

Segmentation

Danna Gurari

The University of Texas at Austin

Fall 2019



<https://www.ischool.utexas.edu/~dannag/Courses/CrowdsourcingForCV/CourseContent.html>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Review

- Last week
 - Scope of “crowdsourcing” in crowdsourcing for computer vision
 - How to recruit a crowd?
 - Who is the crowd?
 - How to collect high quality results with a crowd?
- Assignments (Class Website & Canvas)
 - Reading assignment 4 due yesterday
 - Reading assignment 5 due next week
 - Lab assignment 2 due in two weeks
- Questions?

Today's Topics

- Segmentation definition and applications
- Segmentation evaluation
- Crowdsourcing segmentations
- Class discussion (chosen by YOU 😊)
- Lab: connecting to AMT and submitting HITs

Today's Topics

- Segmentation definition and applications
- Segmentation evaluation
- Crowdsourcing segmentations
- Class discussion (chosen by YOU 😊)
- Lab: connecting to AMT and submitting HITs

Segmentation Definition

Partition image into coherent parts

Segmentation Problems

- Semantic
- Object
- Instance
- Salient
- Stuff
- Panoptic

Segmentation Problems

- Semantic
 - label category that each pixel belongs to

- Object

e.g.,

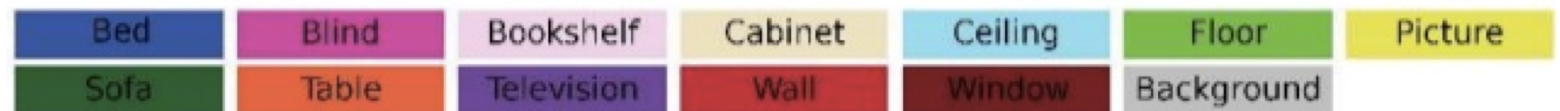


- Instance

- Salient

- Stuff

- Panoptic



Segmentation Problems

- Semantic
- **Object**
 - label all pixels that belong to a given category
- Instance
- Salient
- Stuff
- Panoptic

e.g.,



[Kovashka et al; 2016]

Segmentation Problems

- Semantic
- Object
- **Instance**
 - group all pixels that belong to each object
- Salient
- Stuff
- Panoptic

e.g.,



[Kovashka et al; 2016]

Segmentation Problems

- Semantic
- Object
- Instance
- **Salient**
 - label all pixels belonging to the most prominent object
- Stuff
- Panoptic

e.g.,



<https://http://mmcheng.net/msra10k/>

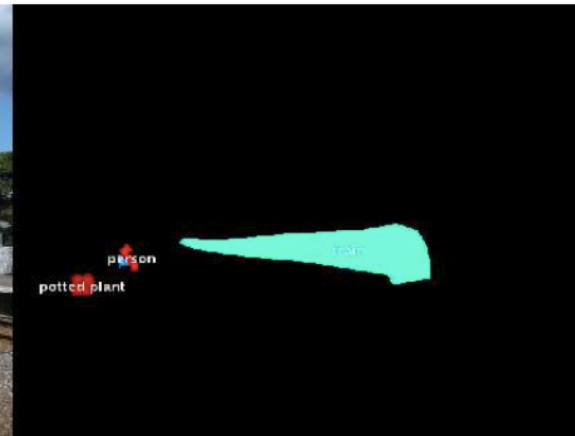
Segmentation Problems

- Semantic
- Object
- Instance
- Salient

e.g.,



“Things” (objects)



“Stuff” added



[Caesar et al; 2018]

- Stuff (as opposed to “things”)
 - label all pixels belonging to regions that are amorphous and uncountable (often background content)
- Panoptic

Segmentation Problems

- Semantic
- Object
- Instance
- Salient
- Stuff
- Panoptic

- unifies both stuff and instance labeling in a single task

e.g.,



(a) image



(b) semantic segmentation



(c) instance segmentation



(d) panoptic segmentation

[Kirillov et al; 2019]

Segmentation Applications

Botanical Species Classification

[White et al; UIST; 2006]



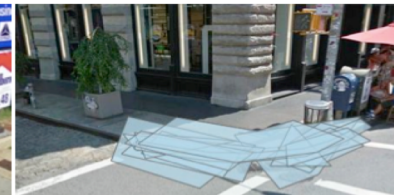
Street Accessibility Problems

[Hara et al; CHI 2013]

Object in Path



Missing Curb Ramp



Surface Problem



Prematurely Ending Sidewalk



Segmentation Applications

Materials Database

[Bell et al; SIGGRAPH; 2013]

Retexturing



(a) Target photo



(b) Retextured

Image Search



Query

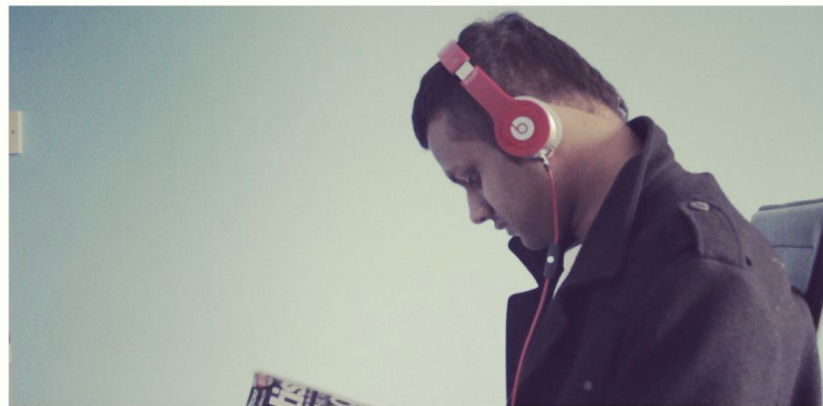


Results: wood floors in kitchens, sorted by diffuse color

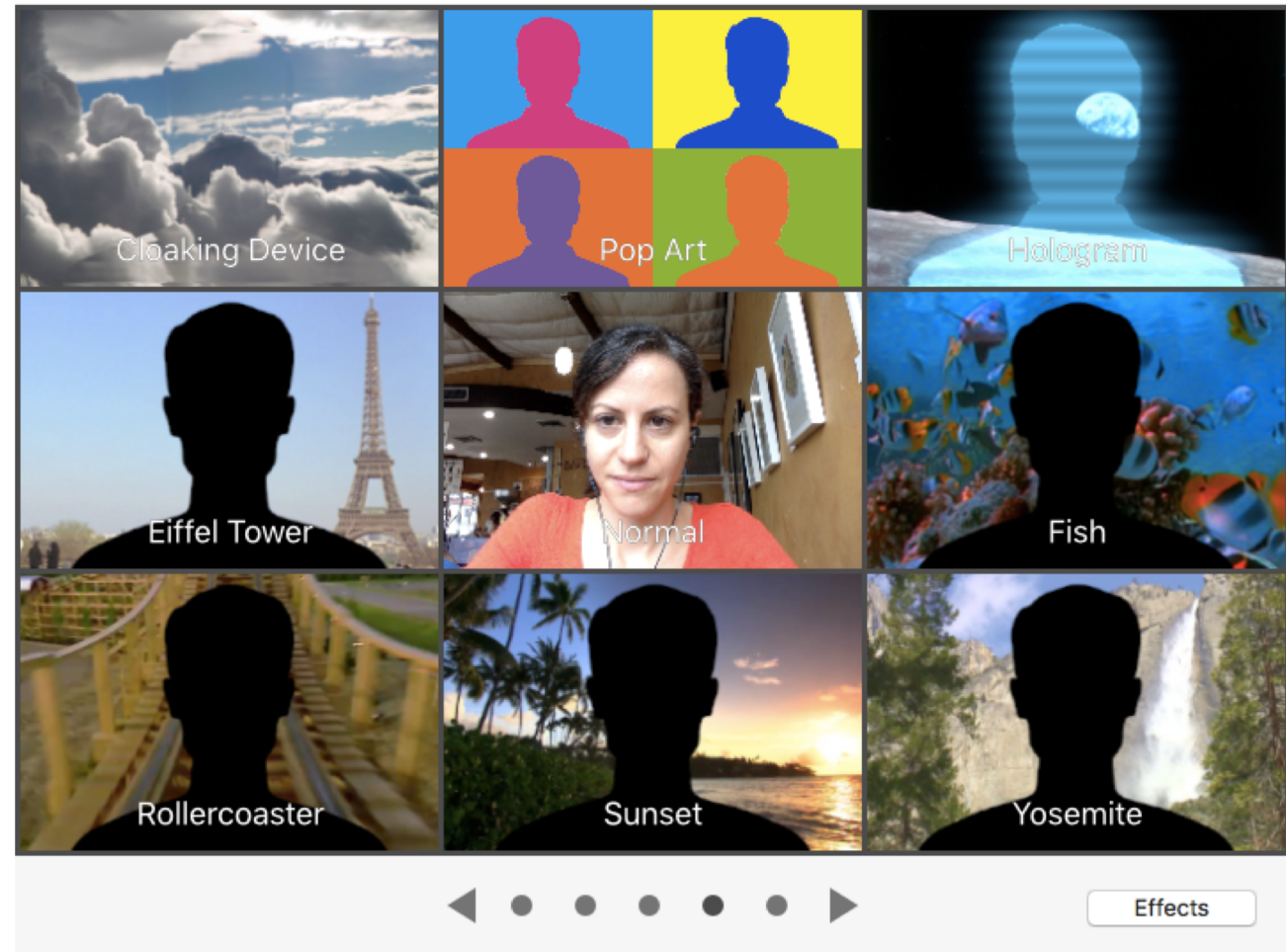
Segmentation Applications

Rotoscoping (more examples on [Wiki](#))

<https://www.starnow.co.uk/ahmedmohammed1/photos/4650871/before-and-after-rotoscopinggreen-screening>



Face Changers (e.g., Photo Booth, phone apps)



