Why a new dataset?

Continue our field’s momentum:
✓ Instance segmentation
✓ Non-iconic Images
Iconic object images
Iconic scene images

http://mscoco.org
Non-iconic images
Bird!
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

http://mscoco.org
Annotation pipeline
Divide and Conquer

1. Category Labeling
dog, bottle

2. Instance spotting

3. Instance segmentation

http://mscoco.org
1. Category Labeling
2. Instance Spotting
3. Instance Segmentation
After training
Properties
Number of categories vs. number of instances

Instances per category

Number of categories

- Caltech Ped
- PASCAL VOC
- COCO
- Caltech 256
- Caltech 101
- SUN
- ImageNet Detection
- ImageNet Classification

100000
10000
1000
100
10
1
- 330,000 images
- >2 million instances (700k people)
- Every instance is segmented
- 7.7 instances per image (3.5 categories)
Detection Performance (DPM V5)

<table>
<thead>
<tr>
<th></th>
<th>Person (mAP)</th>
<th>Average (mAP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PASCAL VOC</td>
<td>41.3</td>
<td>29.6</td>
</tr>
<tr>
<td>MS COCO</td>
<td>17.5</td>
<td>16.9</td>
</tr>
</tbody>
</table>
Beyond detection
Beyond detection

✓ Sentences

two giraffes 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

http://mscoco.org
Beyond detection

✓ Keypoints
  (provided by Facebook)
How to use COCO
Tools

- Python tools

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.
APIs

• Python and MATLAB

```plaintext
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
Going forward
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
Thank you!

Tsung-Yi Lin
Cornell Tech

Michael Maire
TTI Chicago

Serge Belongie

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

http://mscoco.org
Turkopticon rating

<table>
<thead>
<tr>
<th>Trace Contour of Object</th>
<th>View a HIT in this group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requester:</td>
<td>vocds</td>
</tr>
<tr>
<td>HIT Expiration Date:</td>
<td>Sep 16, 2014 (6 days 23 hours)</td>
</tr>
<tr>
<td>Reward:</td>
<td>$0.05</td>
</tr>
<tr>
<td>HITs Available:</td>
<td>1</td>
</tr>
<tr>
<td>communicativeness:</td>
<td>3.42 / 5</td>
</tr>
<tr>
<td>generosity:</td>
<td>3.03 / 5</td>
</tr>
<tr>
<td>fairness:</td>
<td>4.67 / 5</td>
</tr>
<tr>
<td>promptness:</td>
<td>4.61 / 5</td>
</tr>
</tbody>
</table>

What do these scores mean?

Scores based on 76 reviews
Terms of Service violation flags: 0
Report your experience with this requester »

http://mscoco.org
### DPM performance

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