



# Microsoft COCO

Common Objects in Context



**Tsung-Yi Lin**  
Cornell Tech



**Michael Maire**  
TTI Chicago



**Serge Belongie**  
Cornell Tech



**Lubomir Bourdev**



**James Hays**  
Brown University



**Pietro Perona**  
Caltech



**Deva Ramanan**  
UC Irvine



**Ross Girshick**  
Microsoft Research



**Piotr Dollar**  
Microsoft Research



**Larry Zitnick**  
Microsoft Research



Microsoft COCO  
Common Objects in Context

# Why a new dataset?

Continue our field's momentum:

IMAGENET



**Caltech Pedestrian Detection Benchmark**



<http://mscoco.org>

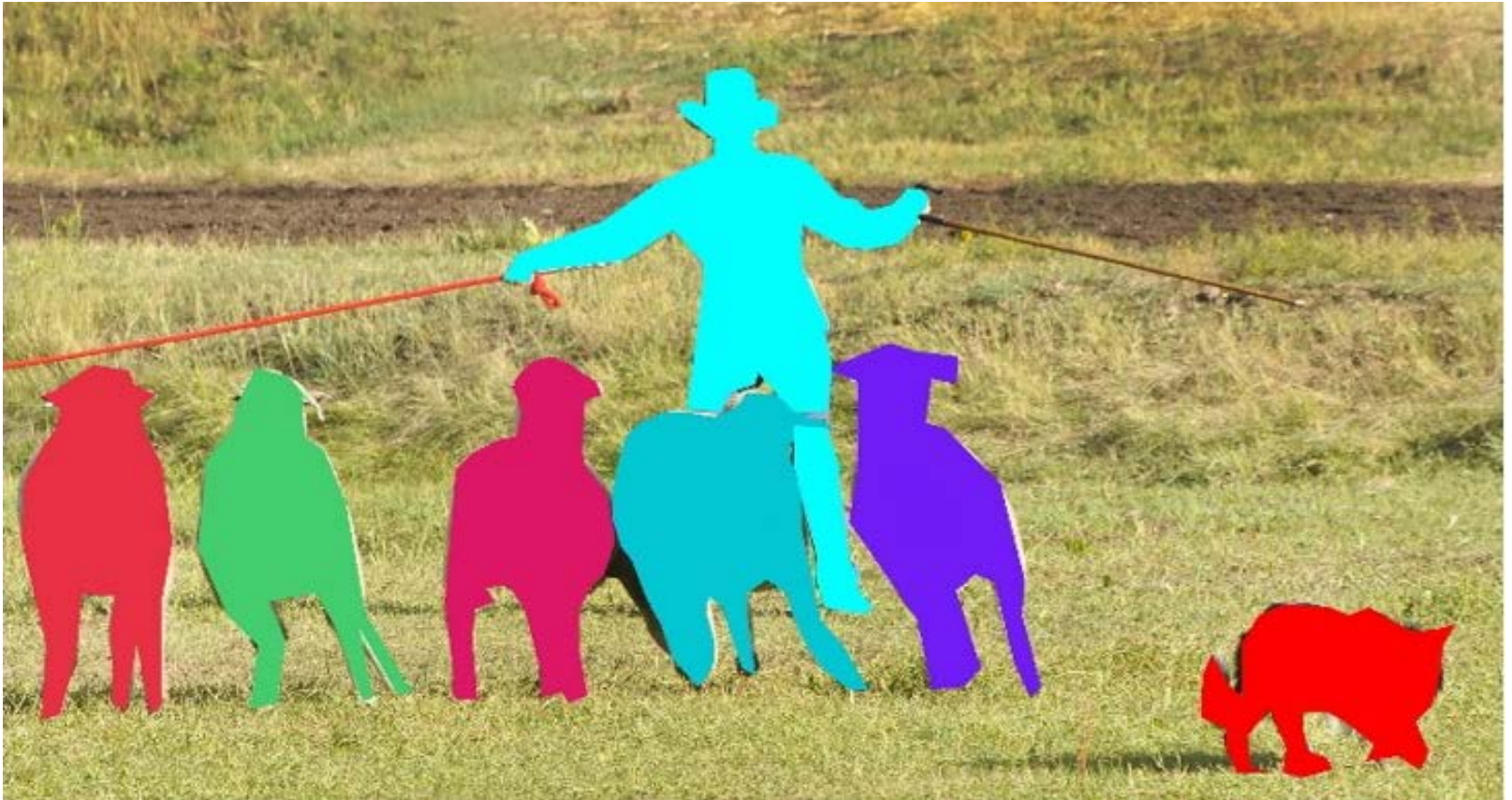








- ✓ Instance segmentation
- ✓ Non-iconic Images



# Iconic object images





# Iconic scene images





# Non-iconic images

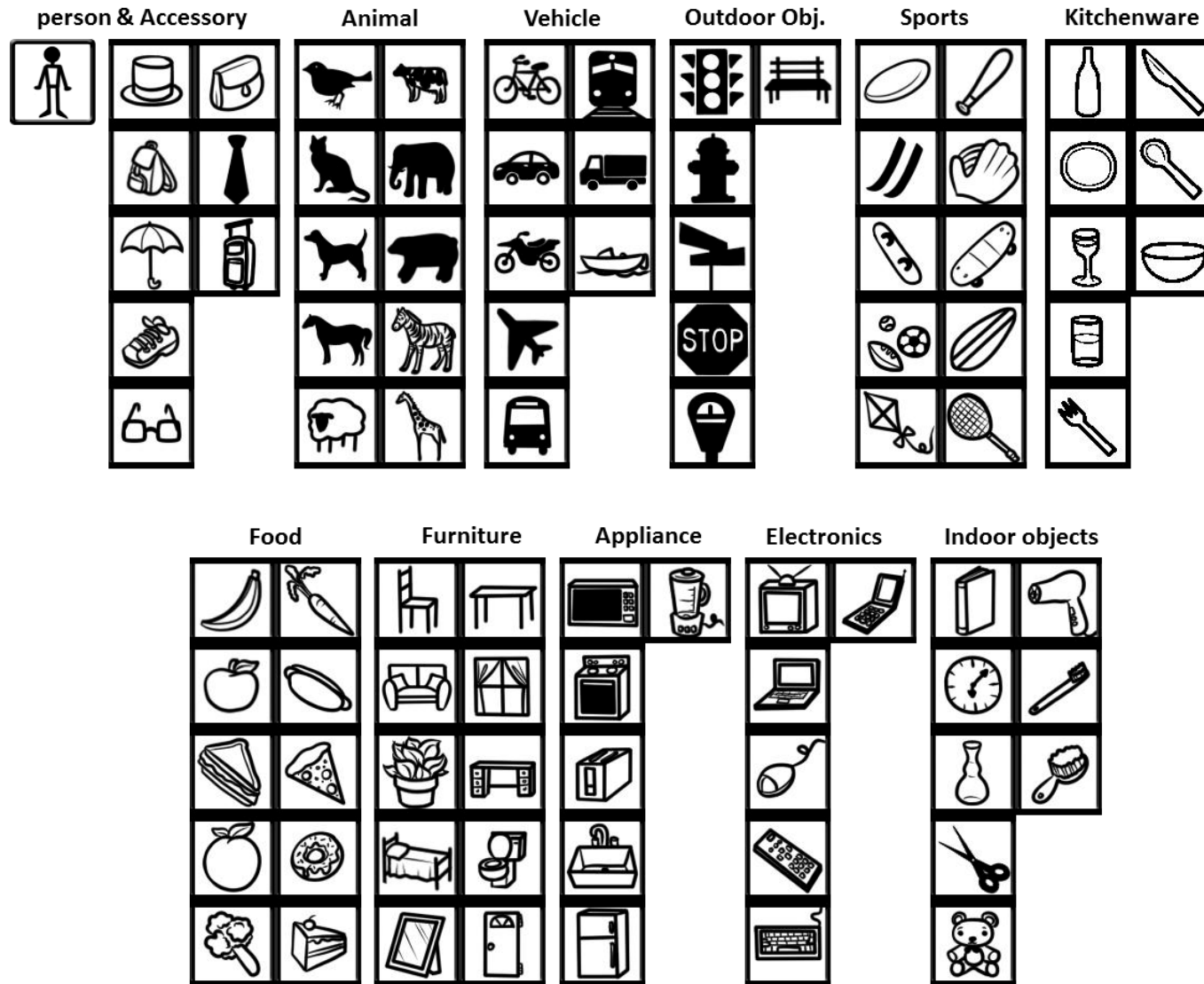








# Object categories (things not stuff)



# flickr

**(All creative commons)**

330,000 images

“Dog”





# “Dog + Car”



*Im2Text: Describing Images Using 1 Million Captioned Photographs, V. Ordonez, G. Kulkarni, T. L. Berg NIPS'11*