Segmentation Applications

Music Videos

[Pin it](#) [Like](#) [Share](#)

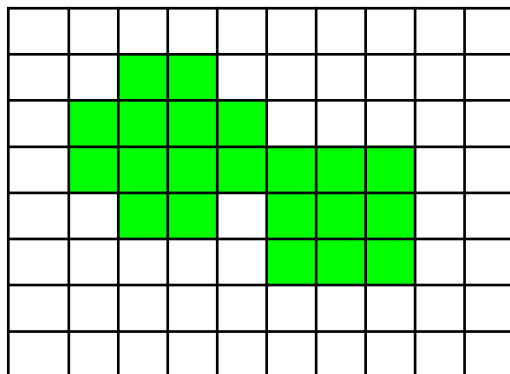


Today's Topics

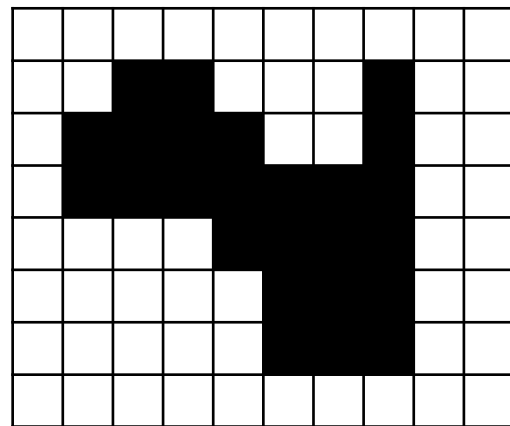
- Segmentation definition and applications
- **Segmentation evaluation**
- Crowdsourcing segmentations
- Class discussion (chosen by YOU 😊)
- Lab: connecting to AMT and submitting HITs

“Things” Segmentation

Ground Truth:

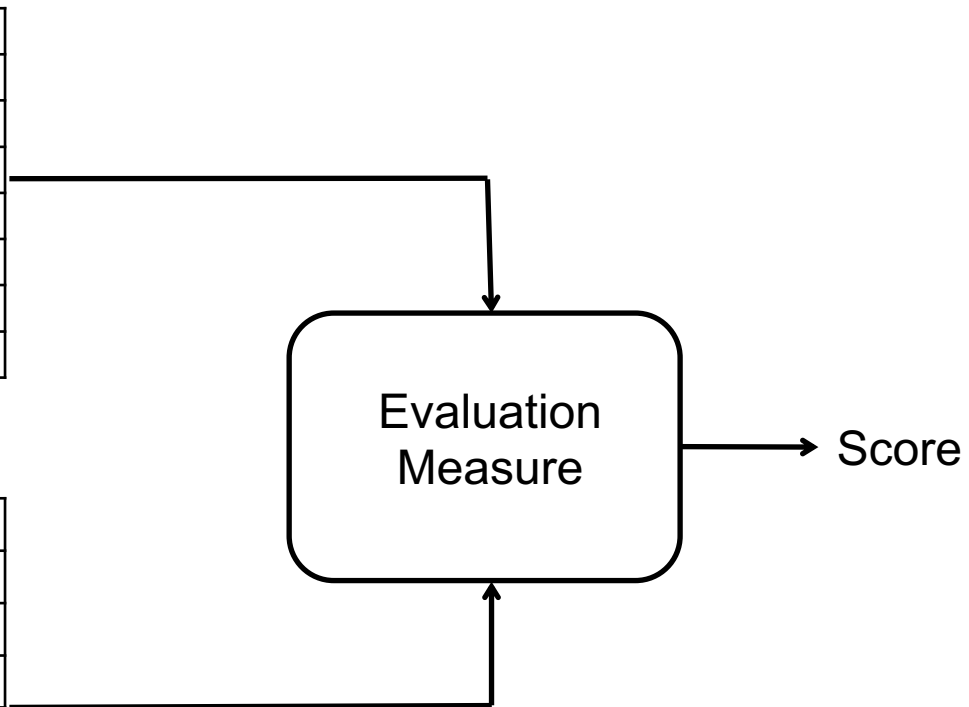


Algorithm:



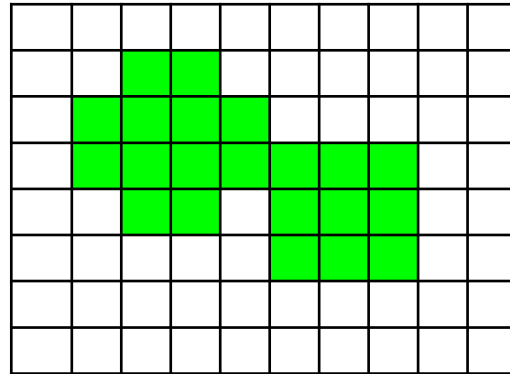
Evaluation Measure

Score

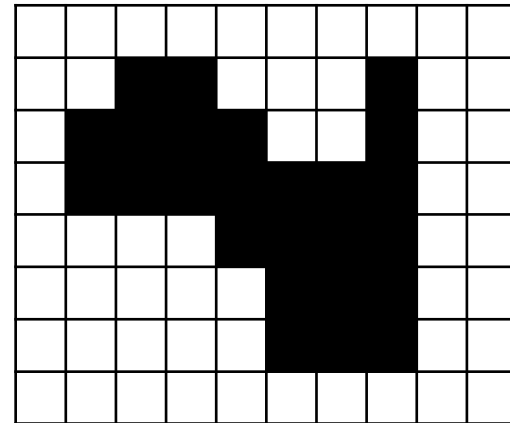


“Things” Segmentation: IoU Metric

Ground Truth:



Algorithm:

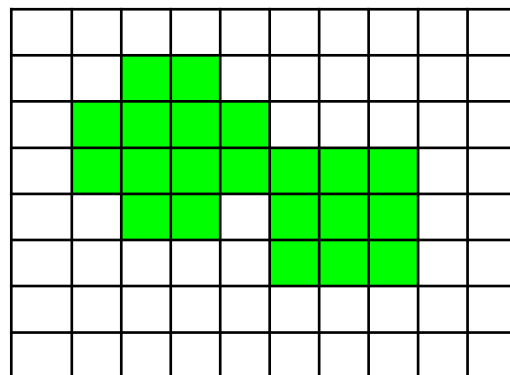


$$\frac{|A \cap B|}{|A \cup B|}$$

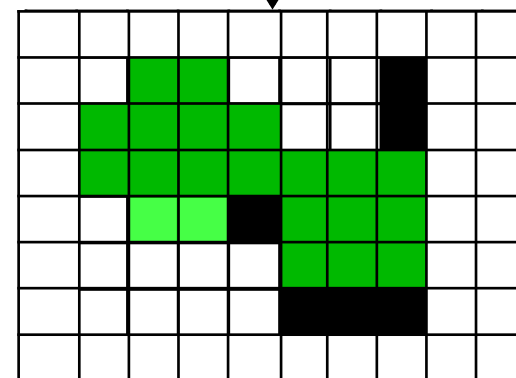
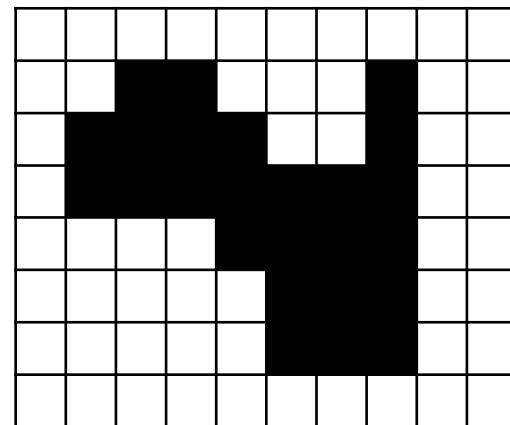
Score

“Things” Segmentation: IoU Metric

Ground Truth:

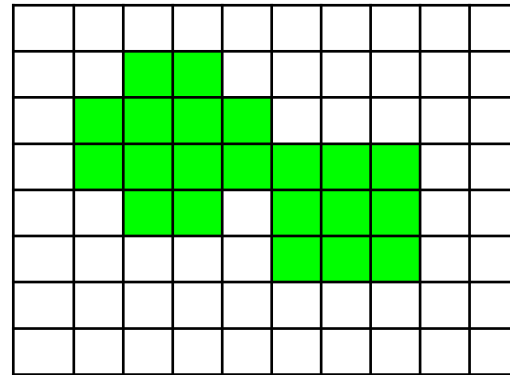


Algorithm:

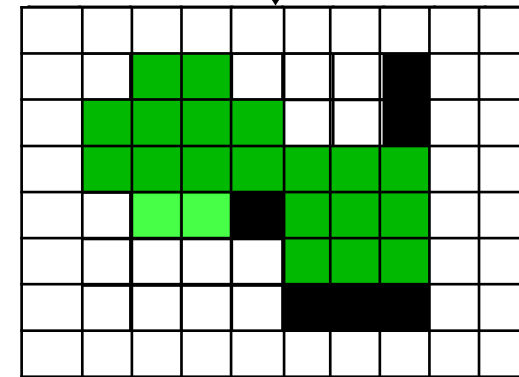
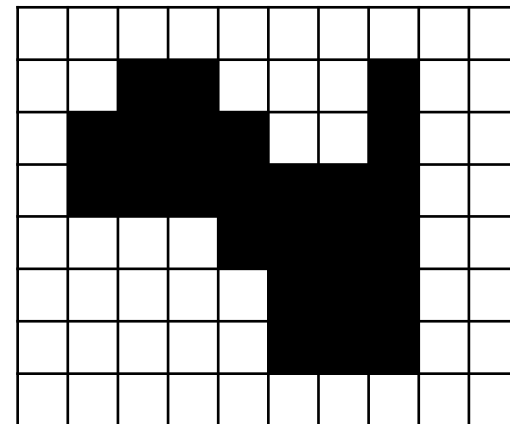


