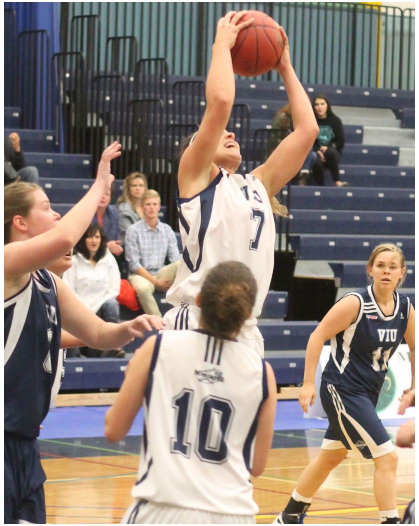
Which images for Open Images ?

- Start from Flickr, then dedupe against the web
  ⟶ biased away from simple images

- Collect all of CC on Flickr
  instead of scraped based on class keywords
  ⟶ natural class statistics, no initial design bias

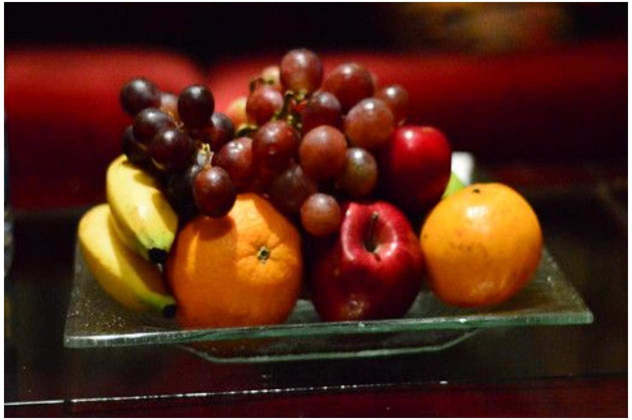


"Ball" in Google Image Search



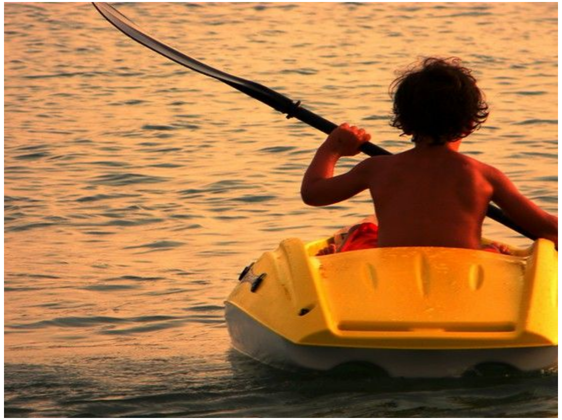
"Ball" in Open Images

Open Images V5: image-level labels



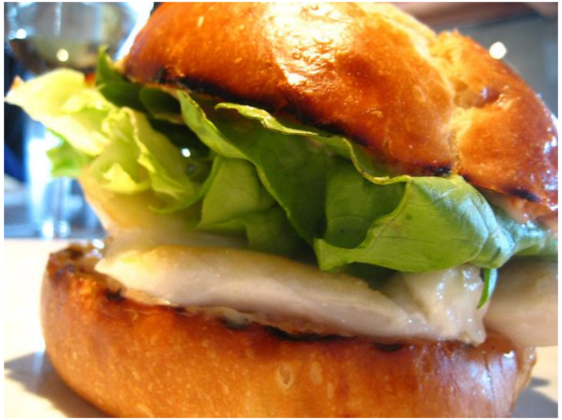
Positive image-level:
Apple, Food, Fruit,
Orange, Grapefruit

Negative image-level:
Plant, Tomato, Vegetable



Positive image-level:
Paddle, Boat, Person,
Watercraft, Vehicle

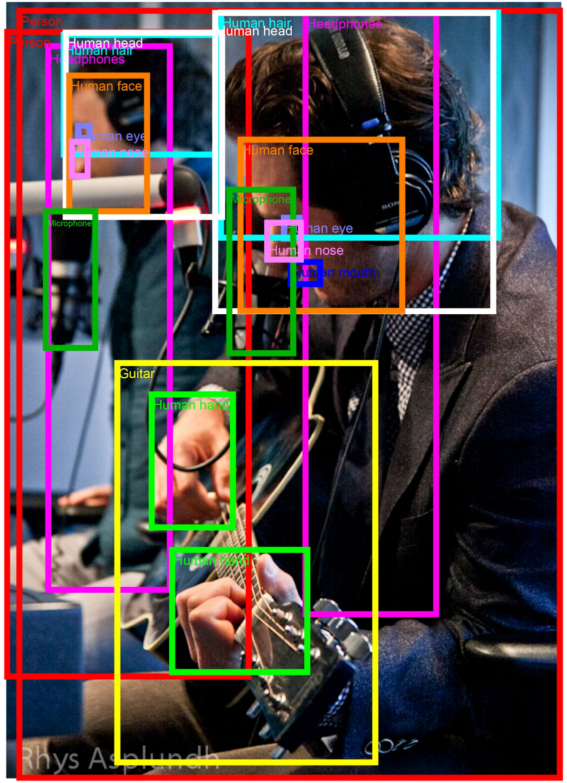
Negative image-level:
Jet ski, Man,
Personal flotation device

