

IMAGE ANNOTATION

Yu-Ting Peng

OUTLINE

- Current object recognition datasets
- Human Computation
- LabelMe
- ESP
- Pickaboom
- CAPTCHA&RECAPTCHA

CURRENT OBJECT RECOGNITION DATASETS

Tight
crops

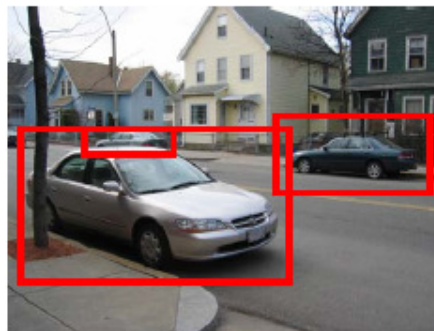


Captions



Car, street, sky, house, road,
Sidewalk, tree bark

Bounding
boxes



HUMAN-BASED COMPUTATION

- Humans can recognize 30000 entry-level object categories.
- Current techniques insufficient- can only recognize a few object categories

WEB-BASED ANNOTATION TOOLS

- Provide a way of building large annotated datasets by relying on the collaborative effort of a large population of users
- Provide a drawing interface that works on many platforms, is easy to use, and allow **instant sharing** of the collected data.
- Examples
 - LabelMe
 - ESP
 - Peekaboom



- Tool went online July 1st, 2005
- Til Feb 11th, 2009

Visitors: 64771

Available Images: 176180

Annotated Images: 51455

Object categories: 4418

STRENGTH

- Object class recognition and localization
- Drawing polygons
- Runs on almost any web browser-Javascript drawing interface
- Resulting labels are stored in the XML file format-makes the annotations portable and easy to extend.
- Matlab toolbox

Zoom
 Erase
 Help
 Make 3D
 Upload image

You are: **emma**
[\(sign out\)](#)

Show me anot

There are **335662** labelled objects

Polygons in this image ([IMG](#), [XML](#))

- [no parking sign](#)
- [window](#)
- [open window](#)
- [fence](#)
- [pickup truck](#)
- [balcony](#)
- [sign](#)
- [telephone pole](#)
- [house](#)
- [sky](#)
- [building](#)
- [building](#)
- [car occluded](#)
- [trash can](#)
- [window](#)
- [air conditioning](#)
- [buildings](#)
- [car occluded](#)
- [column](#)
- [road](#)
- [car side crop](#)
- [car side crop](#)
- [car side crop](#)
- [tree](#)
- [window](#)

What is this object?

Done
 Delete

3D



QUERYING OBJECTS



[Labeling tool](#) [Help](#) [Downloads](#) ^{New!} [About](#) [Matlab Toolbox](#) [Upload Images](#) [Stats](#)

boat

Your query (boat) matches **1155** images.

Parts within each object are shown in color (only when labeled). See [summary statistics](#) for all the objects in the dataset, and the [wordnet hierarchy](#).

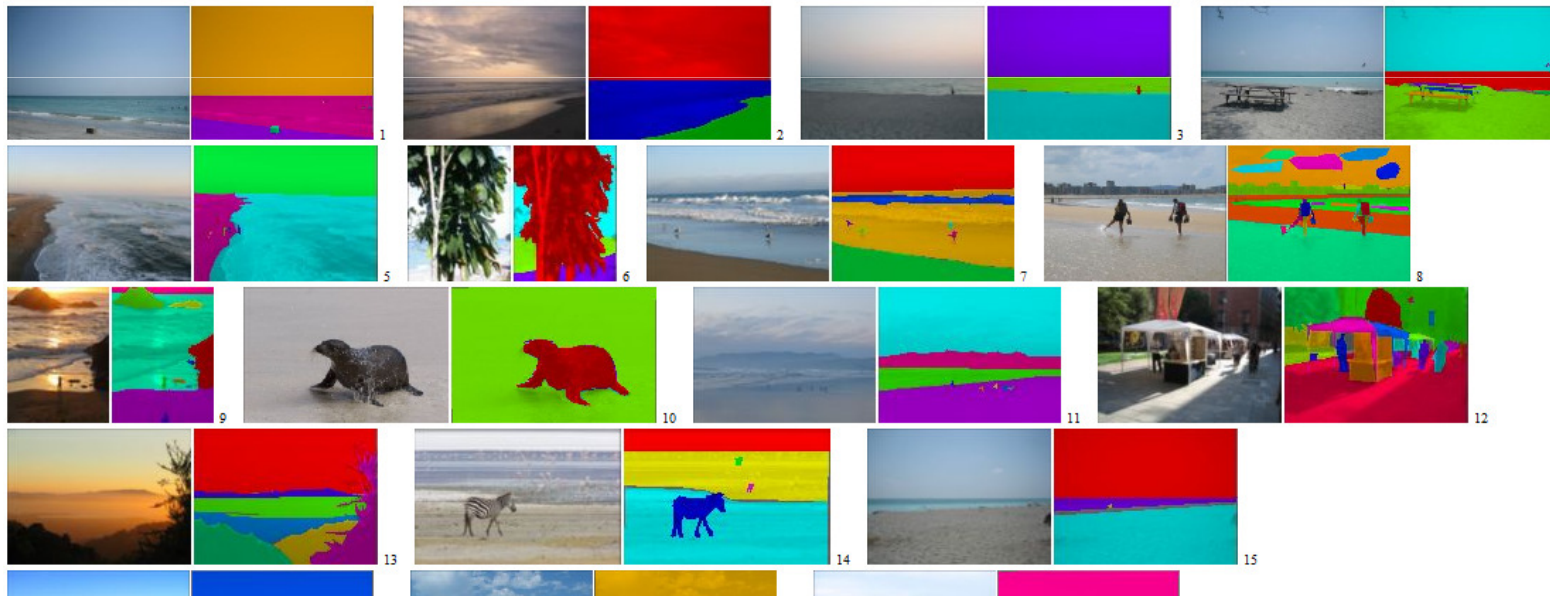


QUERYING SCENES


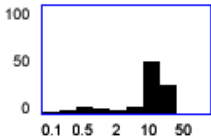
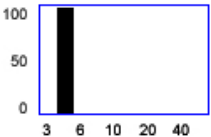