“Things” Segmentation: IoU Metric

Ground Truth:



Algorithm:



$$\frac{19}{27}$$

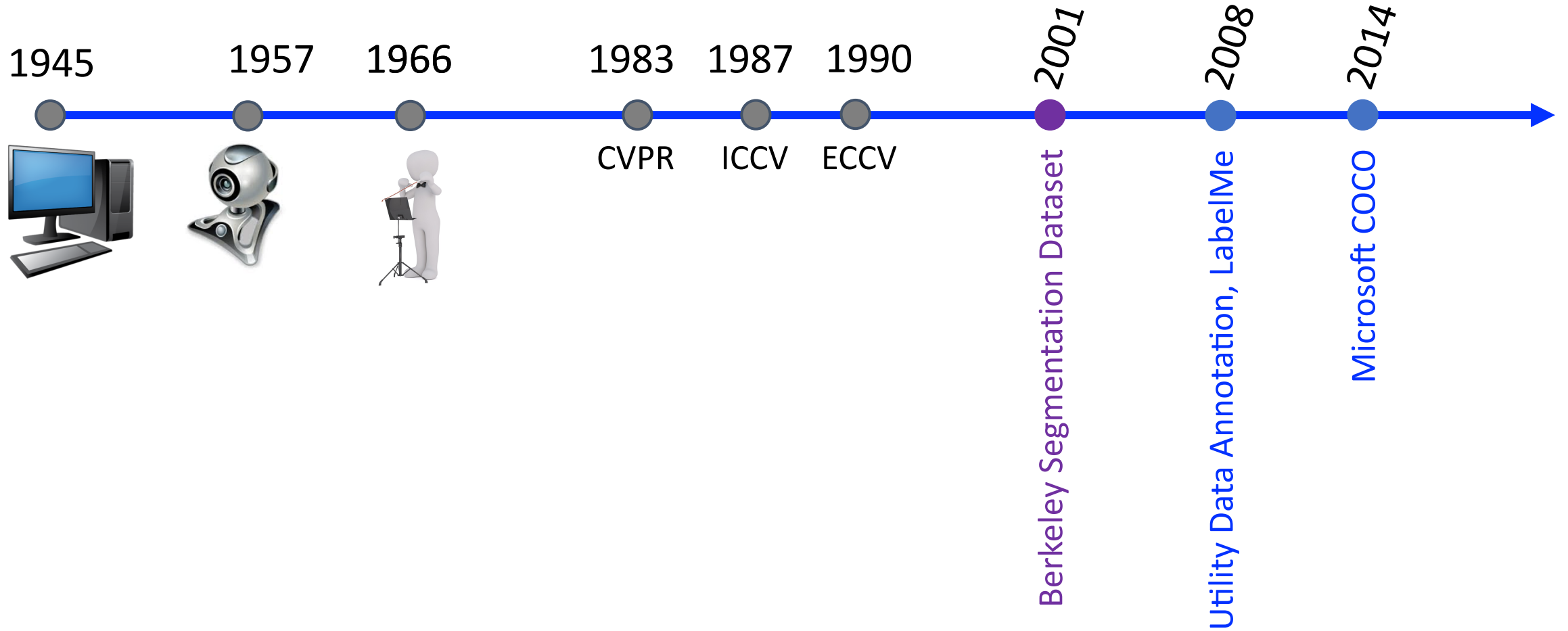
Semantic Segmentation

- **Pixel accuracy:** proportion of correctly classified pixels
- **Mean accuracy:** proportion of correctly classified pixels, averaged over all categories
- **Mean IoU:** IoU between predicted and ground-truth pixels, averaged over all categories
- **Weighted IoU:** IoU weighted by the total pixel ratio of each category

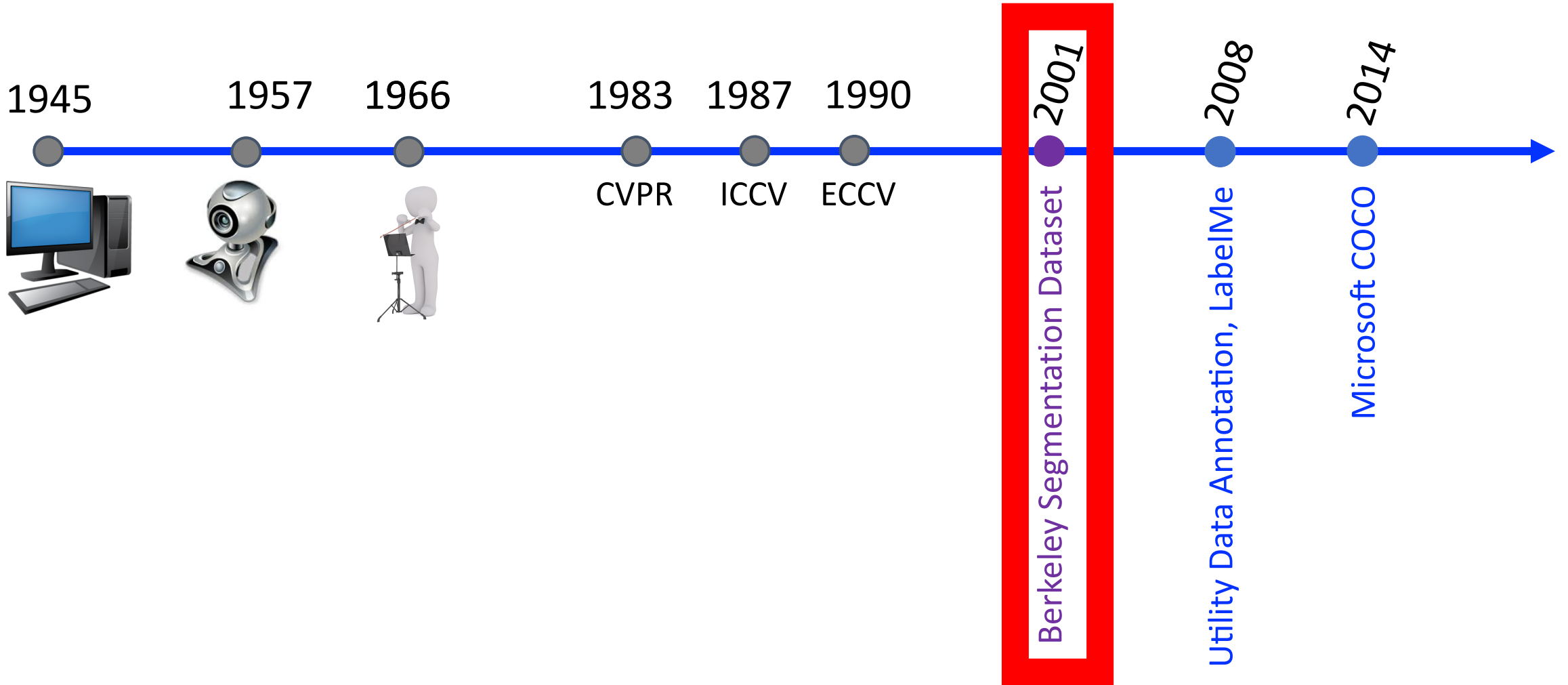
Today's Topics

- Segmentation definition and applications
- Segmentation evaluation
- **Crowdsourcing segmentations**
- Class discussion (chosen by YOU 😊)
- Lab: connecting to AMT and submitting HITs

Segmentation Datasets



Segmentation Datasets



Segmentation Datasets: BSD

1. Image Collection

- 800 481x321 images from Corel image database that contain at least one discernable object

Segmentation Datasets: BSD

1. Image Collection

- 800 481x321 images from Corel image database that contain at least one discernable object

2. Image Annotation

- Completed by students in a graduate-level computer vision class

Segmentation Datasets: BSD

1. Task Design

Instructions:

- “Divide each image into pieces, where each ...”

Interface:



Segmentation Datasets: BSD

1. Task Design

Instructions:

- “Divide each image into pieces, where each ...”

Interface:



“Divide each image into pieces, where each piece represents a distinguished thing in the image. It is important that all of the pieces have approximately equal importance. The number of things in each image is up to you. Something between 2 and 20 should be reasonable for any of our images.”

(intentionally vague so annotators will annotate in a “natural” way)