Microsoft COCO  
Common Objects in Context

# Annotation pipeline







# Divide and Conquer

## 1. Category Labeling



dog, bottle

## 2. Instance spotting



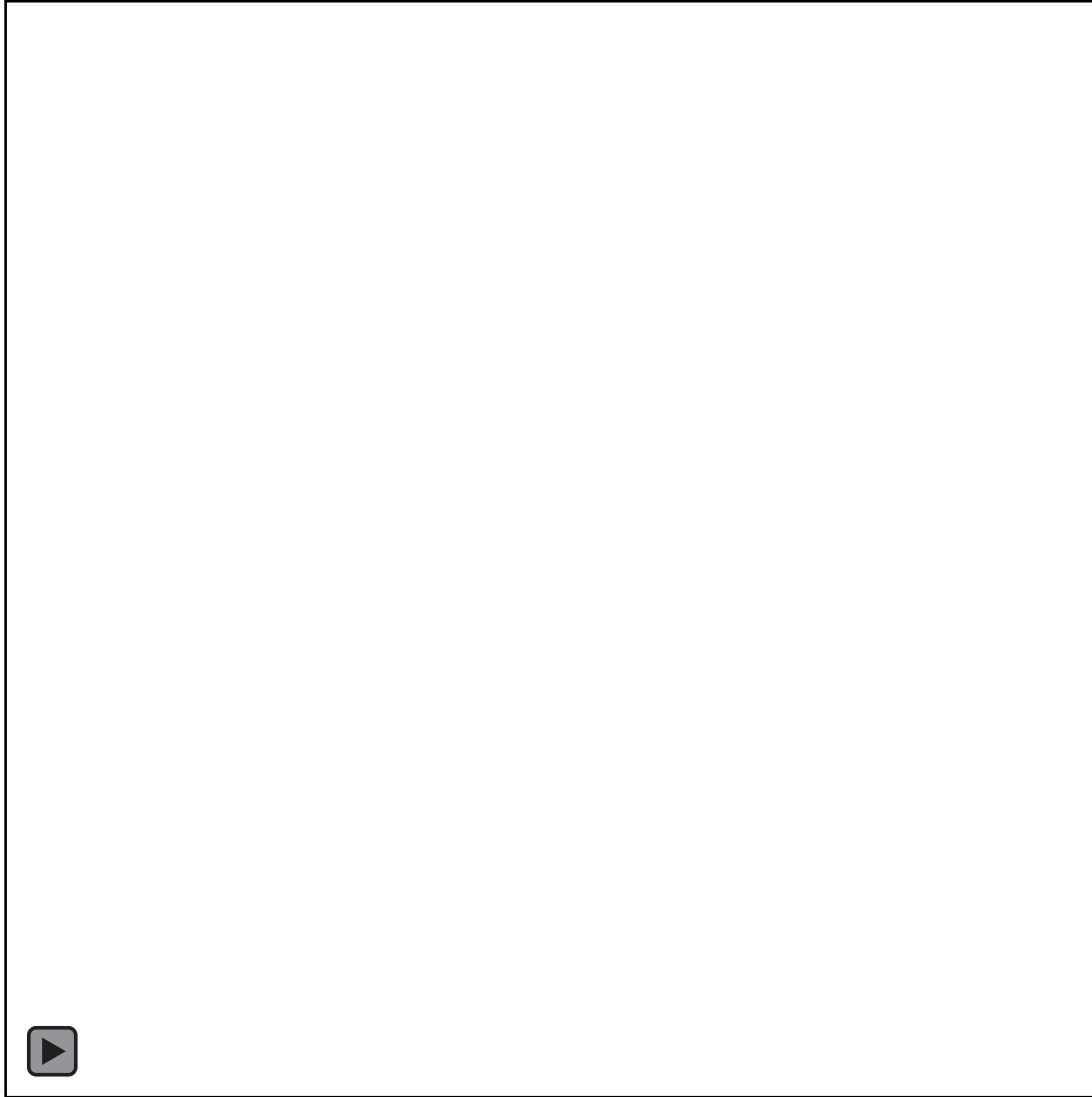
## 3. Instance segmentation



# 1. Category Labeling



## 2. Instance Spotting





# 3. Instance Segmentation



edu/