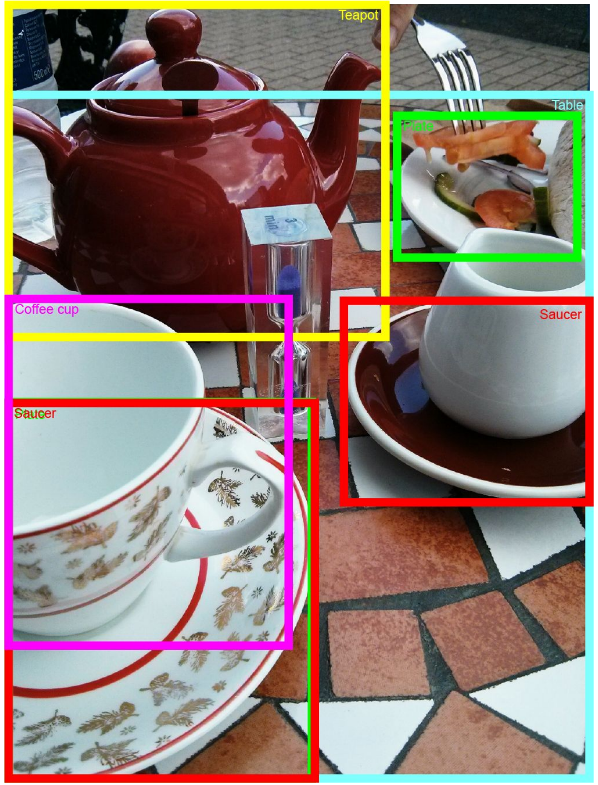


Positive image-level:
Fast food, Food, Baked goods,
Hamburger, Sandwich

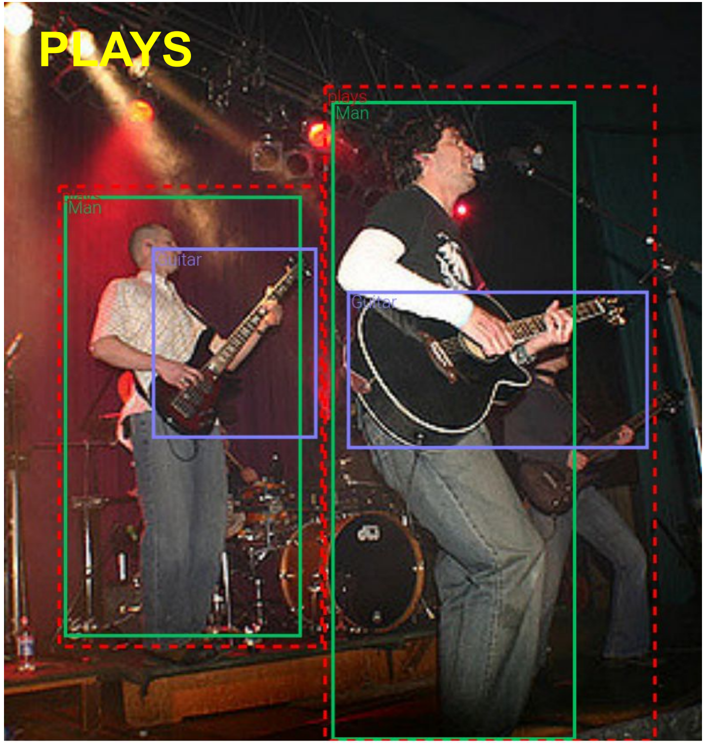
Negative image-level:
Human eye, Dessert,
Human mouth, Plant