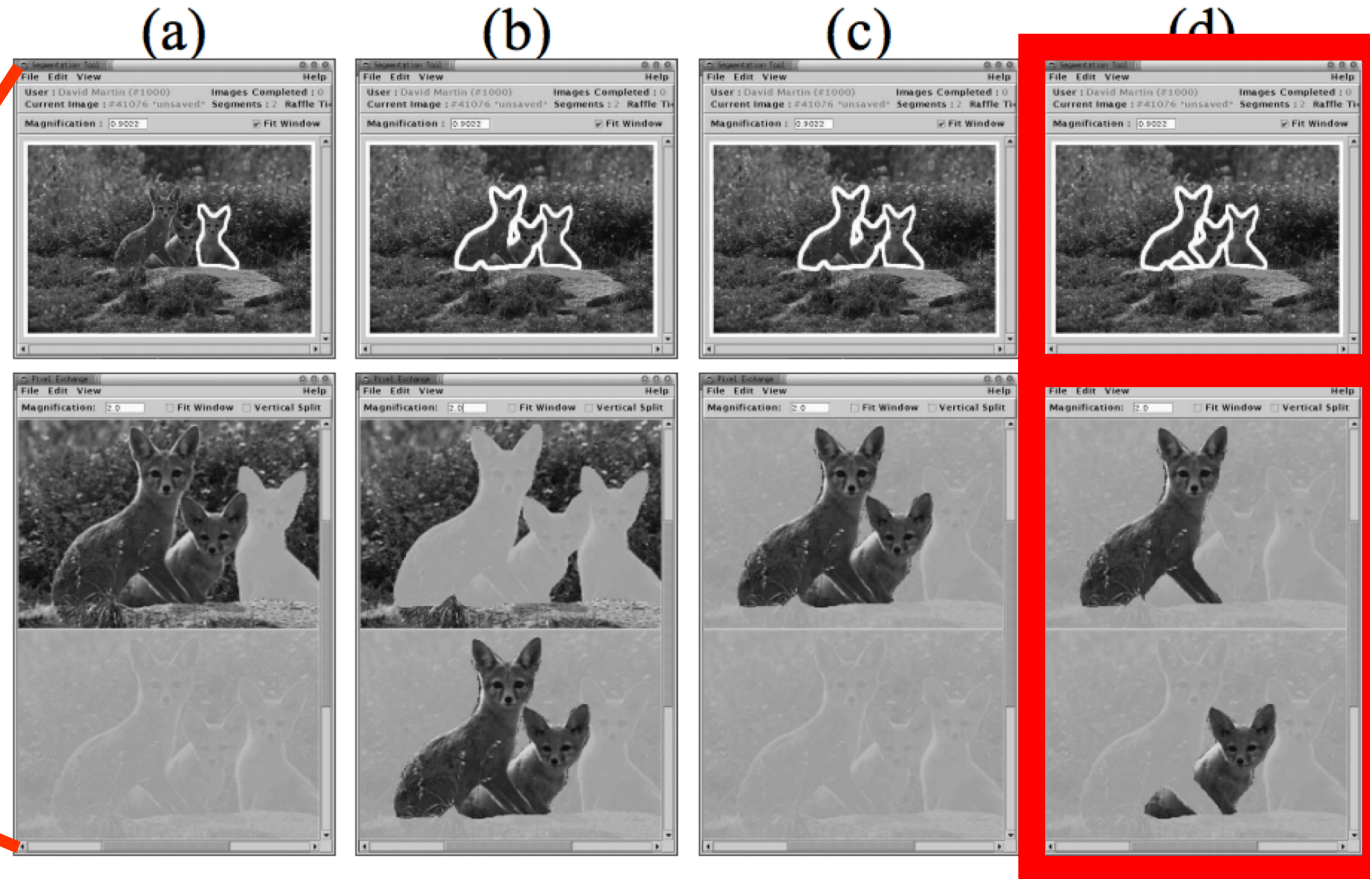
Segmentation Datasets: BSD

1. Task Design

Instructions:

- “Divide each image into pieces, where each ...”

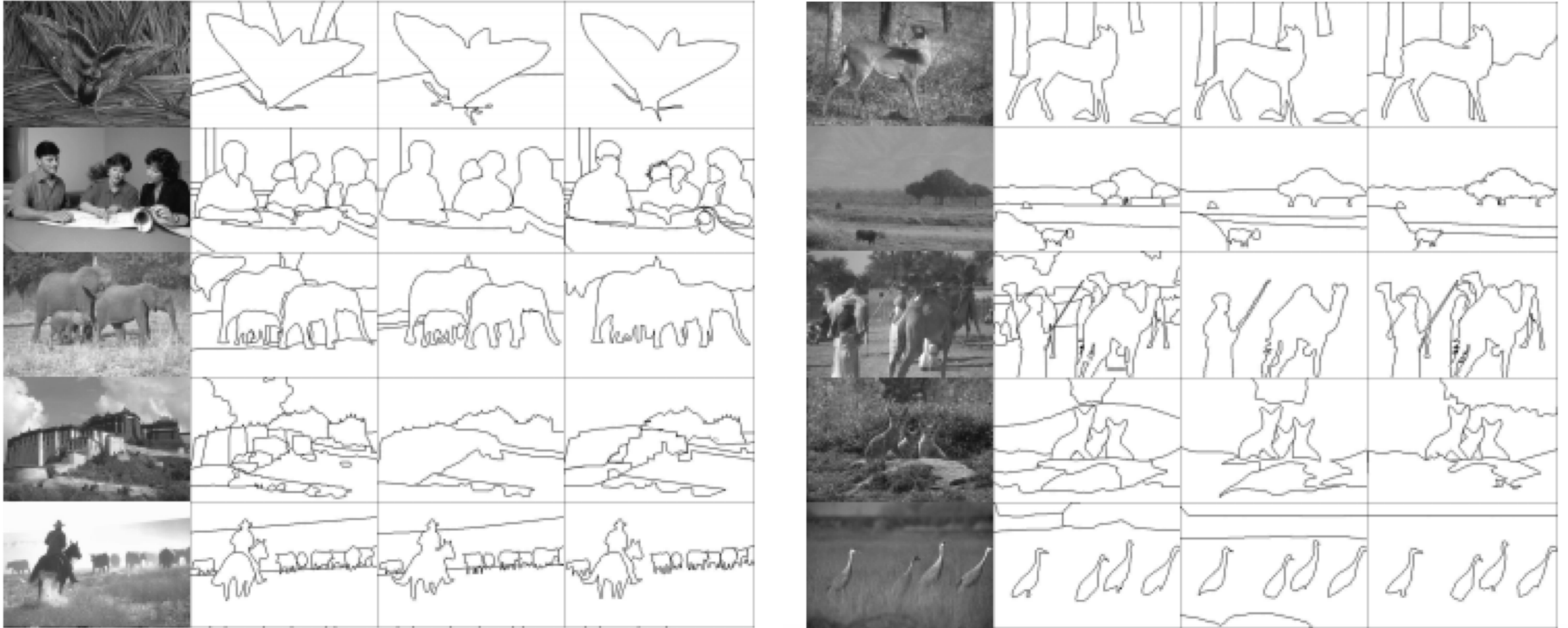
Interface:



Main window overlaid with segmentations

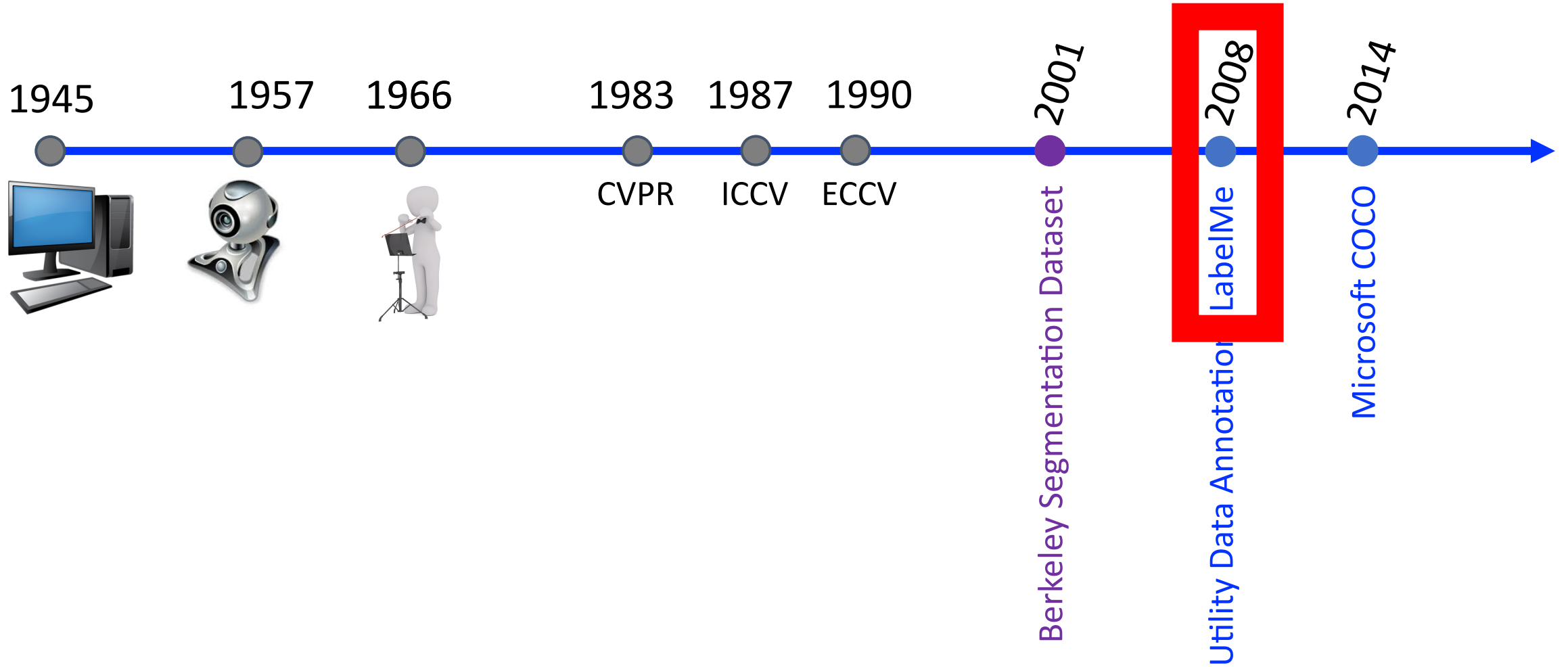
“Splitter” window in which annotator splits a region into two regions

Segmentation Datasets: BSD Results



David Martin, Charless Fowlkes, Doron Tal, Jitendra Milek. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. ICCV 2001.