ca Bala,



FLIP

TASER



1

FLIP

TAREE



After training





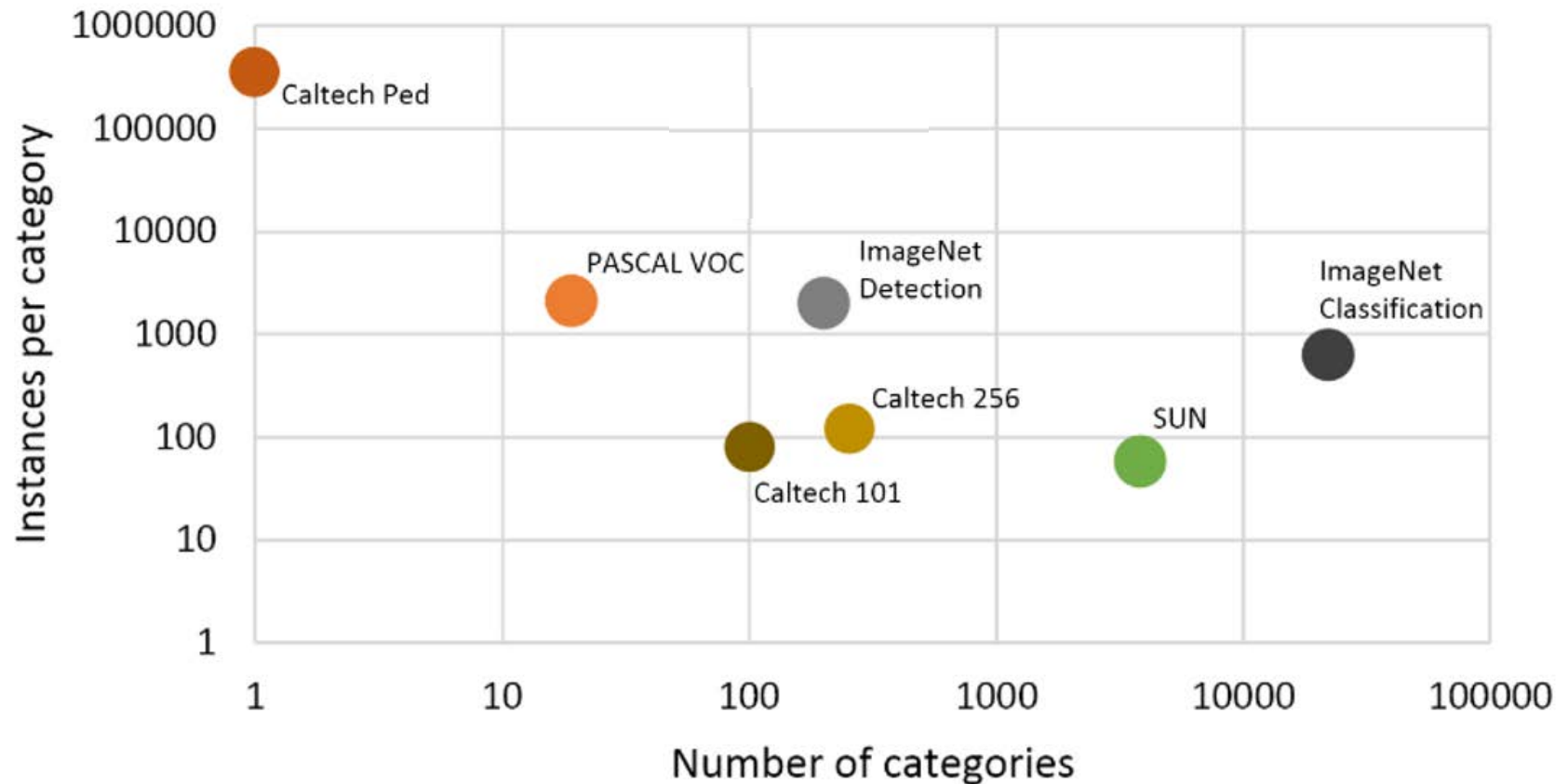


Microsoft COCO  
Common Objects in Context

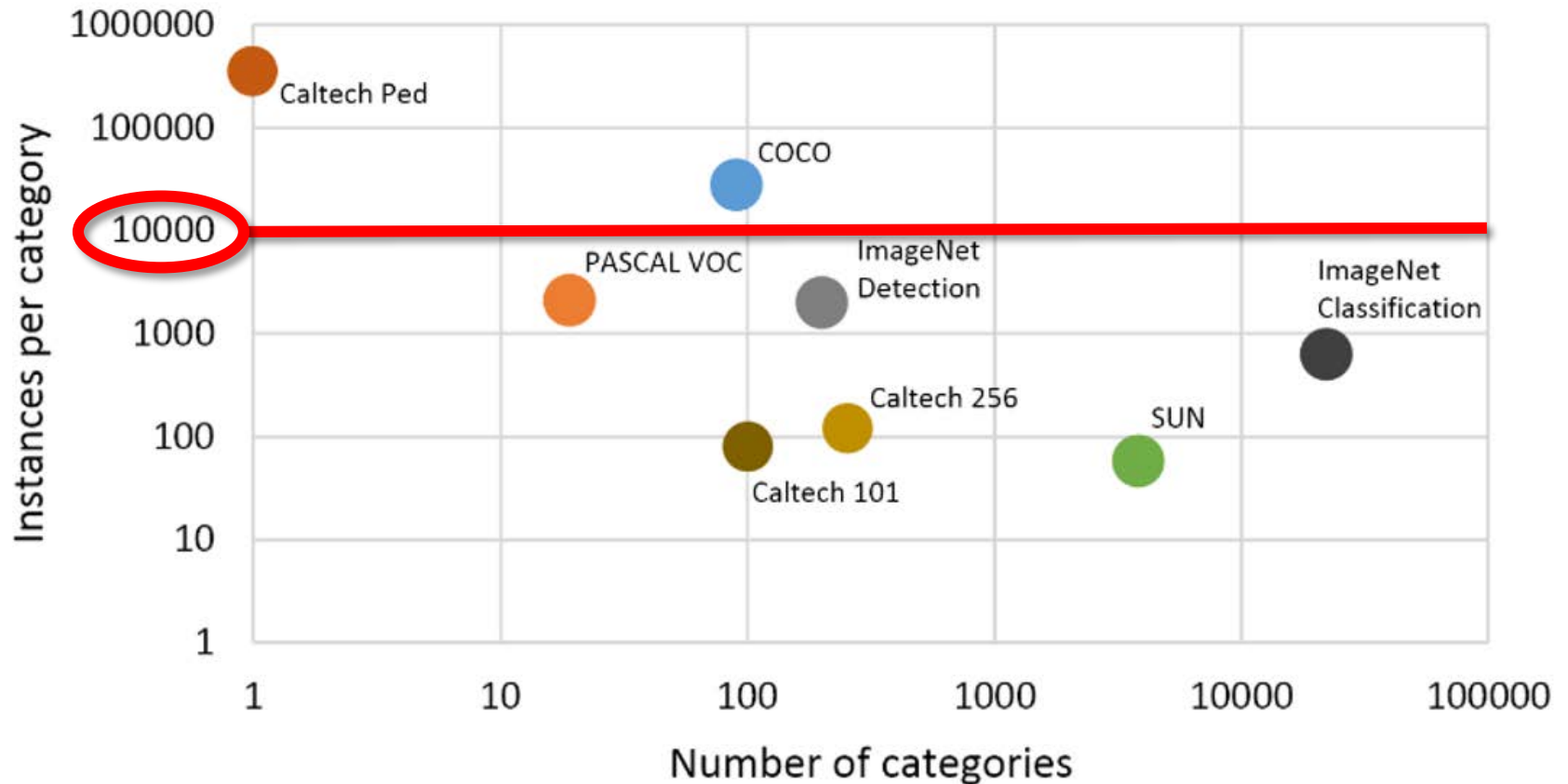
# Properties



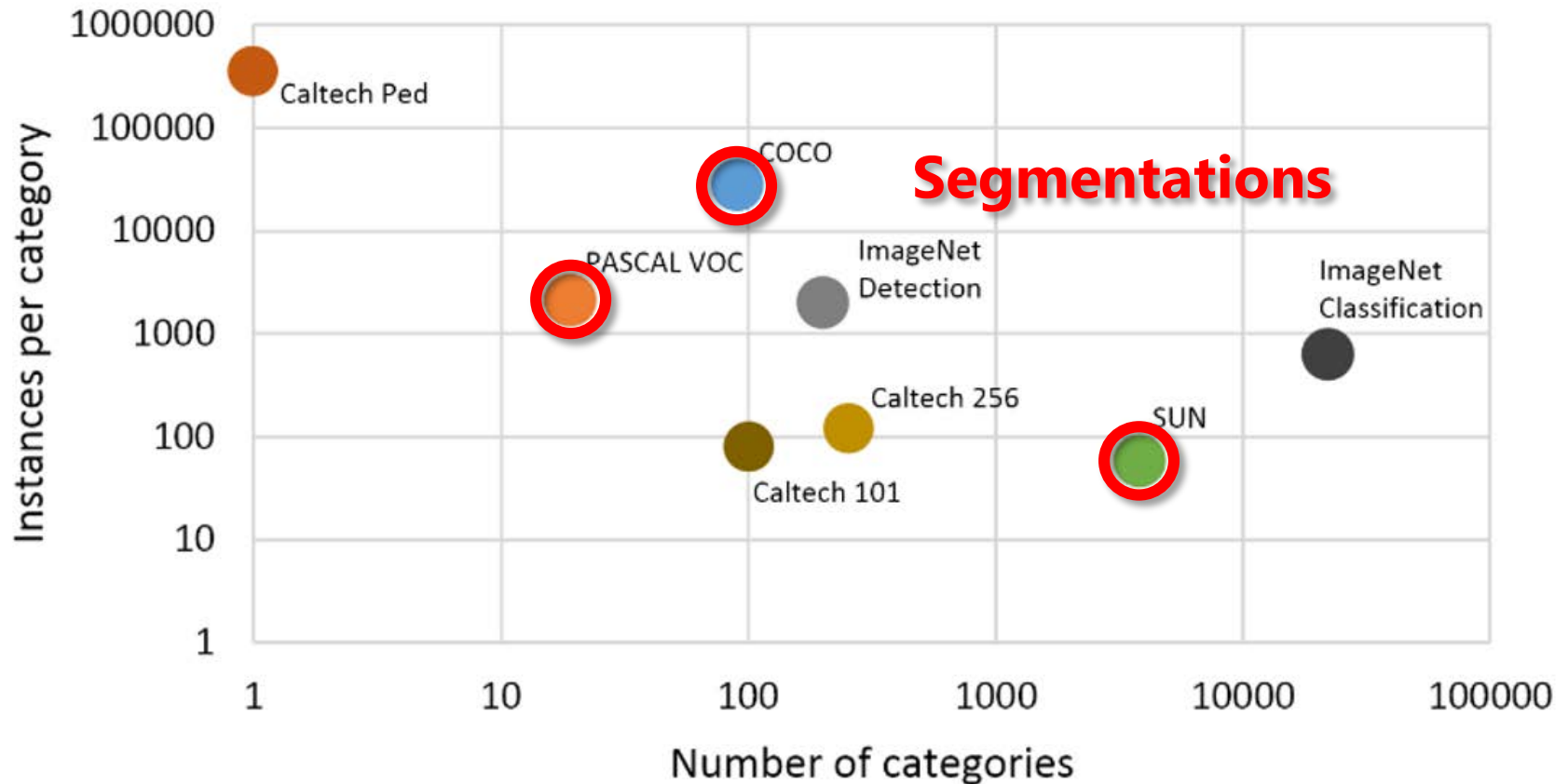
## Number of categories vs. number of instances



# Number of categories vs. number of instances



# Number of categories vs. number of instances





- 330,000 images
- >2 million instances (700k people)
- Every instance is segmented
- 7.7 instances per image (3.5 categories)