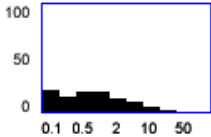
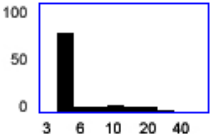
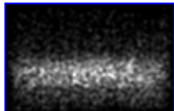
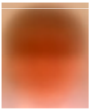
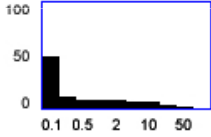
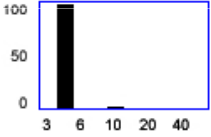
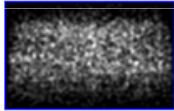

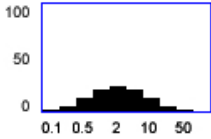
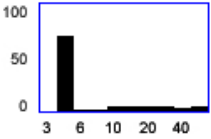
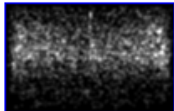

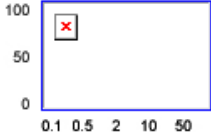
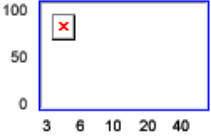


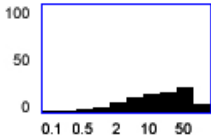
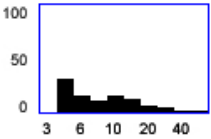
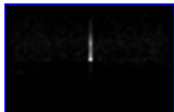
LobeIME

beach

Your query (beach) matches 236 images



STATISTICS

Description	Average	Occupied area	Boundary points	Object location
1) person walking Exact: 25133 Approx: 25801 Words: person(27714) , walking(25828)				
2) car Exact: 5581 Approx: 11158				
3) head Exact: 5509 Approx: 5652				
4) tree Exact: 4779 Approx: 7838				
5) window Exact: 3714 Approx: 4087				
7) building Exact: 2466 Approx: 3216				

Viewpoint labeling

coming soon!



Search objects

Instructions: Match the orientation of the object below to one of the templates. ([LabelMe](#), [XML](#))

Skip

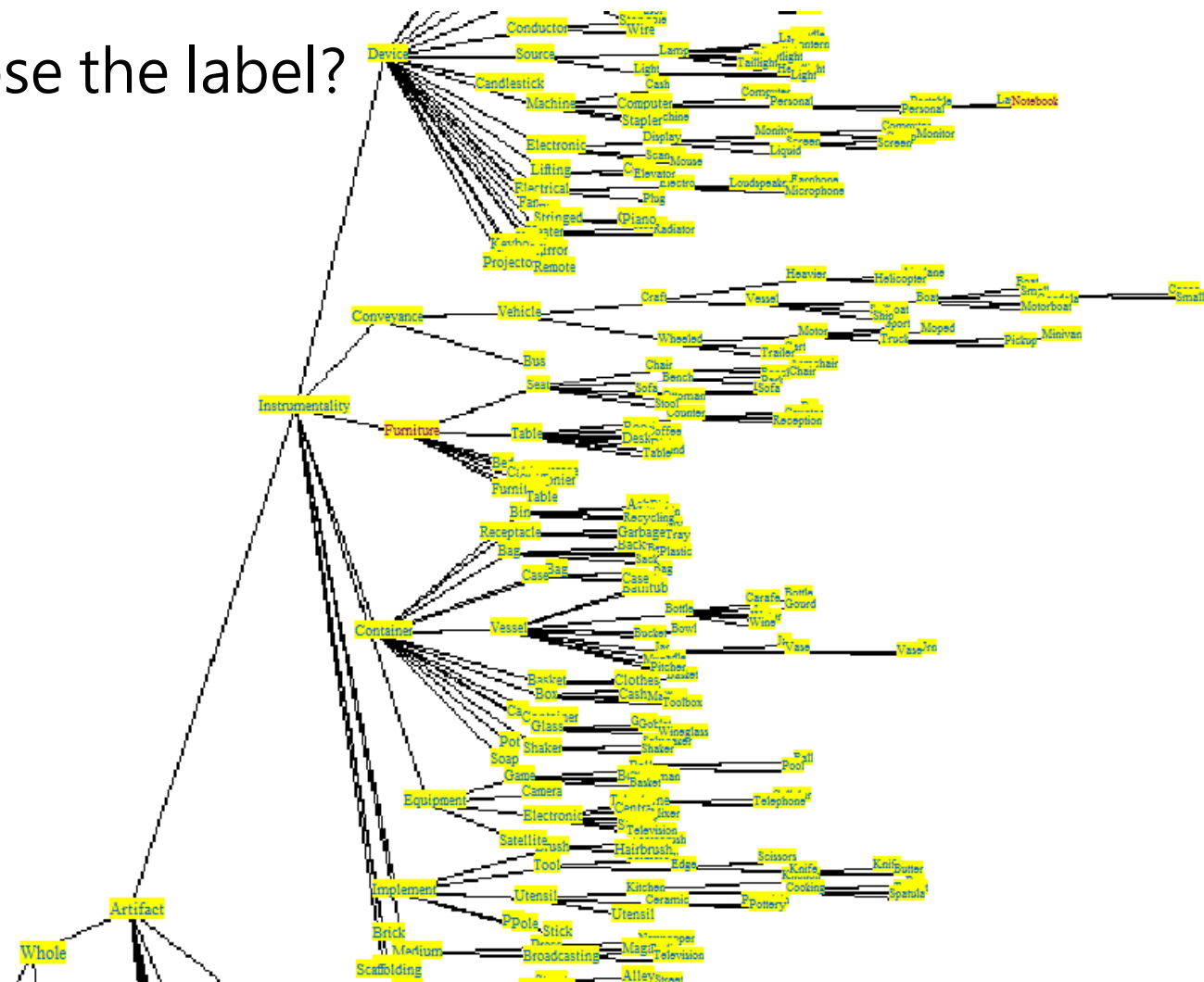


Label: car



WORDNET

- How to choose the label?