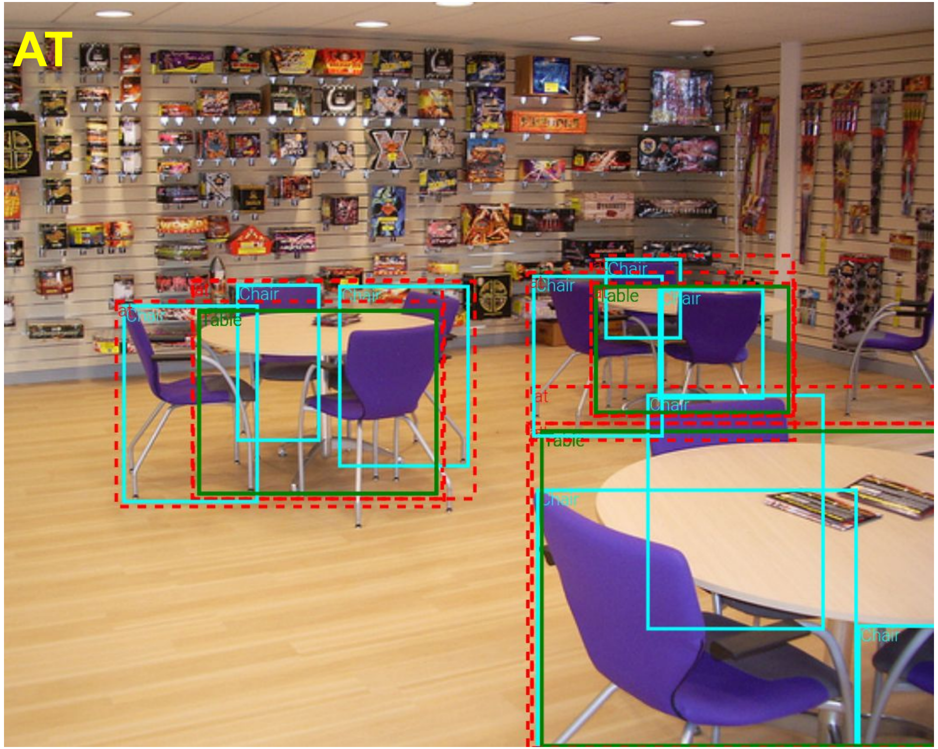
Open Images V5: object bounding-boxes


Person
Person
Human head
Human hair
Headphones
Human face
Human eye
Human nose
Microphone
Human hair
Human head
Headphones
Human face
Microphone
Human eye
Human nose
Human mouth
Guitar
Human hand
Human hand
Rhys Asplundh