Person



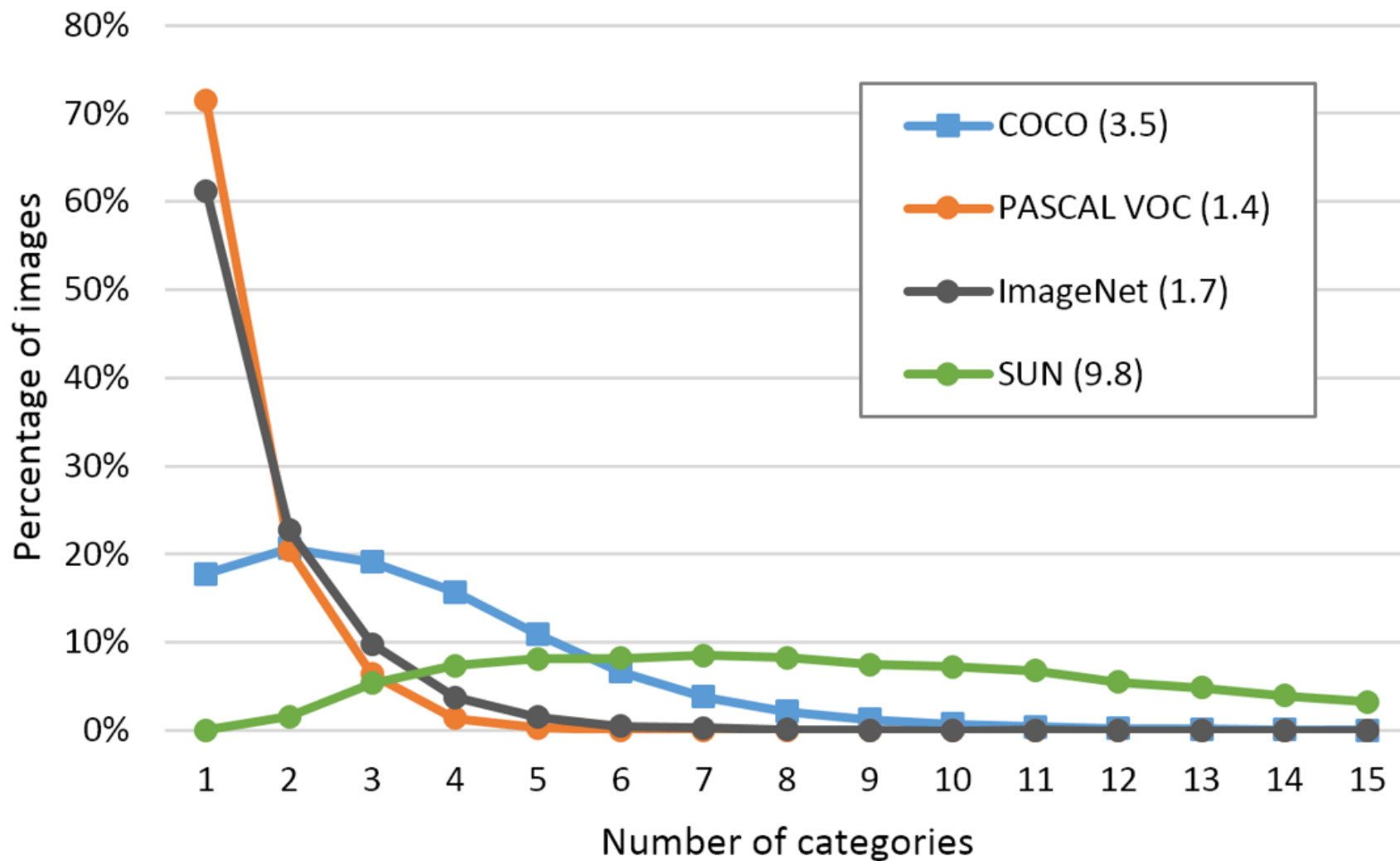
Dog



Cow



# Categories per image



# Detection Performance

( DPM V5 )

	Person (mAP)	Average (mAP)
PASCAL VOC	41.3	29.6
MS COCO	17.5	16.9



# http://mscoco.org



 Microsoft COCO  
Common Objects in Context

cocodataset@outlook.com

Home People **Explore** Download

Disclaimer: Annotations are currently in progress (final annotations will be available soon).



person× elephant× search

899 results

URL    



http://mscoco.org



Microsoft COCO  
Common Objects in Context

# Beyond detection

# Beyond detection

## ✓ Sentences

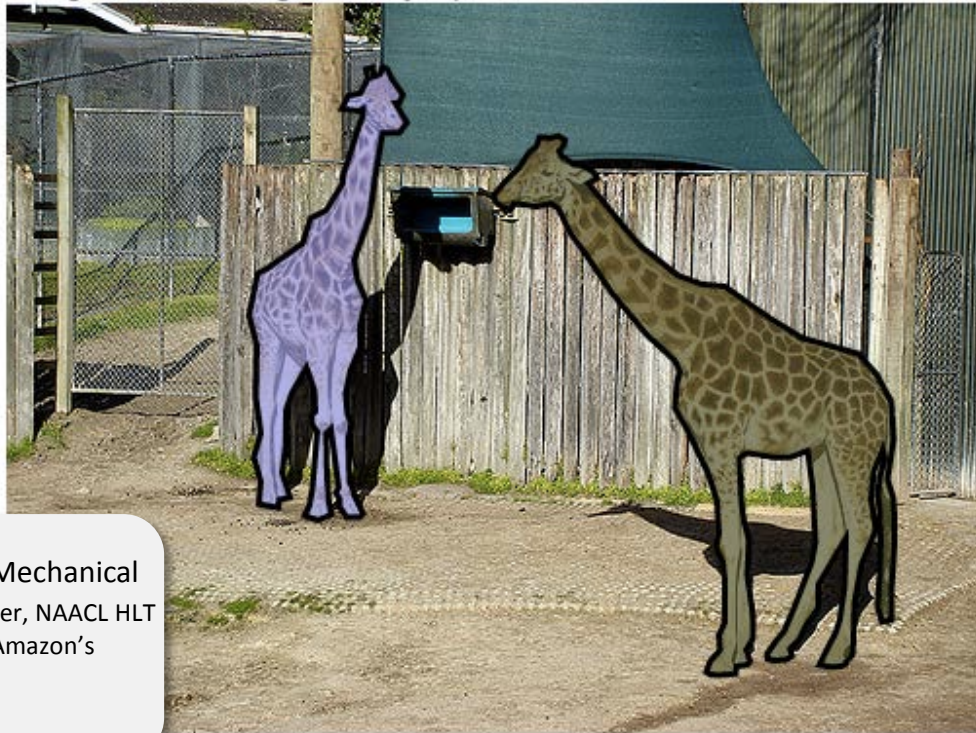
two giraffe standing next to each other in front of a wooden fence.

two giraffes standing in the dirt near a gate.

two giraffes stand by a food box awaiting the goods.

two giraffes are standing next to a wooden fence.