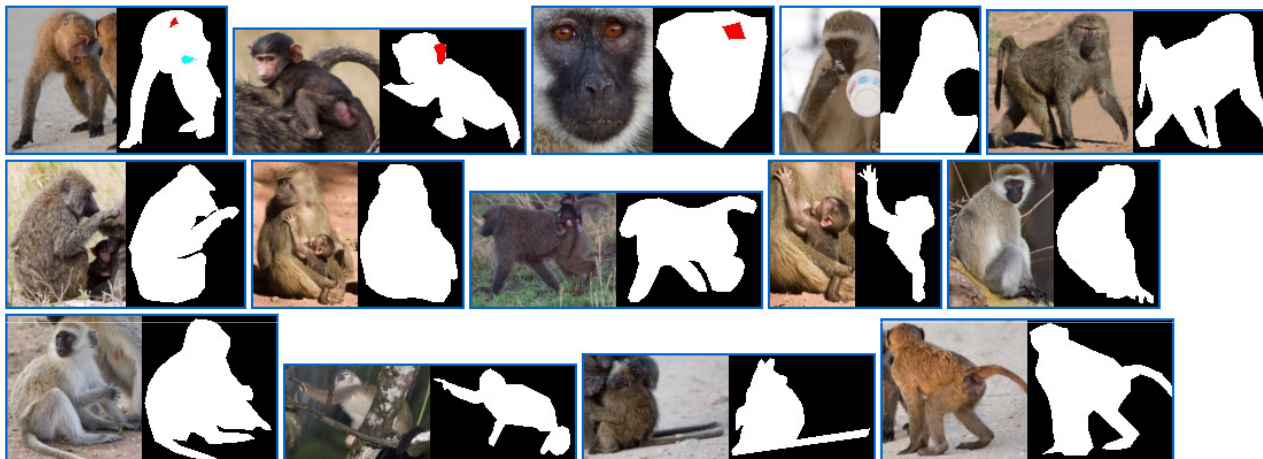




[Labeling tool](#) [Help](#) [Downloads](#) ^{New!} [About](#) [Matlab Toolbox](#) [Upload Images](#) [Stats](#)

key

Your query (key) matches **68** images.
Parts within each object are shown in color (only when labeled). See [summary statistics](#) for all the objects in the dataset, and the [wordnet hierarchy](#).



[Labeling tool](#) [Help](#) [Downloads](#) ^{New!} [About](#) [Matlab Toolbox](#) [Upload Images](#) [Stats](#)

monk

Your query (**monk**) matches **0** images.
Parts within each object are shown in color (only when labeled). See [summary statistics](#) for all the objects in the dataset, and the [wordnet hierarchy](#).

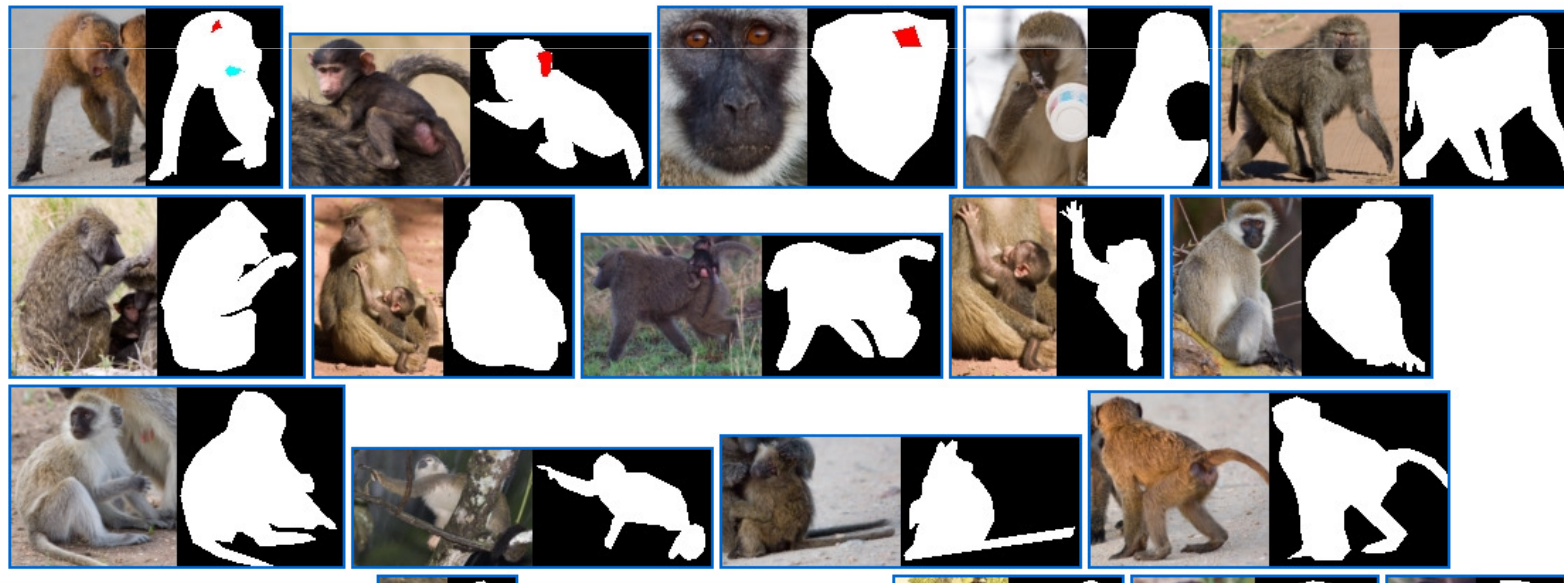


[Labeling tool](#) [Help](#) [Downloads](#) ^{New!} [About](#) [Matlab Toolbox](#) [Upload Images](#) [Stats](#)

onkey

Your query (onkey) matches 50 images.

Parts within each object are shown in color (only when labeled). See [summary statistics](#) for all the objects in the dataset, and the [wordnet hierarchy](#).