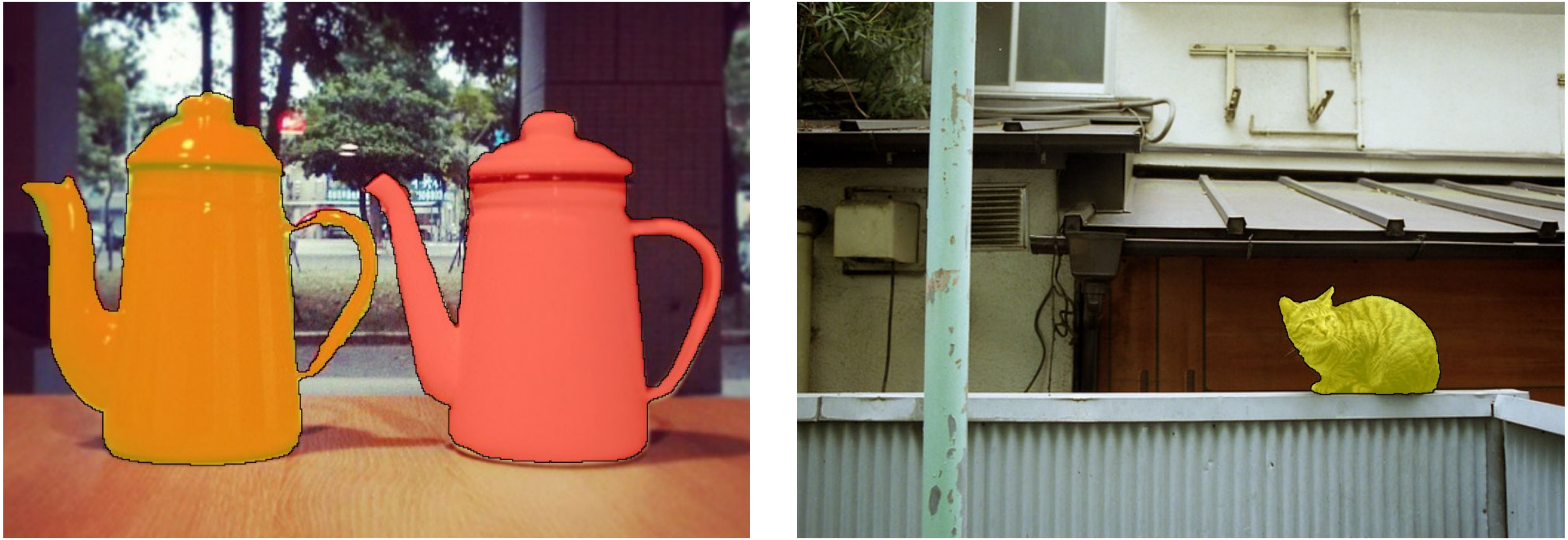

Teapot
Table
Plate
Coffee cup
Saucer
Saucer

Open Images V5: visual relationships

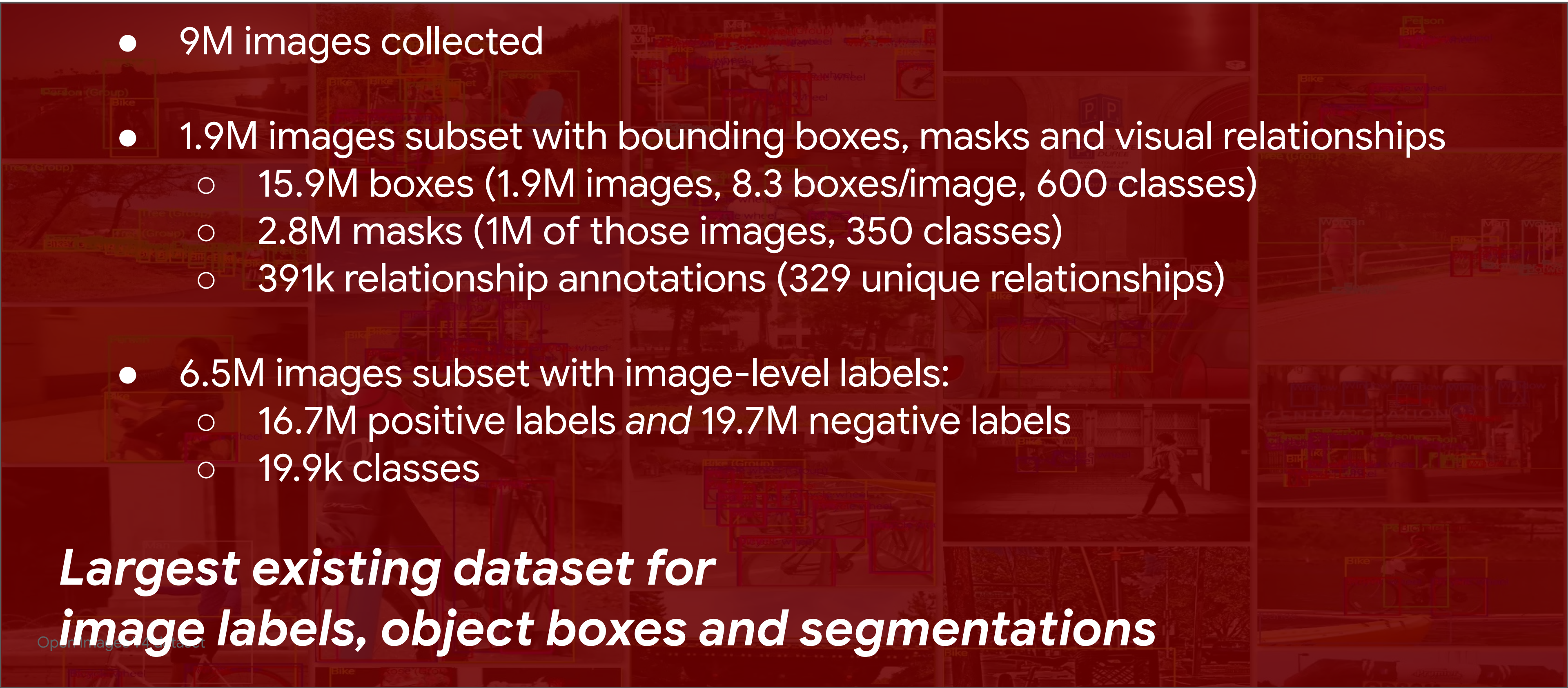

PLAYS
plays
Man
plays
Man
Guitar
Guitar


AT
Chair
Chair
Table
at
Chair
Chair
Chair
Table
Chair
at
Chair
Table
Chair
Chair

Open Images V5: instance segmentation masks (NEW)



Open Images V5: size (train+val+test)



- 9M images collected

- 1.9M images subset with bounding boxes, masks and visual relationships
  - 15.9M boxes (1.9M images, 8.3 boxes/image, 600 classes)
  - 2.8M masks (1M of those images, 350 classes)
  - 391k relationship annotations (329 unique relationships)

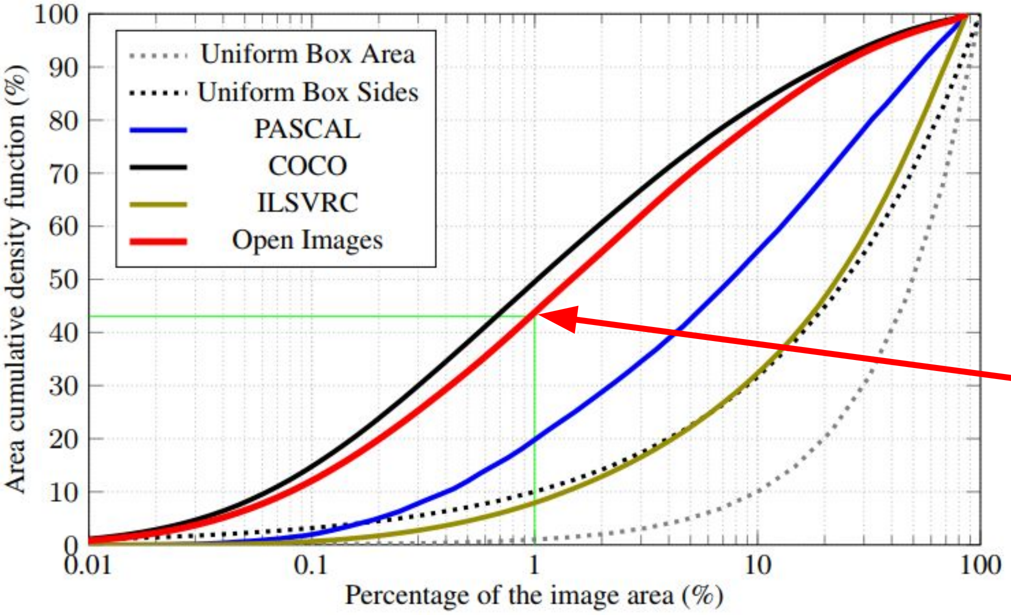
- 6.5M images subset with image-level labels:
  - 16.7M positive labels *and* 19.7M negative labels
  - 19.9k classes