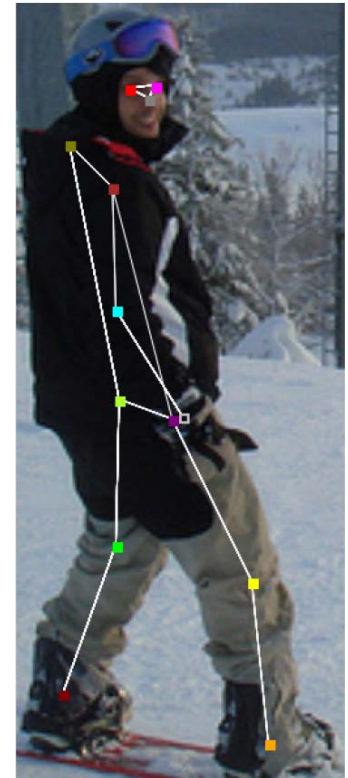
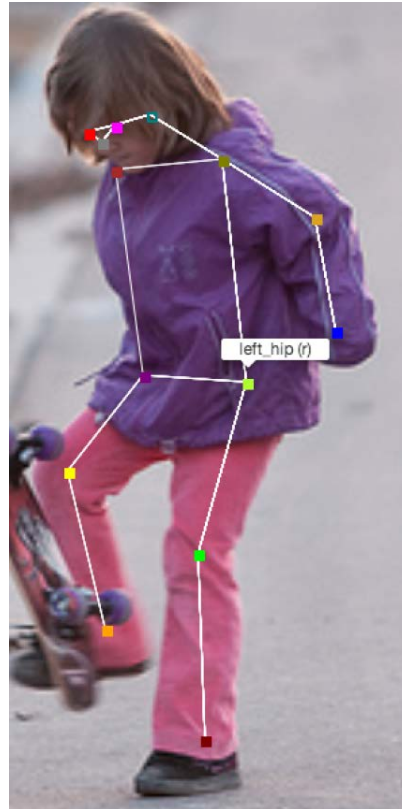
two giraffes standing alone by a picket fence.



Collecting Image Annotations Using Amazon's Mechanical Turk, C. Rashtchian, P. Young, M. Hodosh, J. Hockenmaier, NAACL HLT Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, 2010

# Beyond detection

- ✓ Keypoints  
(provided by Facebook)





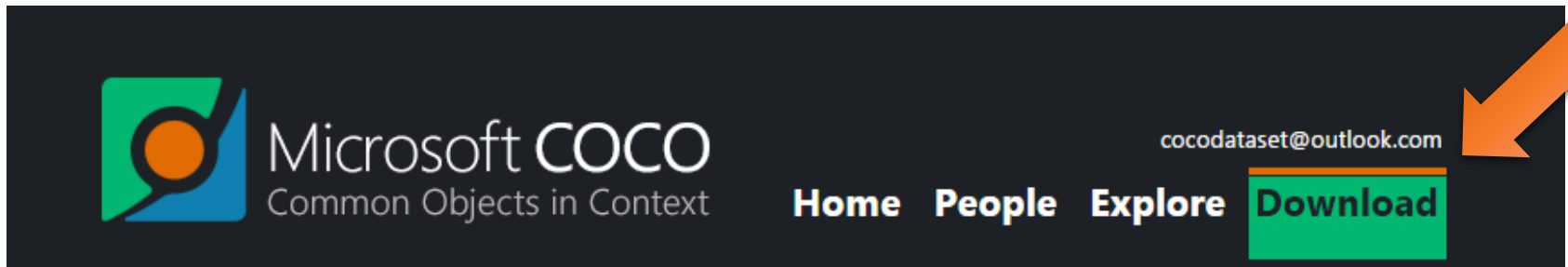


Microsoft COCO  
Common Objects in Context

# How to use COCO

<http://mscoco.org>

# http://mscoco.org



## Tools

[Python tools](#)

Matlab tool is coming soon!

## Images

[Training images \[13GB\] \(v.2014\)](#)

[Validation images \[6.2GB\] \(v.2014\)](#)

The current release contains 82,783 training images and 40,504 validation images. The test set will be released later this year. Note that in 2015 we will be releasing the second half of the dataset which will contain a total of 328,000 images.

## Annotations

[Training instance annotations \[324MB\] \(v.2014 0.9\)](#)

[Validation instance annotations \[157MB\] \(v.2014 0.9\)](#)

[Training sentence annotations \[61MB\] \(v.2014 0.9\)](#)

[Validation sentence annotations \[30MB\] \(v.2014 0.9\)](#)

The annotations are currently on version 0.9 and subject to change.

http://mscoco.org

# APIs

- Python and MATLAB

## `getInfo`

Print the information of the loaded annotation file.

## `getImageIds`

Get a list of image ids from annotation objects that satisfied the specified constraint(s). The constraint can be category id for instance annotations.