Segmentation Datasets



Segmentation Datasets: LabelMe

1. Image Collection

- Most images taken by authors using a variety of hand-held digital cameras
- Includes videos taken with head-mounted web camera

Segmentation Datasets: LabelMe

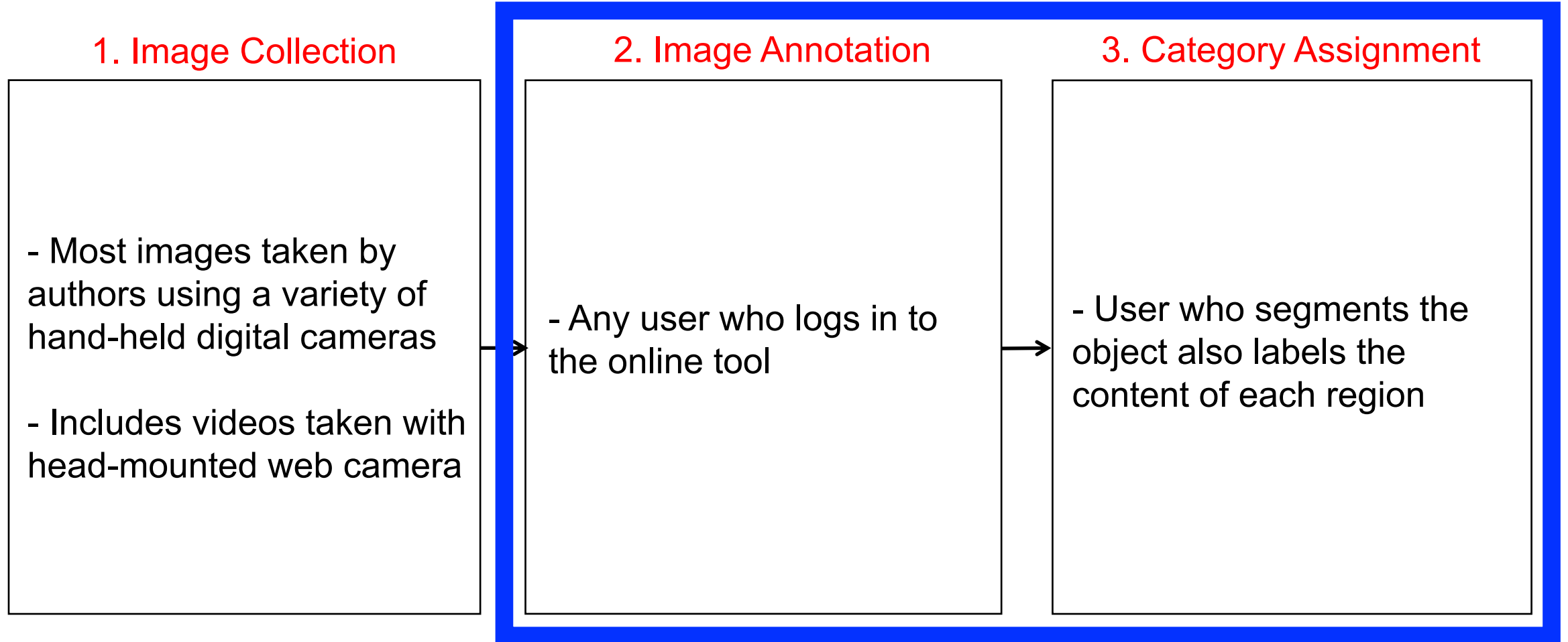
1. Image Collection

- Most images taken by authors using a variety of hand-held digital cameras
- Includes videos taken with head-mounted web camera

2. Image Annotation

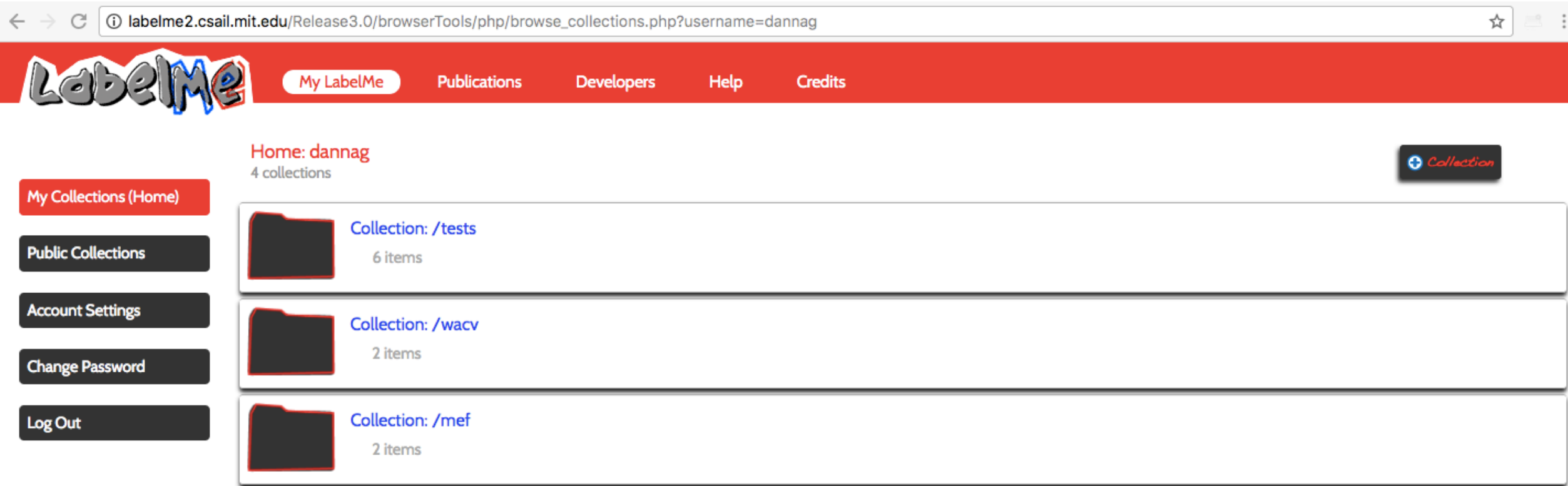
- Any user who logs in to the online tool

Segmentation Datasets: LabelMe



Segmentation Datasets: LabelMe

Online Database and Tool



The screenshot shows the LabelMe web interface. The browser address bar displays the URL: `labelme2.csail.mit.edu/Release3.0/browserTools/php/browse_collections.php?username=dannag`. The page features a red navigation bar with the LabelMe logo and menu items: "My LabelMe", "Publications", "Developers", "Help", and "Credits". On the left side, there is a vertical sidebar with buttons for "My Collections (Home)", "Public Collections", "Account Settings", "Change Password", and "Log Out". The main content area shows the user's profile: "Home: dannag" with "4 collections". A "+ Collection" button is visible in the top right of the main area. Below this, three collection entries are listed, each with a folder icon, the collection name, and the number of items:

- Collection: /tests (6 items)
- Collection: /wacv (2 items)
- Collection: /mef (2 items)

Bryan. C. Russell, Antonio Torralba, Kevin P. Murphy, and William T. Freeman.
LabelMe: a database and web-based tool for image annotation. IJCV 2008.

Segmentation Datasets: LabelMe

“I work in a small clothing shop. The shop is open from 10am to 8pm with only a short break at 2pm. Despite the long working hours I have a lot of free time. As I am the owner of the shop, I can do whatever I want during that time. I am always ready for the clients, however, in such a long day there are many hours of inactivity. I used to read a lot and books passed by my hands a great speed. I was starting to lose the pleasure that one feels when reading a good book. For this reason, when I started working with LabelMe it was very satisfying to know that I was doing something that had some scientific value and that it could be of use for somebody in the future.”



Antonio



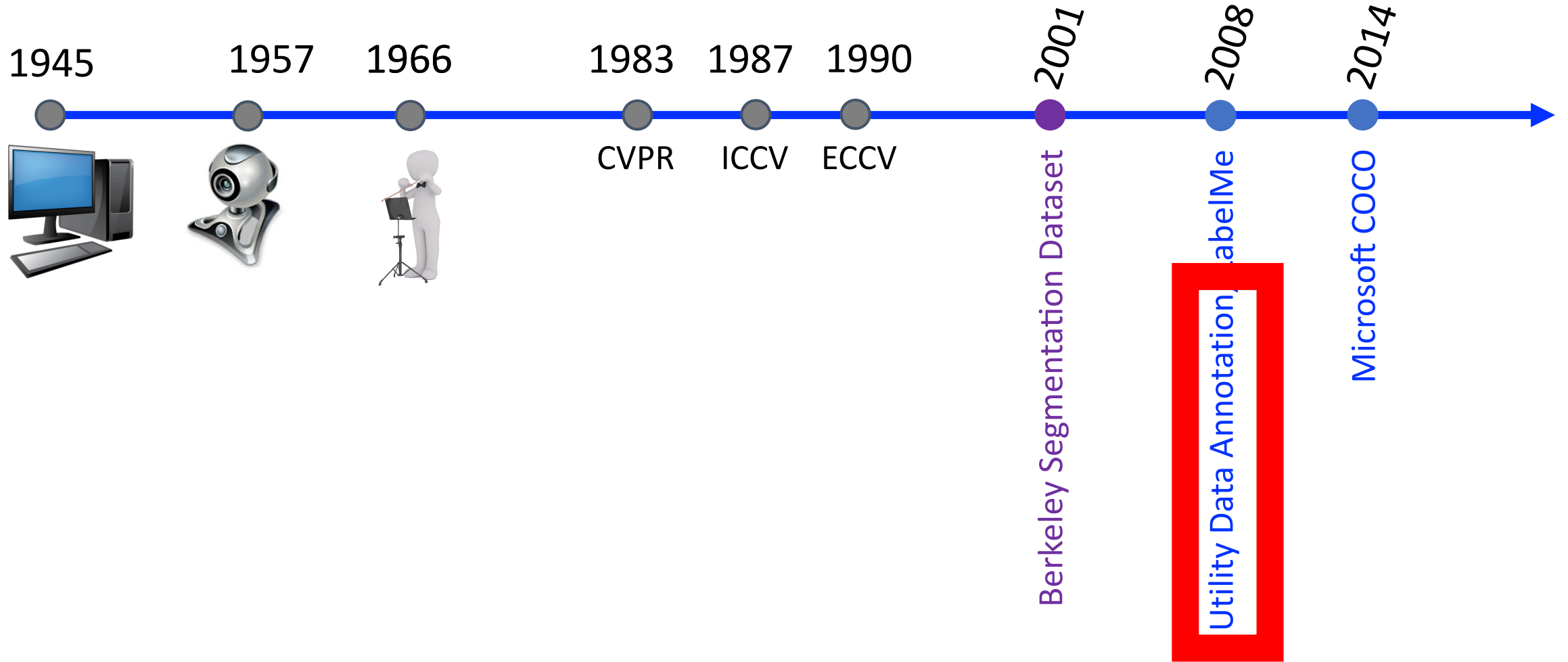
“she has labeled more than 250,000 objects” Antonio’s Mom

Adela Barriuso and Antonio Torralba. Notes on Image Annotation. arXiv 2012.