Largest existing dataset for image labels, object boxes and segmentations

Open Images V4 dataset

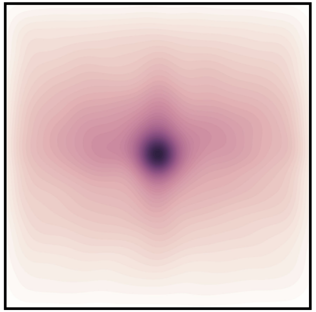
Open Images V5 complexity: objects number and size

- Complex images: 8.3 boxed objects on average (similar to COCO, better than PASCAL VOC and ILSVRC Detection)
- Object size distribution similar to COCO (many small objects)

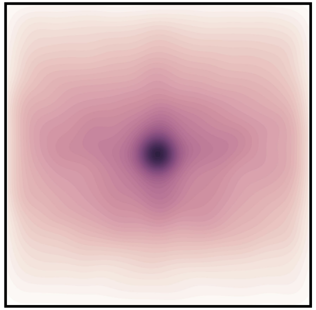


43% of the objects occupy less than 1% of the image area

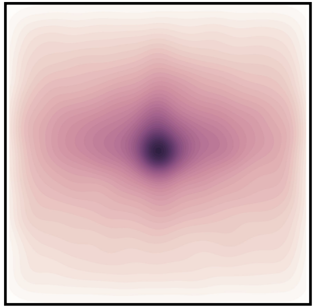
Open Images V5 complexity: objects center



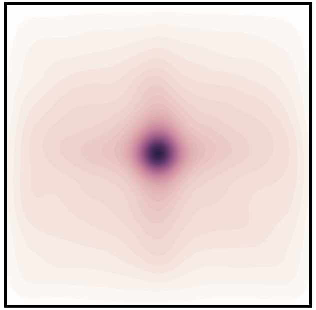
Open Images V4/V5
Train (**1.74M images**)



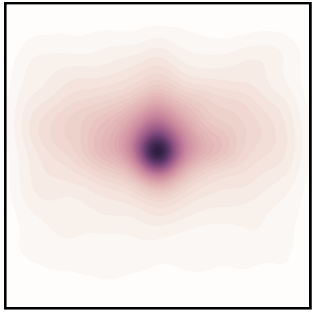
Open Images V4/V5
Challenge (0.1M images)



COCO 2017
Train + Validation
0.12M images



Open Images V5
Validation + Test
0.17M images



PASCAL 2012
TrainVal
0.01M images

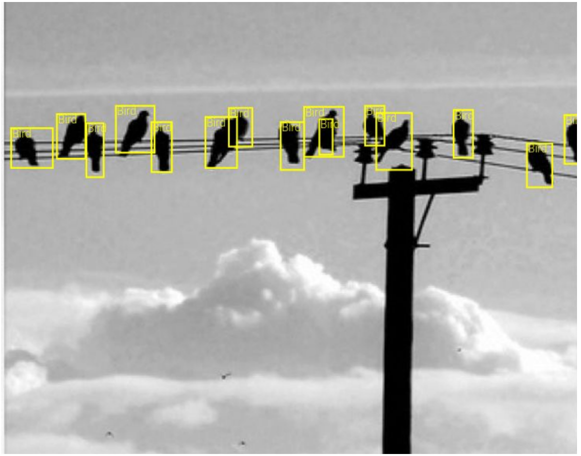
- Open Images train is bulk of data
- train/challenge are rich and diverse ≈ COCO
- val/test simpler, closer to PASCAL than COCO in this metric
  (but still many more objects/image than PASCAL: 7.3 vs 2.4)

Open Images V5: completeness

- Image labels partially annotated:
  positive, negative, unknown
  ⟶ use negative labels for training models!
  (they are hard examples!)

- Boxes: all instances for the positive labels
  (we annotate the most specific label)
  ⟶ boxed *all* bird in the image
  ⟶ enables reliable performance evaluation

- Segmentations:
  ⟶ all big-enough boxes
  (>40x80 pixels, capped 4 classes)
  (evaluation uses boxes to ignore small instances)



Positive: bird
Negative: streetlight
Unknown: all others
(mind hierarchy effects)

Open Images Challenge 2019

Data

- Challenge test set: evaluation *on separate image set* with hidden annotations
- Training set: based on Open Images V5, reduced to classes covered in Challenge

- Most reliable performance indicator is public leaderboard on Kaggle: images+annotations used for evaluation on public and private leaderboard are identically distributed
- Open Images V5 val set can also be used for setting hyper-parameters as indicative of performance on Challenge set, but not identically distributed

Three tracks:

- Object Detection
- Visual Relationship Detection
- Instance Segmentation (NEW)

75K USD prize money fund

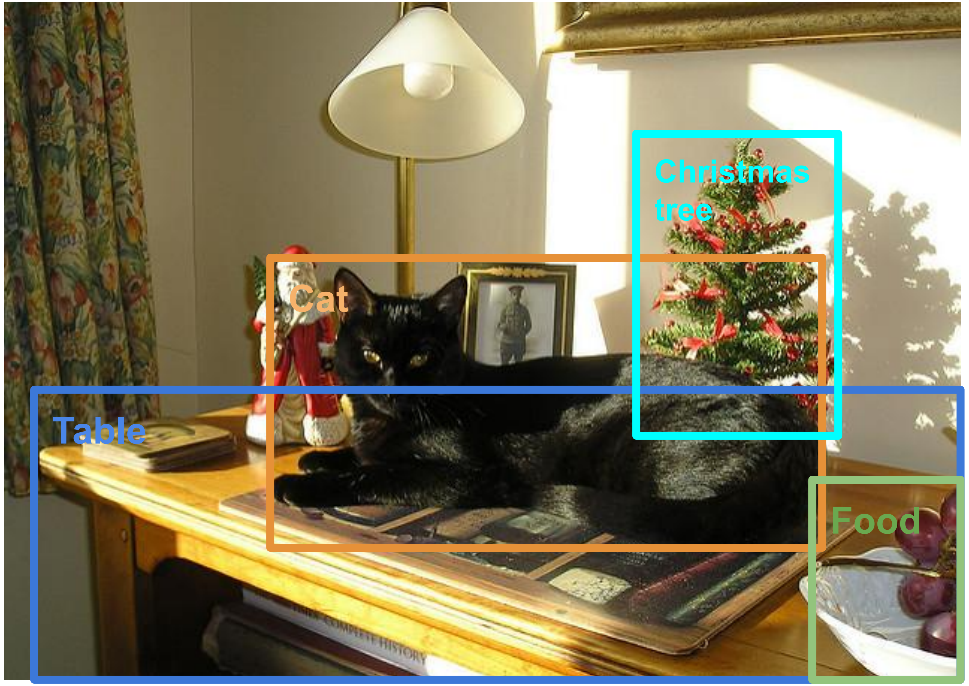
Track 1: Object Detection

Task: detect objects on an image

- object bounding box
- class label

Training set:

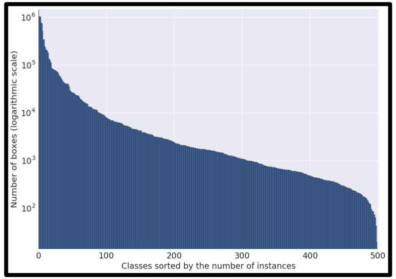
- 12,2M bounding boxes
- 1,7M images
- 500 classes
  removed some broad (e.g. "clothing") and some infrequent classes (e.g. "paper cutter")

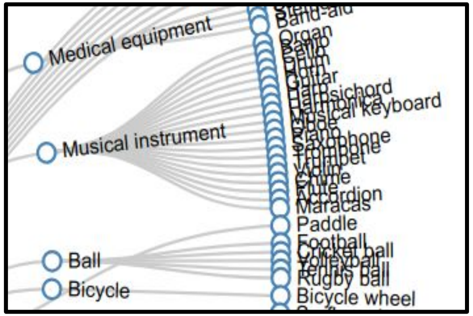

Christmas tree
Cat
Table
Food

Track 1: Object Detection - why challenging

- Complex images
- Many classes (500)
- Imbalanced class distribution
- Class hierarchy
- Non-exhaustive annotations for image labels (but all objects annotated for those we have)

FUSION
TOURISTIQUE
TRANSPORT




Medical equipment
Musical instrument
Ball
Bicycle
Band-aid
Organ
Guitar
Harpsichord
Musical keyboard
Saxophone
Trumpet
Accordion
Maracas
Paddle
Football
Cricket ball
Rugby ball
Bicycle wheel

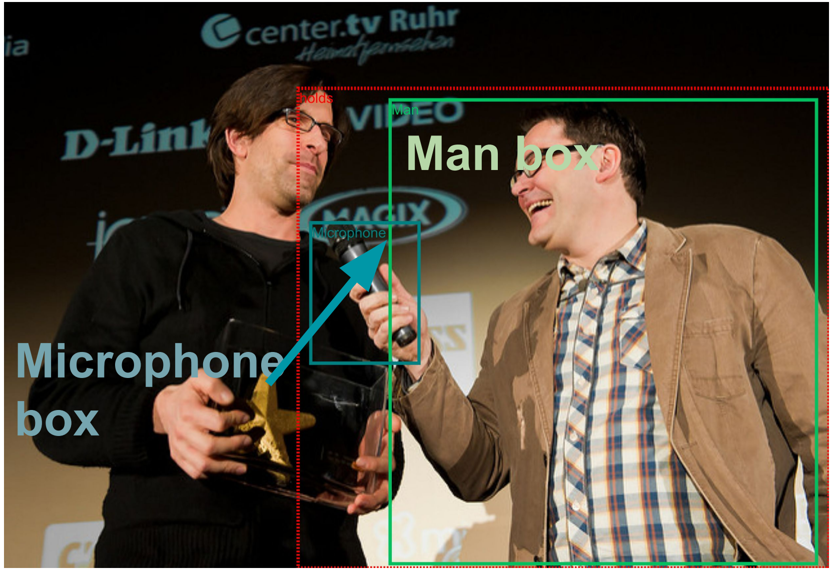
Track 2: Visual Relationships Detection

Task: detect relations and attributes between objects

- Two objects boxes and classes
- Relationship between them
- Also, object attributes in the form of “is” relationship