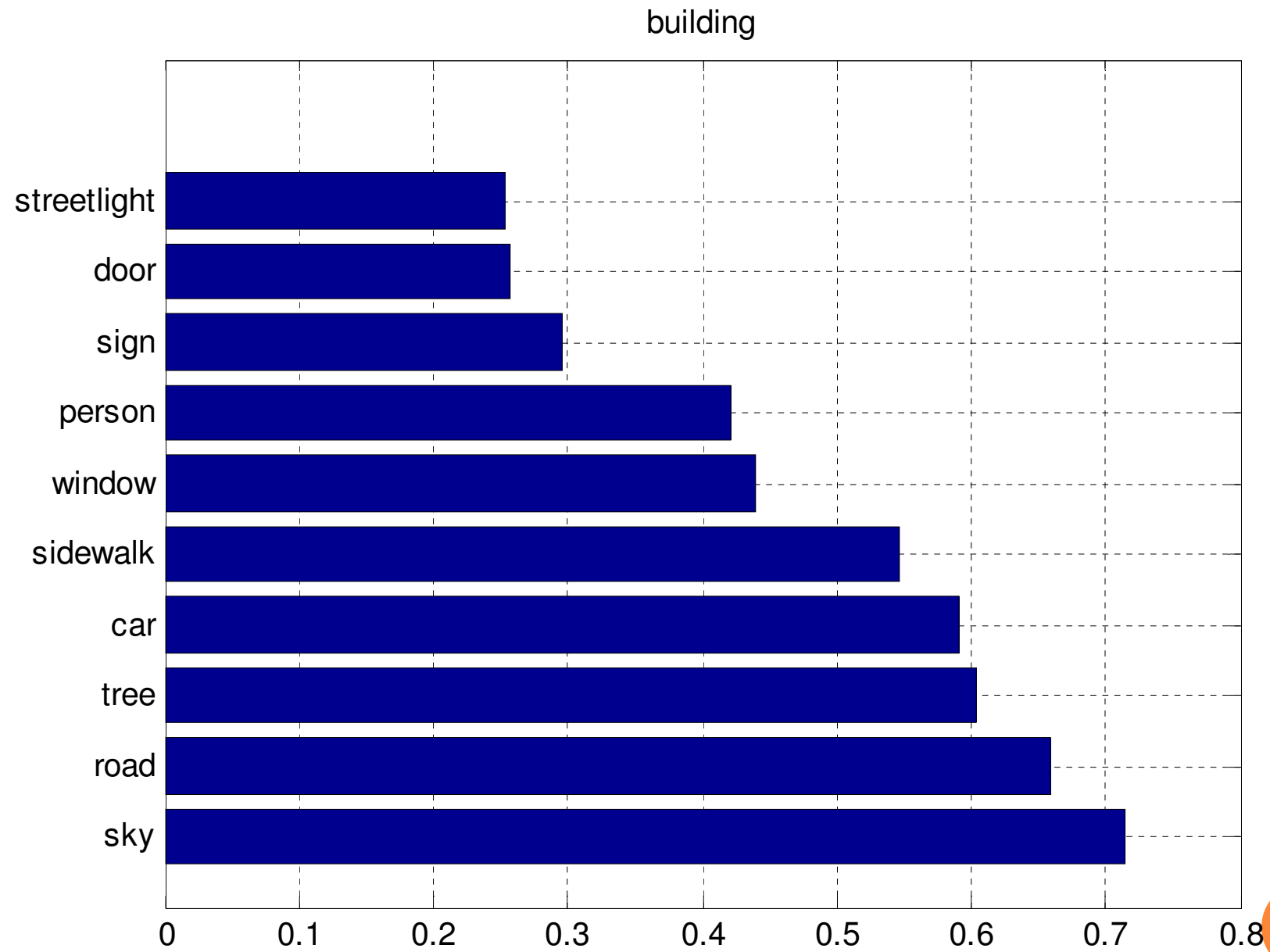


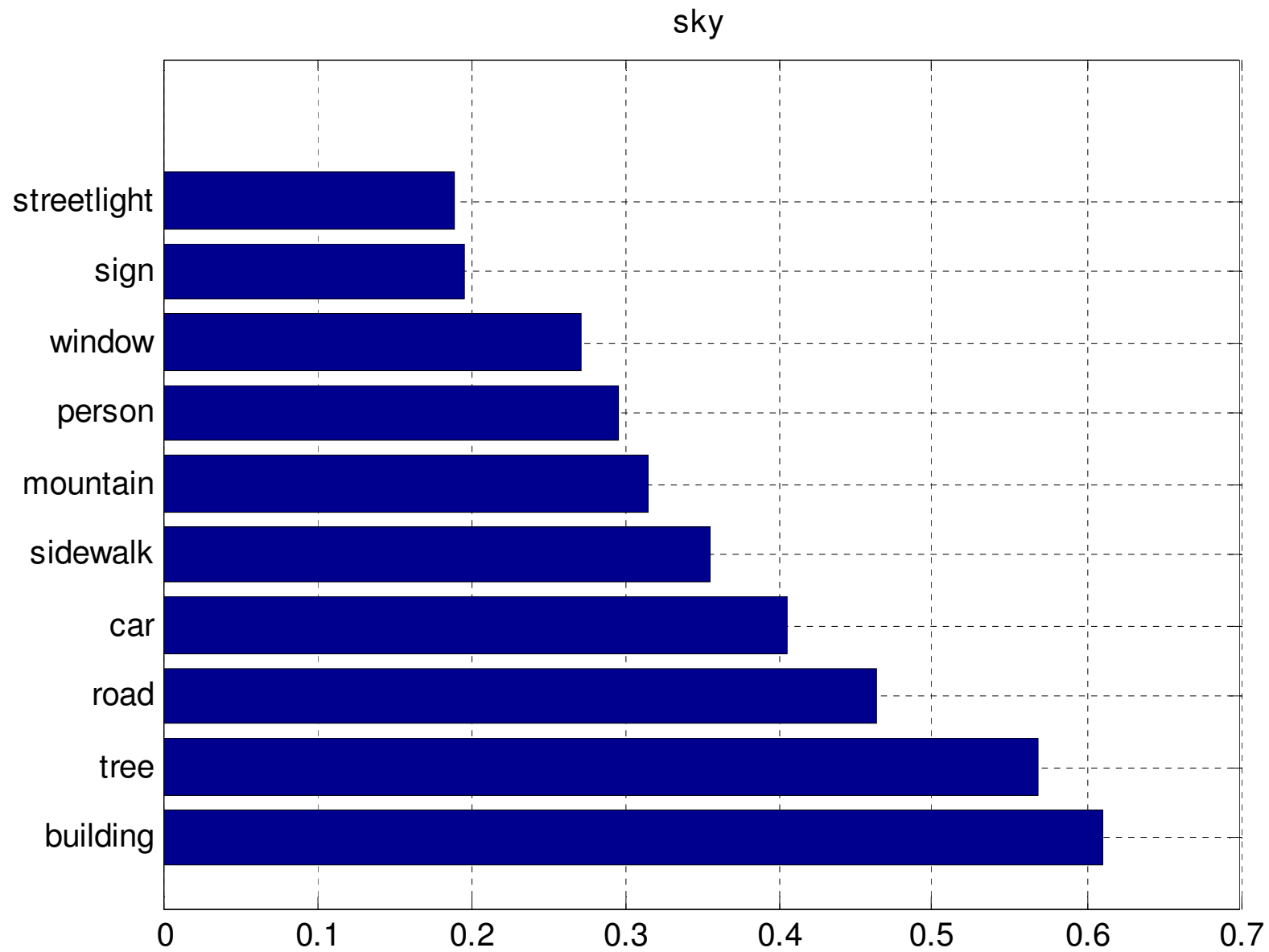
EXPERIMENTS IN THE PAPER

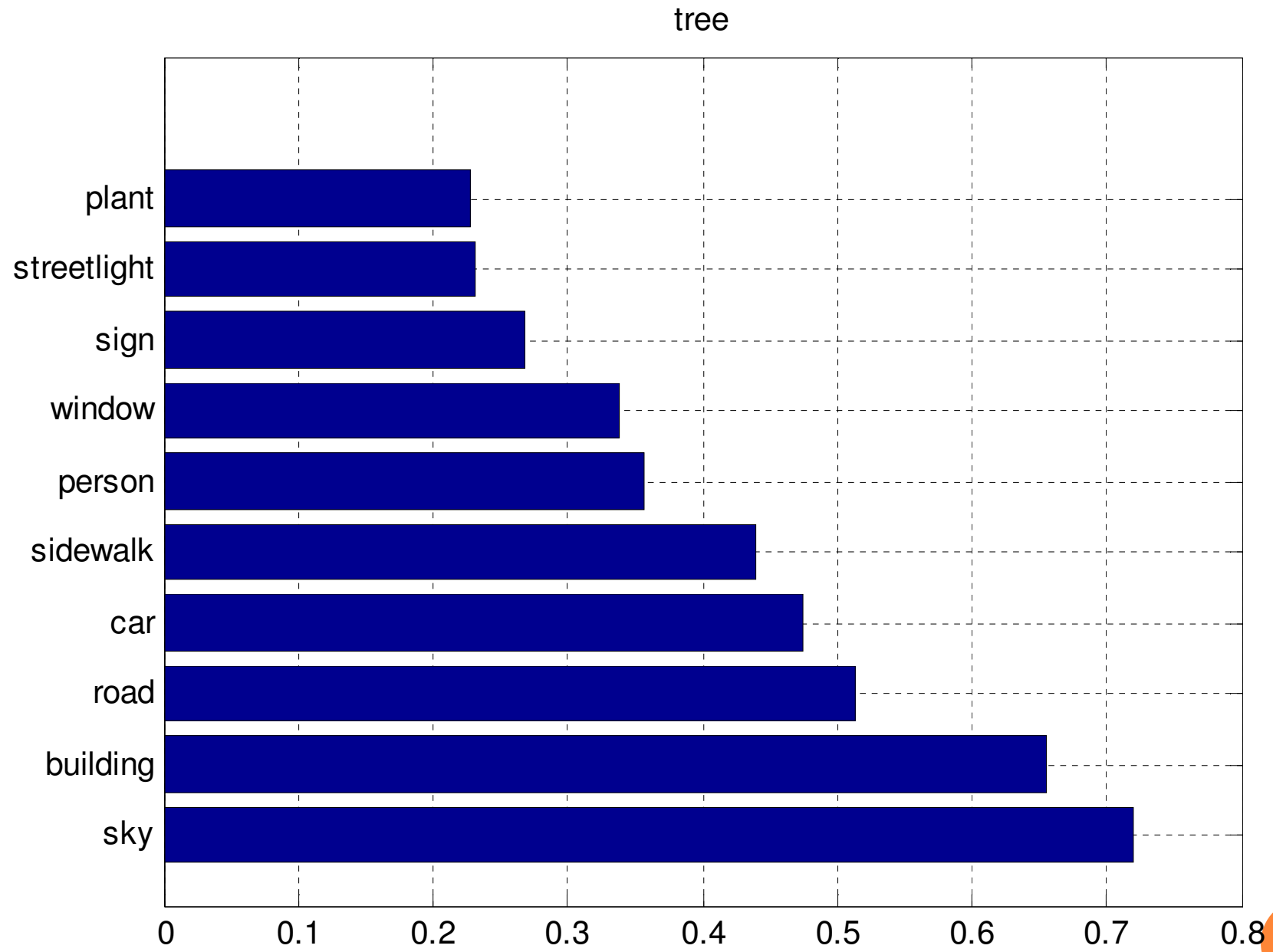
- percentage of pixels labeled per image.
- the number of labeled objects per image.
- the average number of control points clicked for each category.
- Distributions of object location showing where in the image each instance occurs
- object sizes, relative to the image size showing what is the typical size that the object has in the LabelMe dataset.
- how many parts an object has
- Depth ordering