Bryan. C. Russell, Antonio Torralba, Kevin P. Murphy, and William T. Freeman.

LabelMe: a database and web-based tool for image annotation. IJCV 2008.

Segmentation Datasets



Segmentation Datasets: Utility Data Annotation

1. Image Collection

- Images from YouTube as well as vision datasets (LabelMe, Weizmann, and authors' images)

2. Image Annotation

- Workers employed from AMT to annotate images with multiple protocols

Segmentation Datasets: Utility Data Annotation

Regular Grid of Circles

Protocol 1

This is an experimental prototype. The tasks/pay of the HITs may change later.



Instructions

1. Find the person above in the image on the left.
2. Click on all small circles that overlap with the person. The circles will expand.
3. Submit the results by clicking "submit" button.

<< page (1 of 2)

Unfortunately this does not work with Internet Explorer yet. It will soon



Segmentation Datasets: Utility Data Annotation

Grid of Centers of Computed Superpixels

Protocol 2

This is an experimental prototype. The tasks/pay of the HITs may change later.

If there the person is too large, click only circles on the boundary.

Instructions

1. Find the person above in the image on the left.
2. Click on all small circles that overlap with the person. The circles will expand.
3. Submit the results by clicking "submit" button.

<< page (1 of 2)



Segmentation Datasets: Utility Data Annotation

Protocol 3

This is an experimental prototype. The tasks/pay of the HITs may change later.



Instructions
Draw a boundary around the person in the image.
<< page (1 of 3)

Label:
person

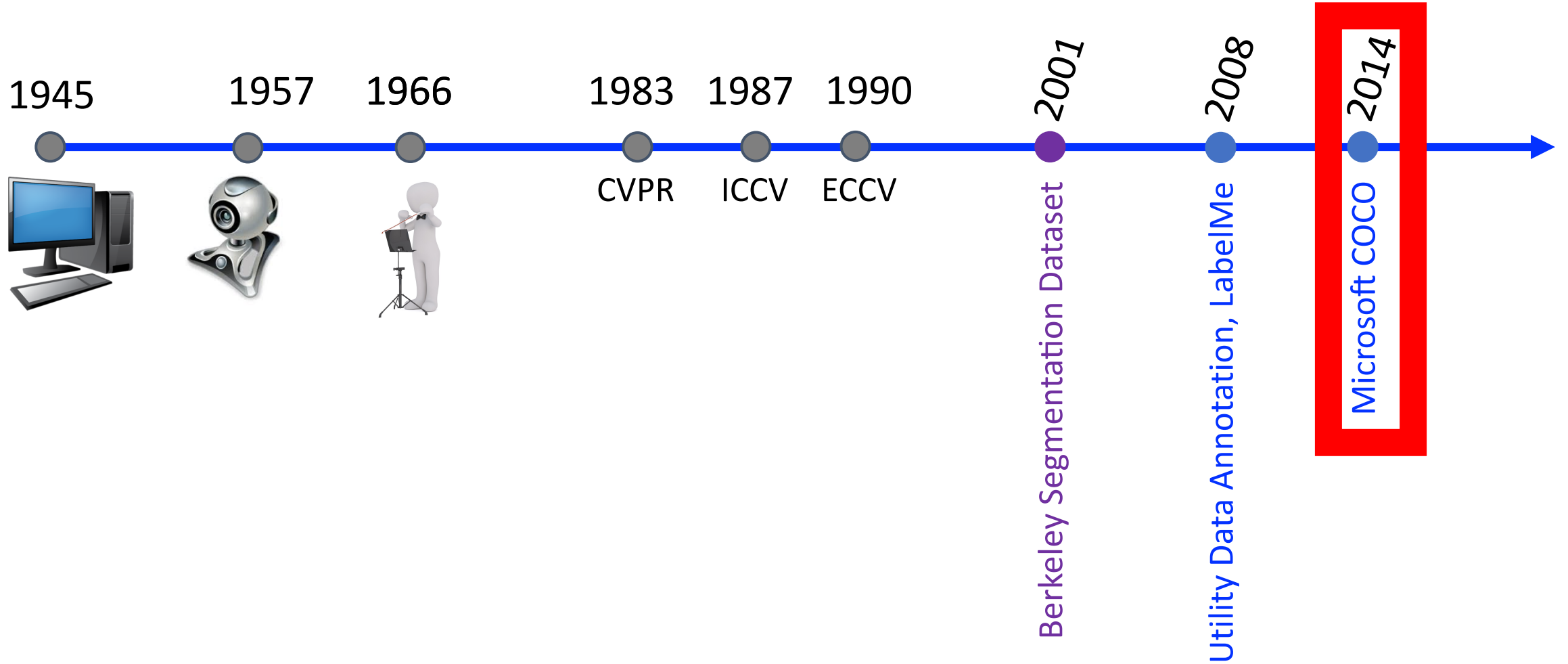
Add shape

Existing shapes

Delete shape



Segmentation Datasets



Segmentation Datasets: MSCOCO

Include “things”: objects that can easily be labeled; e.g., person, chair

1. Category Selection

- 272 candidate categories chosen from:
 - 1) WordNet, SUN, VOC, ...
 - 2) Most frequent words describing visual objects
 - 3) 4-8 yr olds listing objects in indoors/outdoors
- 91 categories chosen by author votes + coverage



Exclude “stuff”: objects with no clear boundaries; e.g., sky, grass,



Rationale: primary interest is in precise localization of object instances

Segmentation Datasets: MSCOCO

Selected 91 from 272 categories in bold (without *)

person	bicycle	car	motorcycle	bird	cat	dog	horse	sheep	bottle
chair	couch	potted plant	tv	cow	airplane	hat*	license plate	bed	laptop
fridge	microwave	sink	oven	toaster	bus	train	mirror*	dining table	elephant
banana	bread	toilet	book	boat	plate*	cell phone	mouse	remote	clock
face	hand	apple	keyboard	backpack	steering wheel	wine glass	chicken	zebra	shoe*
eye	mouth	scissors	truck	traffic light	eyeglasses*	cup	blender*	hair drier	wheel
street sign*	umbrella	door*	fire hydrant	bowl	teapot	fork	knife	spoon	bear
headlights	window*	desk*	computer	refrigerator	pizza	squirrel	duck	frisbee	guitar
nose	teddy bear	tie	stop sign	surfboard	sandwich	pen/pencil	kite	orange	toothbrush
printer	pans	head	sports ball	broccoli	suitcase	carrot	chandelier	parking meter	fish
handbag	hot dog	stapler	basketball hoop	donut	vase	baseball bat	baseball glove	giraffe	jacket
skis	snowboard	table lamp	egg	door handle	power outlet	hair	tiger	table	coffee table
skateboard	helicopter	tomato	tree	bunny	pillow	tennis racket	cake	feet	bench
chopping board	washer	lion	monkey	hair brush*	light switch	arms	legs	house	cheese
goat	magazine	key	picture frame	cupcake	fan (ceil/floor)	frogs	rabbit	owl	scarf
ears	home phone	pig	strawberries	pumpkin	van	kangaroo	rhinoceros	sailboat	deer
playing cards	towel	hyppo	can	dollar bill	doll	soup	meat	window	muffins
tire	necklace	tablet	corn	ladder	pineapple	candle	desktop	carpet	cookie
toy cars	bracelet	bat	balloon	gloves	milk	pants	wheelchair	building	bacon
box	platypus	pancake	cabinet	whale	dryer	torso	lizard	shirt	shorts
pasta	grapes	shark	swan	fingers	towel	side table	gate	beans	flip flops
moon	road/street	fountain	fax machine	bat	hot air balloon	cereal	seahorse	rocket	cabinets
basketball	telephone	movie (disc)	football	goose	long sleeve shirt	short sleeve shirt	raft	rooster	copier
radio	fences	goal net	toys	engine	soccer ball	field goal posts	socks	tennis net	seats
elbows	aardvark	dinosaur	unicycle	honey	legos	fly	roof	baseball	mat
ipad	iphone	hoop	hen	back	table cloth	soccer nets	turkey	pajamas	underpants
goldfish	robot	crusher	animal crackers	basketball court	horn	firefly	armpits	nectar	super hero costume
jetpack	robots								

Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. Microsoft COCO: Common Objects in Context. ECCV 2014.

Segmentation Datasets: MSCOCO

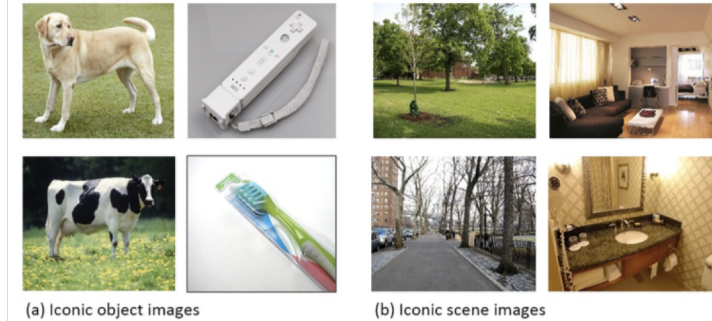
1. Category Selection

- 272 candidate categories chosen from:
 - 1) WordNet, SUN, VOC, ...
 - 2) Most frequent words describing visual objects
 - 3) 4-8 yr olds listing objects in indoors/outdoors
- 91 categories chosen by author votes + coverage

2. Image Collection

- Images scraped from Flickr because it is believed to often have non-iconic images
- Query: object + object or scene + scene
- Query: unusual categories
- Crowd workers flag images that meet selection criteria of containing multiple objects

Iconic images commonly retrieved with Google, Bing, etc:



Goal: images with **contextual** information and taken from **non-canonical** viewpoints



MSCOCO: Task 1

Grid of 128 images:

Task: select images that contain BOTH a person AND a bicycle

Instructions:
Please click and select images that contain **BOTH** a person(s) **AND** a bicycle(s).
Do **NOT** select an image that contains **ONLY** a person(s) or **ONLY** a bicycle(s).
(It is right to not select any image if none of image contains both categories.)