Training set:

- 375k relationship annotations
- 3,2M boxes
- 1,7M images
- 329 distinct relation triplets
- 62 different object classes

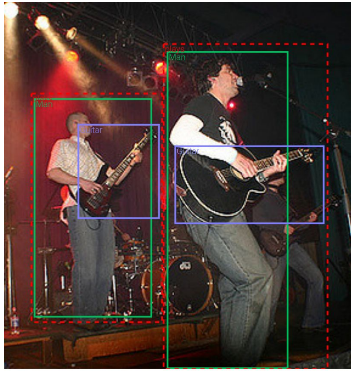

holds
Man
Man box
Microphone
Microphone box

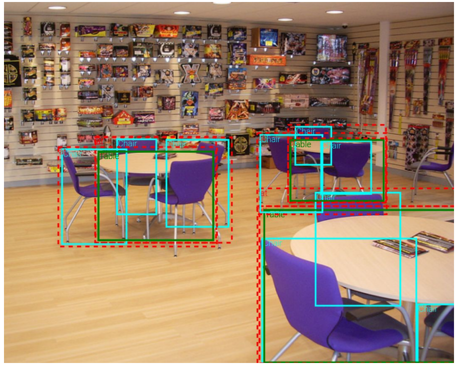
Must identify which of the two men holds the microphone!

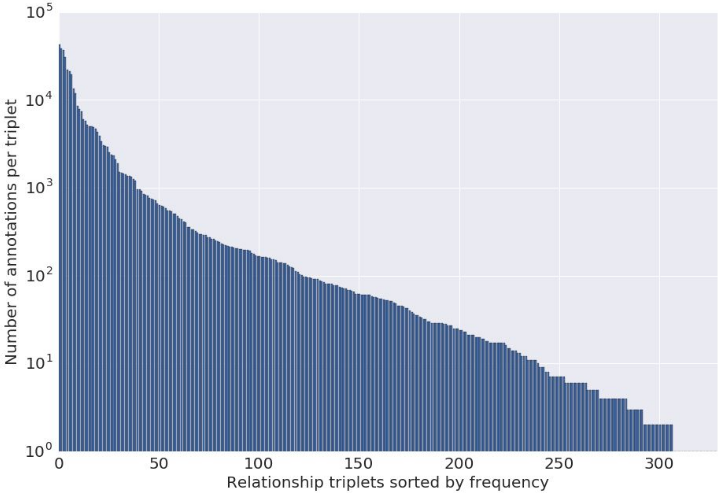
Track 2: Visual Relationships Detection - **why challenging**

Challenges:

- Complex images
- Class co-occurrence not sufficient to predict the relationship
- High relationship triplets imbalance