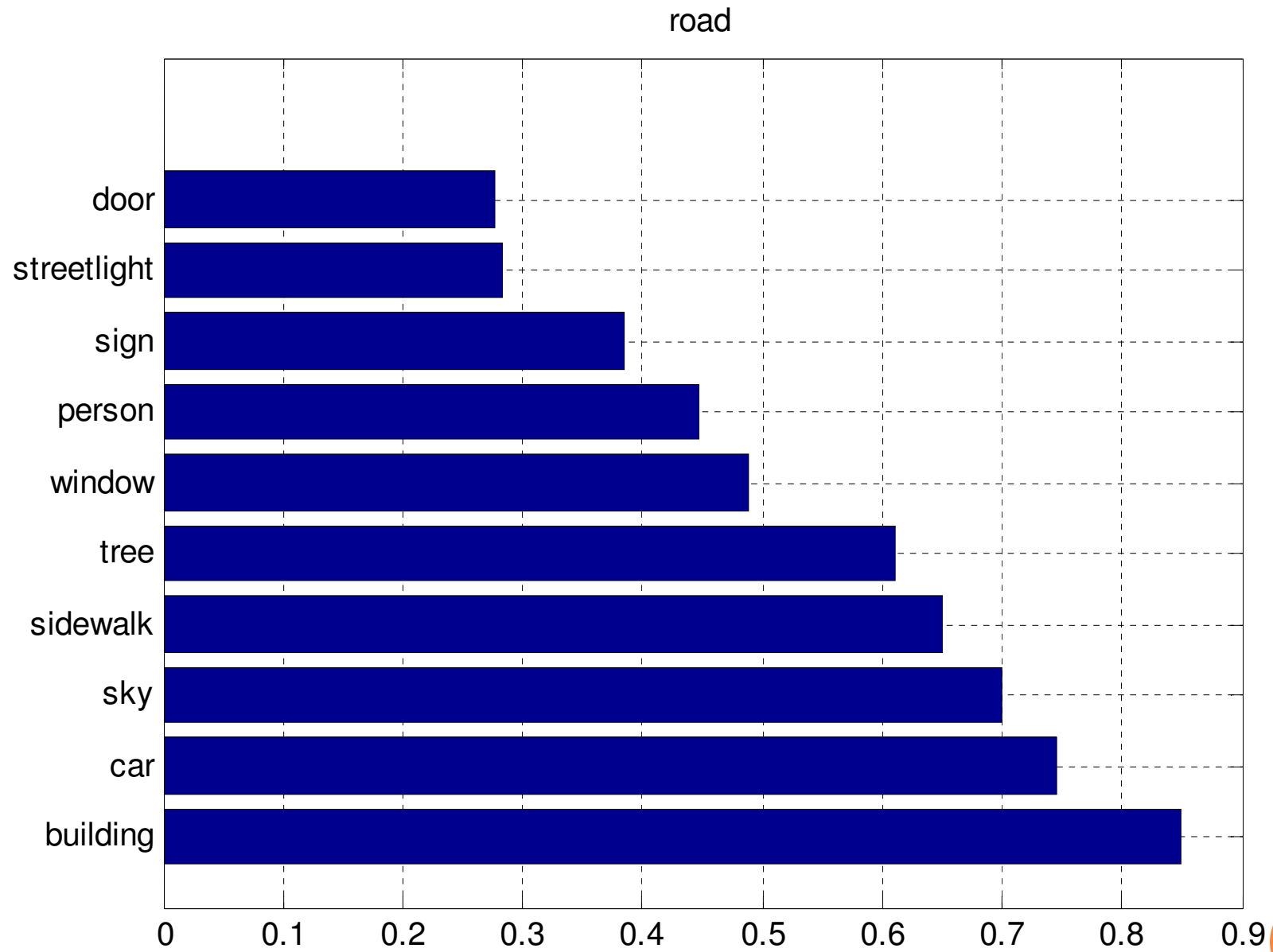
CORRELATION-TOP PAIRS

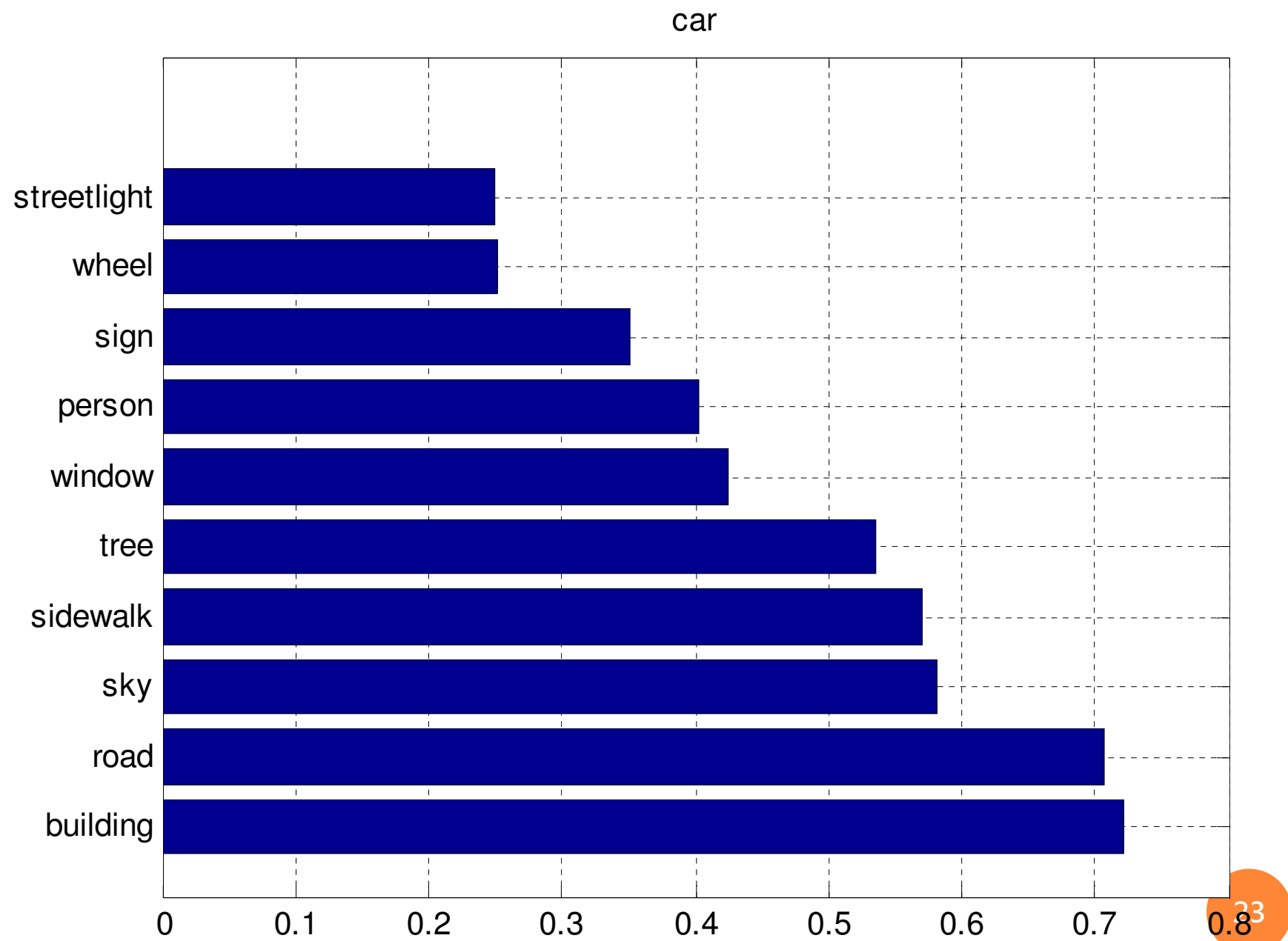
1	'building'	'sky'	1234
2	'sky'	'tree'	1147
3	'building'	'road'	1137
4	'building'	'tree'	1044
5	'building'	'car'	1019
6	'car'	'road'	998
7	'building'	'sidewalk'	944
8	'road'	'sky'	937
9	'road'	'sidewalk'	870
10	'car'	'sky'	821
11	'road'	'tree'	819
12	'car'	'sidewalk'	804
13	'building'	'window'	759
14	'car'	'tree'	756
15	'building'	'person'	726
16	'sidewalk'	'sky'	717
17	'sidewalk'	'tree'	700
18	'road'	'window'	654
19	'mountain'	'sky'	636
20	'car'	'window'	600