 + 

You can de-select the image by clicking on it again.
Please do not select cartoons or paintings.

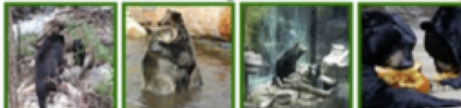


MSCOCO: Task 2

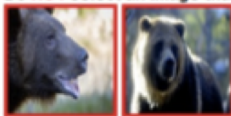
Grid of 128 images:

Task: select images that contain a bear(s)


Instructions:
Please click and select images that contain **MULTIPLE** objects **AND** at least one bear.



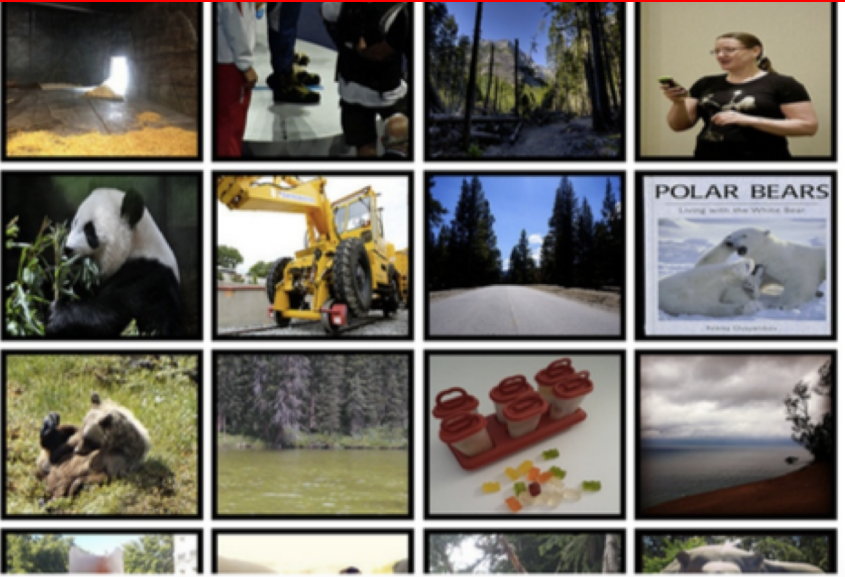

Do **NOT** select an image that contains **ONLY** a bear(s).



Do **NOT** select an image that contains **NO** bear(s).



You can de-select the image by clicking on it again.
Please do not select cartoons or paintings.



Segmentation Datasets: MSCOCO

1. Category Selection

- 272 candidate categories chosen from:
 - 1) WordNet, SUN, VOC, ...
 - 2) Most frequent words describing visual objects
 - 3) 4-8 yr olds listing objects in indoors/outdoors
- 91 categories chosen by author votes + coverage

2. Image Collection

- Images scraped from Flickr because it is believed to often have non-iconic images
- Query: object + object or scene + scene
- Query: unusual categories
- Crowd workers flag images that meet selection criteria of containing multiple objects

3. Image Annotation

- Task of having AMT crowd workers delineate boundary of objects belonging to each category completed over a series of multiple microtasks

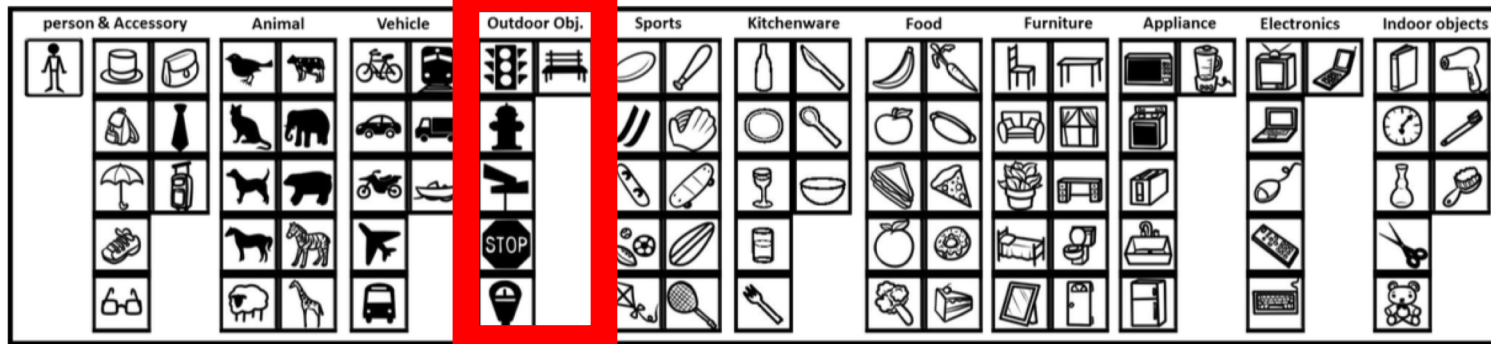
Segmentation Task Decomposition

1. Category Assignment

- AMT crowd workers assign the categories present in each image by locating one instance of each category

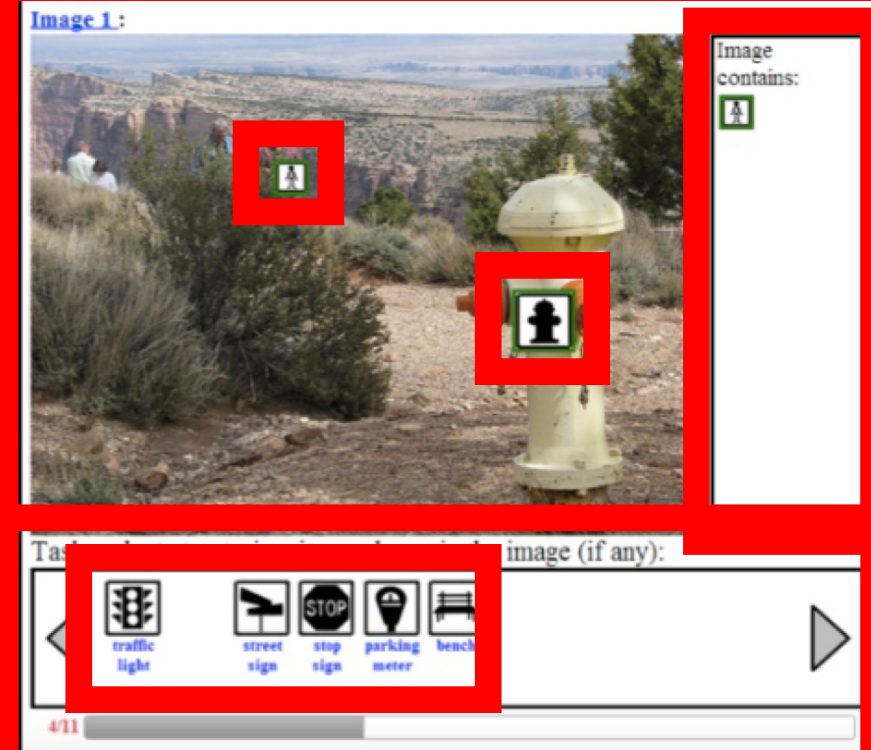
Category Assignment Task

11 Groupings



For high recall, 8 people per image were solicited to do this task

Instructions (PLEASE ACCEPT THE HIT TO GET STARTED):
Please drag and drop icons from the bottom panel to matching objects in the image. If an icon matches multiple objects you can drag the icon onto any of the objects. There are 11 sets of objects to drag onto the image. Use the buttons or arrow keys to cycle through them. There are total of 8 images to label.
Please drag and drop ICONS to matching objects in the image.



Segmentation Task Decomposition

1. Category Assignment

- AMT crowd workers assign the categories present in each image by locating one instance of each category

2. Instance Tagging


- AMT crowd workers mark location of each instance of the “thing” with a crosshair

Instance Tagging Task


Instructions (PLEASE ACCEPT THE HIT TO GET STARTED):

- Mark **each occurrence** (if any) of the following object: cow.
- You only need to mark up to 10 instances if multiple cow(s) exist in the image. It is possible for some images that this object does not appear.
- The blinking icon (Hint) shows where one instance of the object could be. The Hint is **NOT ALWAYS** correct.
- Type **N** to go to the next image and **B** to go back.
- There are 50 images in this HIT.


Good Example




Bad Example (Do not click)



Left Click: Add marker Right Click: Delete marker Drag & Drop: Move marker

 7 cow(s) found in this image.

Back [B] Next [N] Hint [H]



“magnifying glass” feature: doubles resolution of currently selected region to assist with small objects.

Segmentation Task Decomposition

1. Category Assignment

- AMT crowd workers assign the categories present in each image by locating one instance of each category

2. Instance Tagging

- AMT crowd workers mark location of each instance of the “thing” with a crosshair

3. Object Segmentation

- AMT crowd workers delineate boundary of specified object(s)
- Another group of AMT workers verify quality of object segmentations

Object Seg.

Instructions: carefully trace around regions that have a **single sports ball** indicated by the icon. (1/3)



Not sure what object sports ball is? Click on [here](#) to see examples!

Draw (D) Adjust (A) Undo (Ctrl-Z) Redo (Ctrl-Y) Close (Right-click) Delete (Delete)

Move to Target (M) Zoom In (I) Zoom Out (O) Reset Zoom (ESC)

Please Accept HIT to get started! [Examples](#) [Instructions](#)

Tips: Using "Move to target" (M) and "Zoom In" (I) for the small object!
Please pay attentions to trace boundary carefully. Work will be rejected if not follow the instruction.