Man
Guitar
Man
Guitar


Chair
Table
Chair
Table



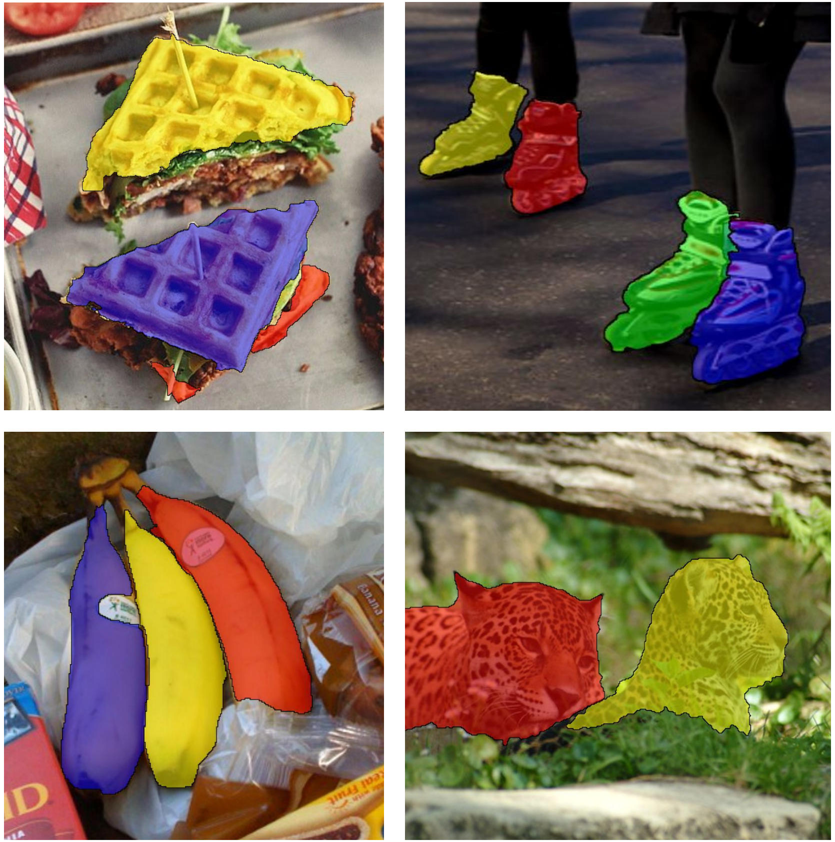
Track 3: Instance segmentation

Task: segment objects on an image

- object mask
- class label

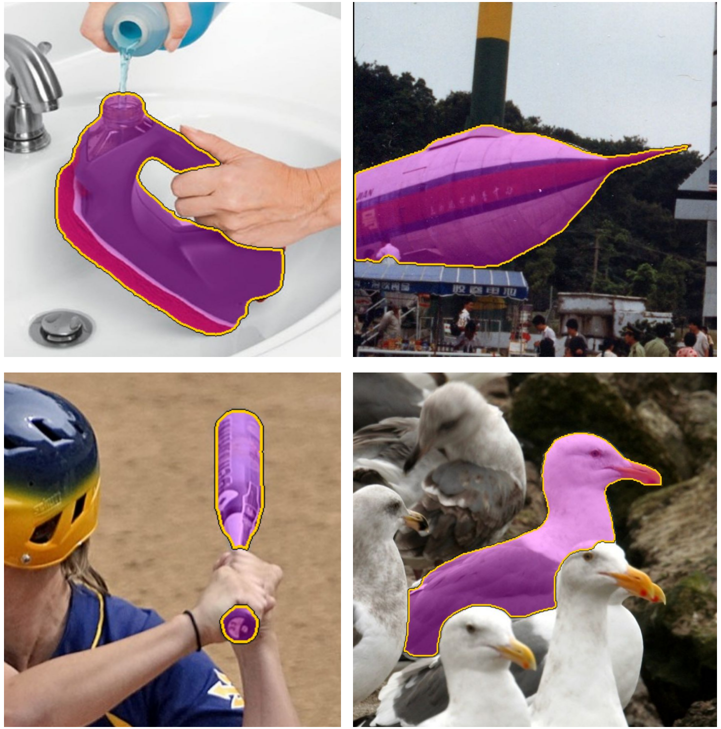
Training set:

- 2,1M masks
- 848k images
- 300 classes



Track 3: Instance segmentation - why challenging

- Complex images
- Many classes (300)
- Imbalanced class distribution
- Diverse family of shapes across classes
- Annotations include topologically complex objects



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

Challenge organizers

Vittorio
Ferrari

Alina
Kuznetsova

Rodrigo
Benenson

Victor
Gomes

Matteo
Malloci

Many contributors to the Open Images Dataset (25+ without counting the annotators armada)

Project Lead & Coordination

Vittorio Ferrari
Tom Duerig
Victor Gomes

Image collection

Ivan Krasin
David Cai

Image-level labels

Neil Alldrin
Ivan Krasin
Shahab Kamali
Tom Duerig
Zheyun Feng
Anurag Batra
Alok Gunjan

Bounding boxes

Hassan Rom
Alina Kuznetsova
Jasper Uijlings
Stefan Popov
Matteo Malloci
Sami Abu-El-Haija
Vittorio Ferrari

Segmentations

Rodrigo Benenson
Stefan Popov
Matteo Malloci
Vittorio Ferrari

Visual Relationships

Alina Kuznetsova
Matteo Malloci
Vittorio Ferrari

Website & visualizer

Jordi Pont-Tuset

Classes & hierarchy

Chen Sun
Kevin Murphy
Tom Duerig
Vittorio Ferrari

Challenge

Vittorio Ferrari
Alina Kuznetsova
Rodrigo Benenson
Victor Gomes
Matteo Malloci
Jordi Pont-Tuset
Jasper Uijlings
Jake Walker

Advisers

Andreas Veit
Serge Belongie
Abhinav Gupta
Dhyanesh Narayanan
Gal Chechik

Thanks to ...

Common Visual Data Foundation
For hosting the data through AWS
Special thanks to Tsung-Yi Lin

Figure Eight
For hosting the data

Kaggle
For hosting the competition

Julia Elliott

Thanks to our partners

Thanks to our price sponsors

Google AI

kaggle

Thank you to our sponsors

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

Room: 402

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

[**Prisms**, **imagesearch**, **MMfruit**]

14:50 - 15:05 Instance segmentation track

Settings, metrics, winners, analysis

15:05 - 15:50 Presentations by three winners of the Instance segmentation track

[**PFDet**, **n01z3**, **MMfruitSeg**]

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

[**Very Random team**, **Layer6 AI**]

17:25 - 17:30 Concluding remarks