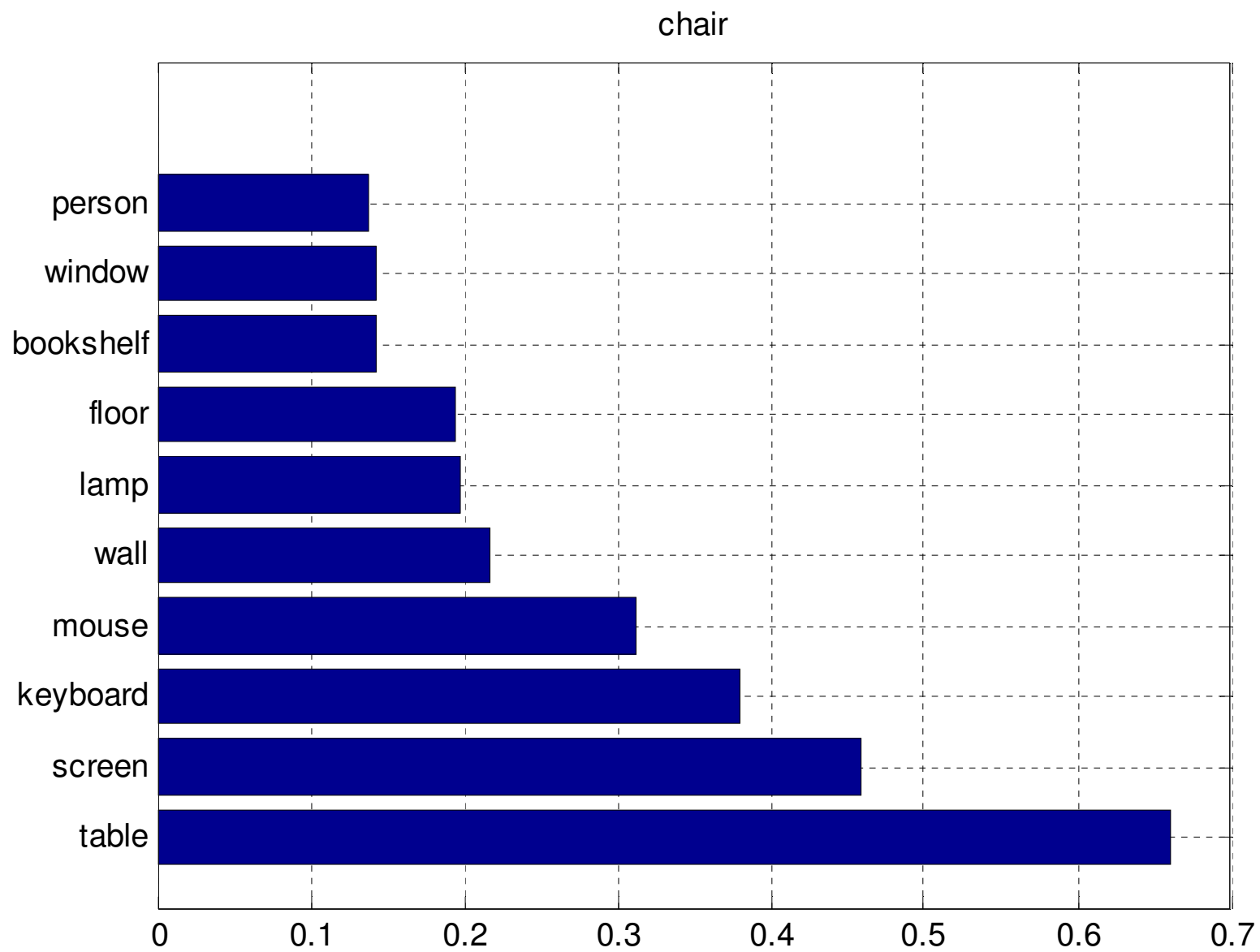


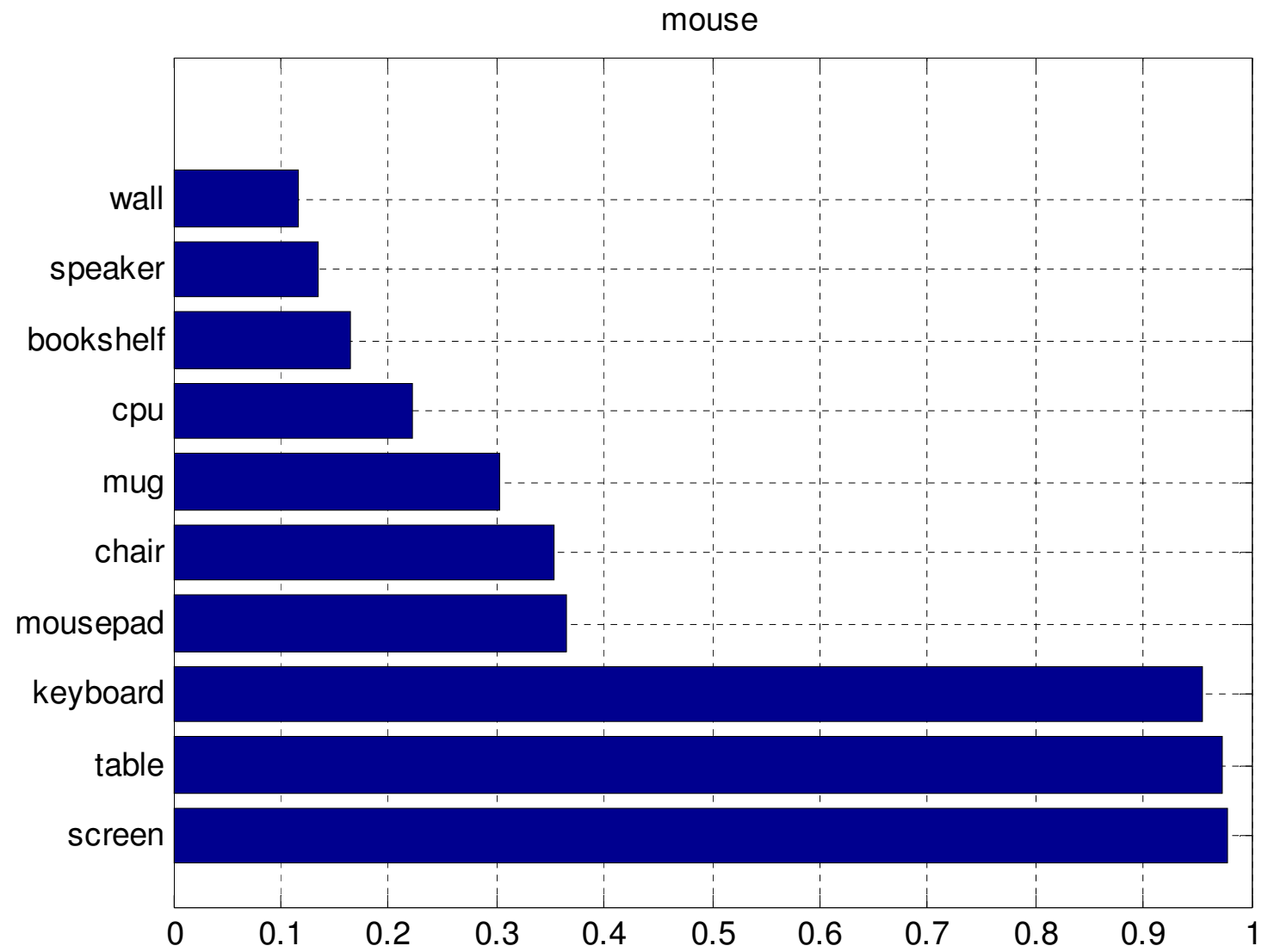












LEAST FREQUENT PAIRS

1	'ac unit'	'air conditioning'	1
2	'airplane'	'airport'	1
3	'another one'	'arbitrary'	1
4	'arcade'	'arcade occluded'	1
5	'airplane'	'arch'	1
6	'arch'	'arch crop'	1
7	'airplane'	'arm'	1
8	'airport cart'	'arm'	1
9	'arcade'	'arm'	1
10	'another brick'	'arrow'	1
11	'arch'	'arrow'	1
12	'arm'	'arrow'	1
13	'arm'	'art'	1
14	'arm'	'asm'	1
15	'air conditioning'	'attic'	1
16	'arm'	'attraction'	1
17	'aerial'	'avenue'	1
18	'baby carriage'	'bag'	1
19	'backsplash'	'bag'	1