## `loadAnnotations`

Get a list of annotation objects that satisfies the specified constraint(s). The constraints can be one or a combination of followings: list of image ids, a category id, and the size of instance. The function overloads for different types of annotations.

## `showAnns`

Take a list of annotation objects and visualize them. The function overloads for different types of annotations.

## `loadImage`

Load a color image by its id.

# MS COCO 2014 release

(half of COCO)

Over 77,000 worker hours (8+ years)

- 160k images
- 80 object categories (things not stuff)
- 1M+ instances (300k people)
- Every instance segmented
- 5 sentences per image
- Separate train and validation set



<http://mscoco.org>





Microsoft COCO  
Common Objects in Context

# Going forward

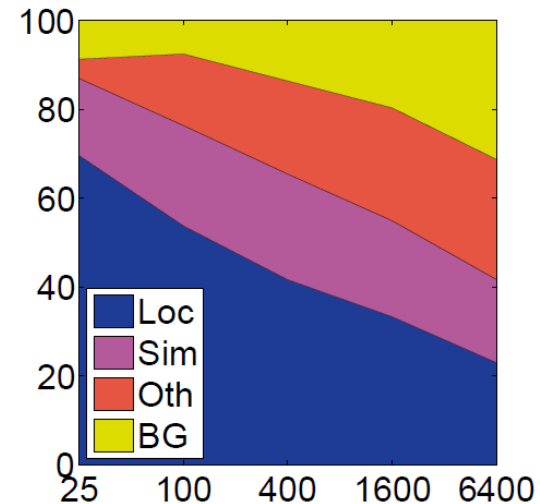
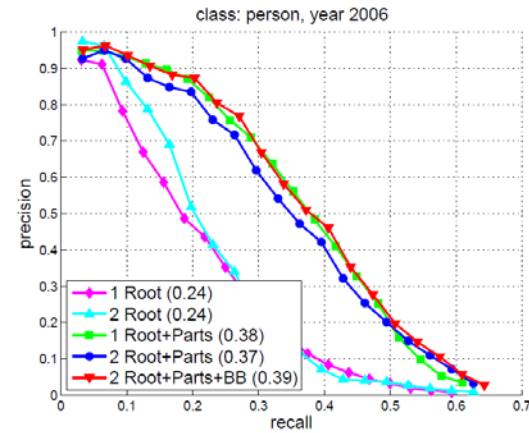
<http://mscoco.org>

# Algorithm Evaluation

Still debating...

The metric should be:

- Simple
- Relevant
- Robust



# MS COCO 2015

(full release)

## Early 2015

- 80-100 object categories
- 330k images
- 2M+ instances (700k people)
- Every instance segmented
- 5 sentences per image
- Keypoint annotations

<http://mscoco.org>





Microsoft COCO  
Common Objects in Context

Visit [mscoco.org](http://mscoco.org)

# Thank you!



**Tsung-Yi Lin**  
Cornell Tech



**Michael Maire**  
TTI Chicago



**Serge Belongie**  
MIT



**Lubomir Bourdev**  
Microsoft



**James Hays**  
Brown University



**Pietro Perona**  
Caltech



**Deva Ramanan**  
UC Irvine



**Ross Girshick**  
Microsoft Research



**Piotr Dollar**  
Microsoft Research



**Larry Zitnick**  
Microsoft Research

<http://mscoco.org>





# Turkopticon rating

[Trace Contour of Object](#) [View a HIT in this group](#)

**Requester:** ▼ vocds **HIT Expiration Date:** Sep 16, 2014 (6 days 23 hours) **Reward:** \$0.05

**HITs Available:** 1

communicativity:	<div style="width: 68%; background-color: #76b82a; border: 1px solid #ccc;"></div>	3.42 / 5
generosity :	<div style="width: 60%; background-color: #76b82a; border: 1px solid #ccc;"></div>	3.03 / 5
fairness :	<div style="width: 93%; background-color: #76b82a; border: 1px solid #ccc;"></div>	4.67 / 5
promptness :	<div style="width: 92%; background-color: #76b82a; border: 1px solid #ccc;"></div>	4.61 / 5


[What do these scores mean?](#)

Scores based on [76 reviews](#)

Terms of Service violation flags: 0

[Report your experience with this requester »](#)

[Developers](#) | [Press](#) | [Policies](#) | [Blog](#)

[Associates](#) | An  company

# DPM performance

	plane	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	moto	person	plant	sheep	sofa	train	tv	Avg.
DPMv5-P	45.6	49.0	11.0	11.6	27.2	50.5	43.1	23.6	17.2	23.2	10.7	20.5	42.5	44.5	41.3	8.7	29.0	18.7	40.0	34.5	29.6
DPMv5-C	43.7	50.1	11.8	2.4	21.4	60.1	35.6	16.0	11.4	24.8	5.3	9.4	44.5	41.0	35.8	6.3	28.3	13.3	38.8	36.2	26.8
DPMv5-P	35.1	17.9	3.7	2.3	7	45.4	18.3	8.6	6.3	17	4.8	5.8	35.3	25.4	17.5	4.1	14.5	9.6	31.7	27.9	16.9
DPMv5-C	36.9	20.2	5.7	3.5	6.6	50.3	16.1	12.8	4.5	19.0	9.6	4.0	38.2	29.9	15.9	6.7	13.8	10.4	39.2	37.9	19.1