Training task per object category required.

Object Seg.

- Crowd labeling is similar to semantic segmentation as object instances are not individually identified.
- Crowd labeling is employed for images containing 10+ instances of an object category.

Draw all unlabeled **person(s)** in the image.

- Find and draw on **all person(s)** that haven't been labeled.
- It's okay to overlap to labeled region.
- You need to label two images that contain unlabeled person(s) to complete
- Work will be rejected if **not carefully** drawn or unlabeled person(s) remain.

Submit

No Unlabeled Person(s)

Draw (D)

Erase (E)

Zoom In (Z)

Zoom Out (X)



Object Seg.

Seeded gold standards: 4 of 64 segmentation known to be bad; a worker must identify 3 of the 4 known bad segmentations to complete the task.

Verification step: 3-5 workers judged each segmentation and indicate whether it matched the instance well or not.

Blocked workers: those who often produced poor segmentations were blocked and their work not used

64 examples

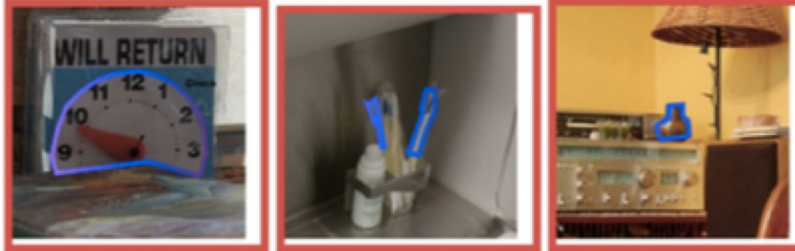
Task: select images that have **WRONG** object contour for **toothbrush**.

Examples:

Right Object Contour



Wrong Object Contour (not toothbrush, only contains parts of visible object contour, or multiple objects)



Tips: use **n** and **b** keys to move between rows of image.



Segmentation Datasets: MSCOCO Summary

1. Category Selection

- 272 candidate categories chosen from:
 - 1) WordNet, SUN, VOC, ...
 - 2) Most frequent words describing visual objects
 - 3) 4-8 yr olds listing objects in indoors/outdoors
- 91 categories chosen by author votes + coverage

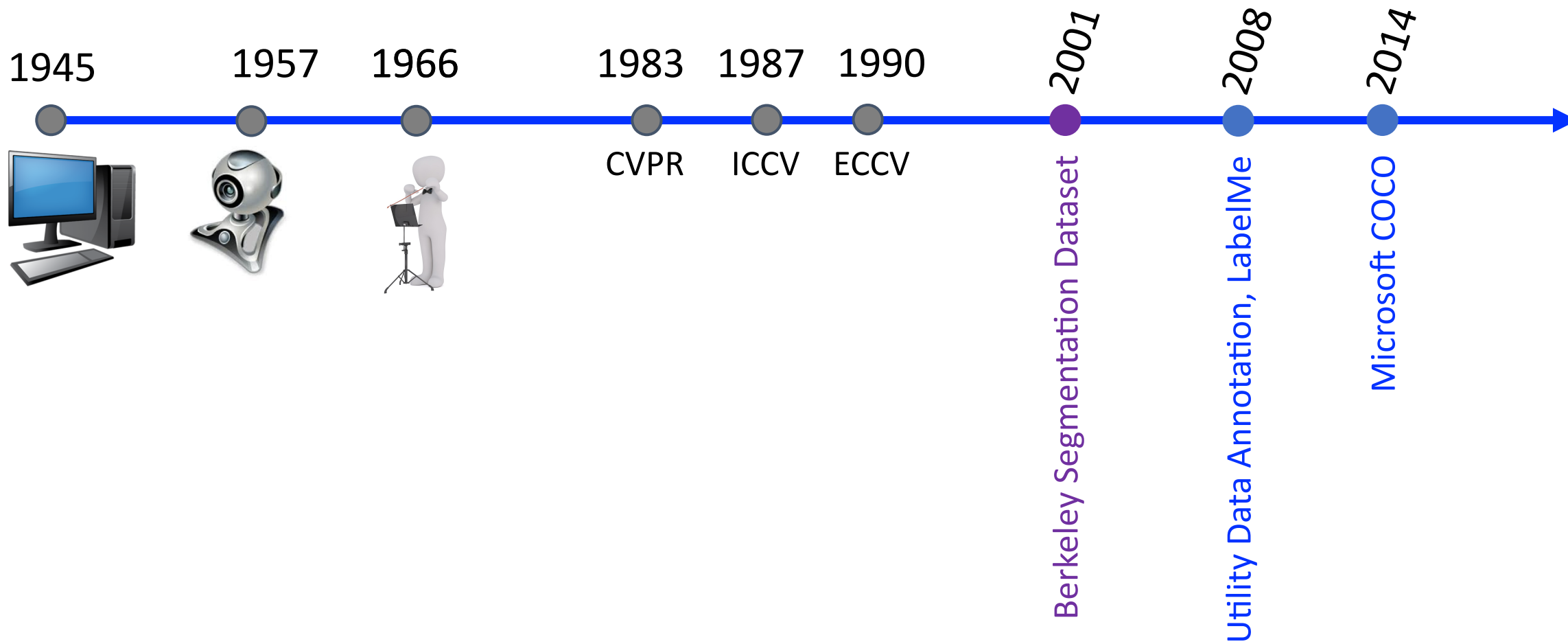
2. Image Collection

- Images scraped from Flickr because it is believed to often have non-iconic images
- Query: object + object or scene + scene
- Query: unusual categories
- Crowd workers flag images that meet selection criteria of containing multiple objects

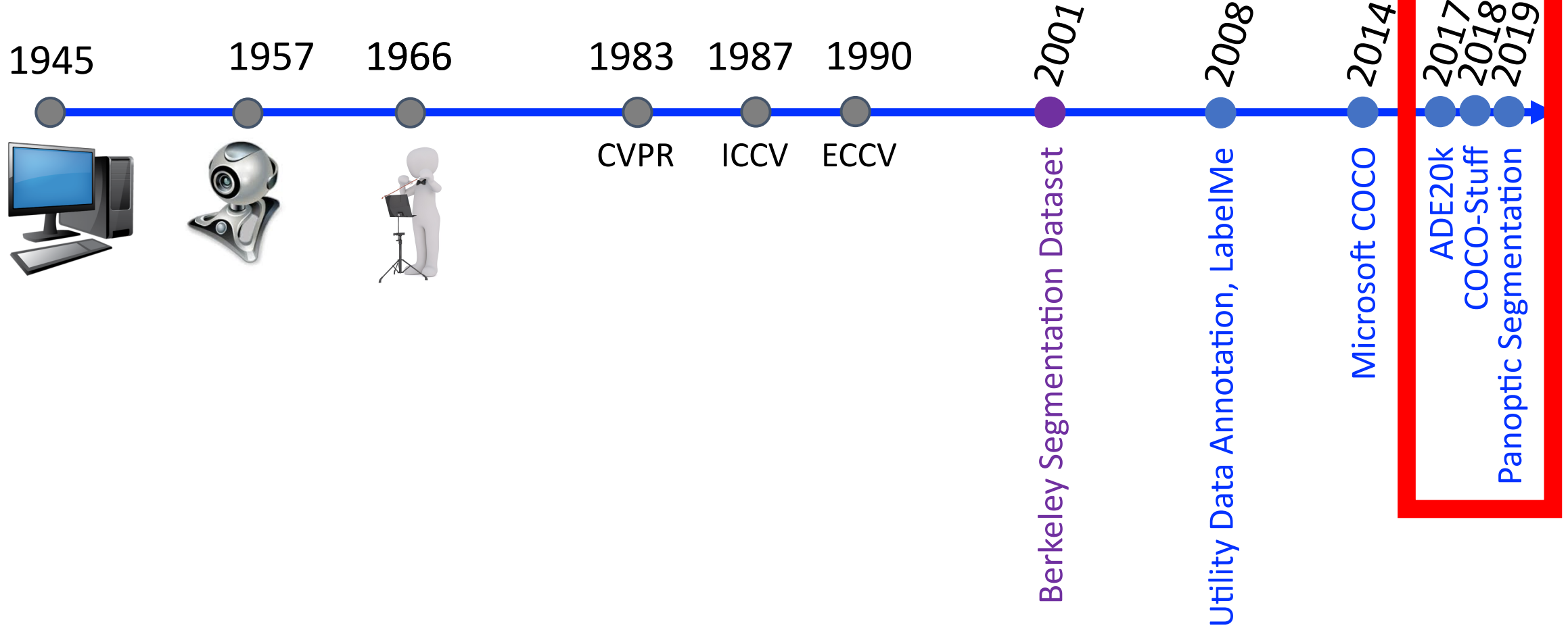
3. Image Annotation

- Task of having AMT crowd workers delineate boundary of objects belonging to each category completed over a series of multiple microtasks

Segmentation Datasets



Segmentation Datasets



And many more datasets such as the following...

Segmentation Datasets: Summary

- Key steps in creating dataset:

1. Category Selection



2. Image Collection



3. Image Annotation



How is this process different from the process used for object/scene/attribute recognition?

Segmentation Datasets: Summary

- Key steps in creating dataset:

1. Category Selection



2. Image Collection



3. Image Annotation



1. Annotate ON image
2. Task decomposition

Today's Topics

- Segmentation definition and applications
- Segmentation evaluation
- Crowdsourcing segmentations
- **Class discussion (chosen by YOU 😊)**
- Lab: connecting to AMT and submitting HITs

Today's Topics

- Segmentation definition and applications
- Segmentation evaluation
- Crowdsourcing segmentations
- Class discussion (chosen by YOU 😊)
- **Lab: connecting to AMT and submitting HITs**