20	'badge'	'bag'	1

HIGH VARIANCE OF CLICKED POINTS

	STD	Mean
'leaf'	101.63	100.75
'animal crop'	66.68	40.44
'lion'	52.78	74.00
'parrot'	37.42	32.67
'stair'	29.70	37.00
'stage'	28.29	21.33
'castle'	26.81	52.00
'lobster'	12.02	19.50
'deer'	11.87	12.00
'motorbikes'	11.56	21.56

LOW VARIANCE OF CLICKED POINTS

	STD	Mean
'land'	0.33	4.12
'label'	0.35	4.13
'paper'	0.39	4.18
'cigar'	0.40	4.19
'leaves'	0.71	4.50
'teeth'	0.71	5.50
'trash'	0.71	6.50
'whiteboard'	1.60	4.88
'furniture occluded'	1.69	4.93
'painting'	2.12	5.50

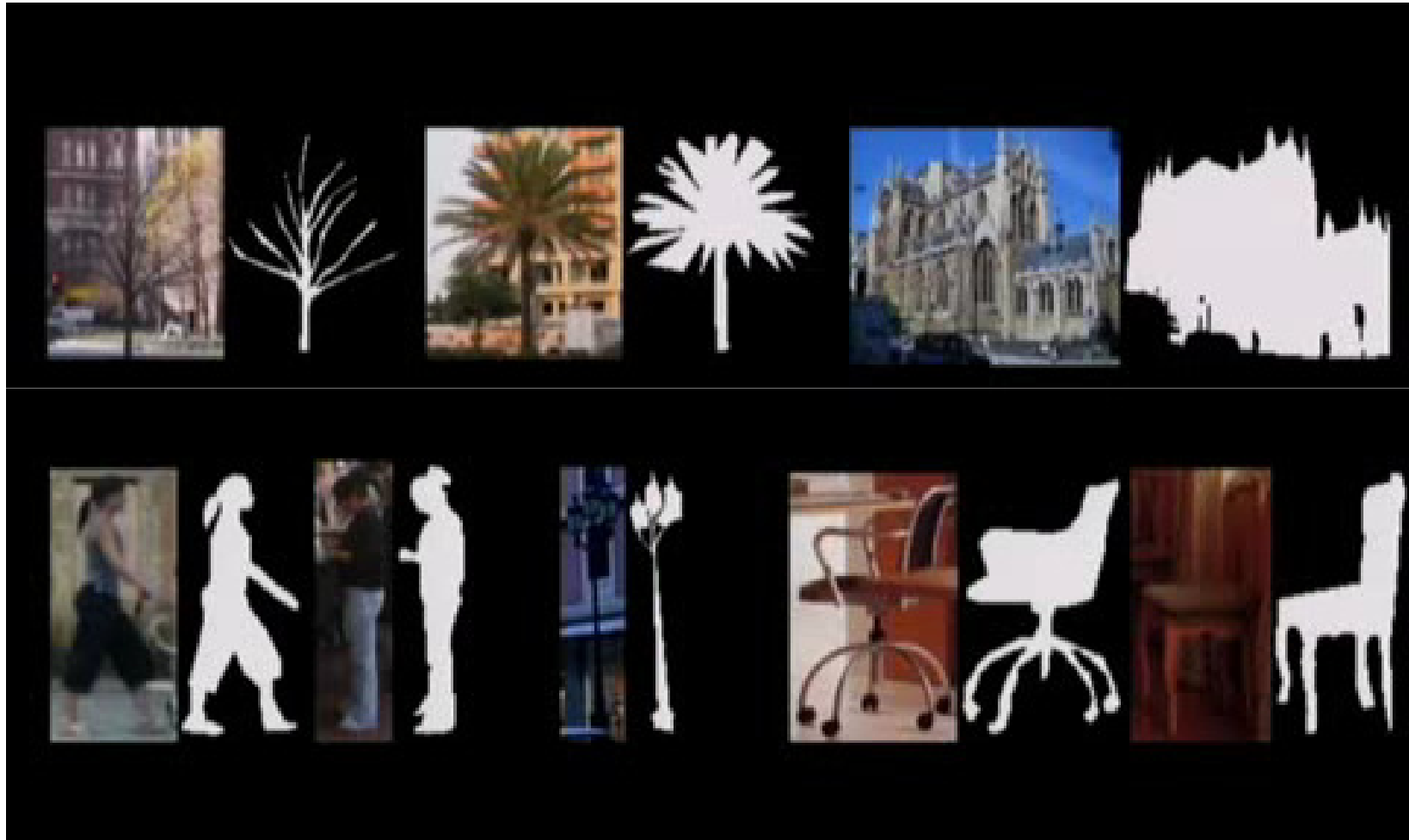


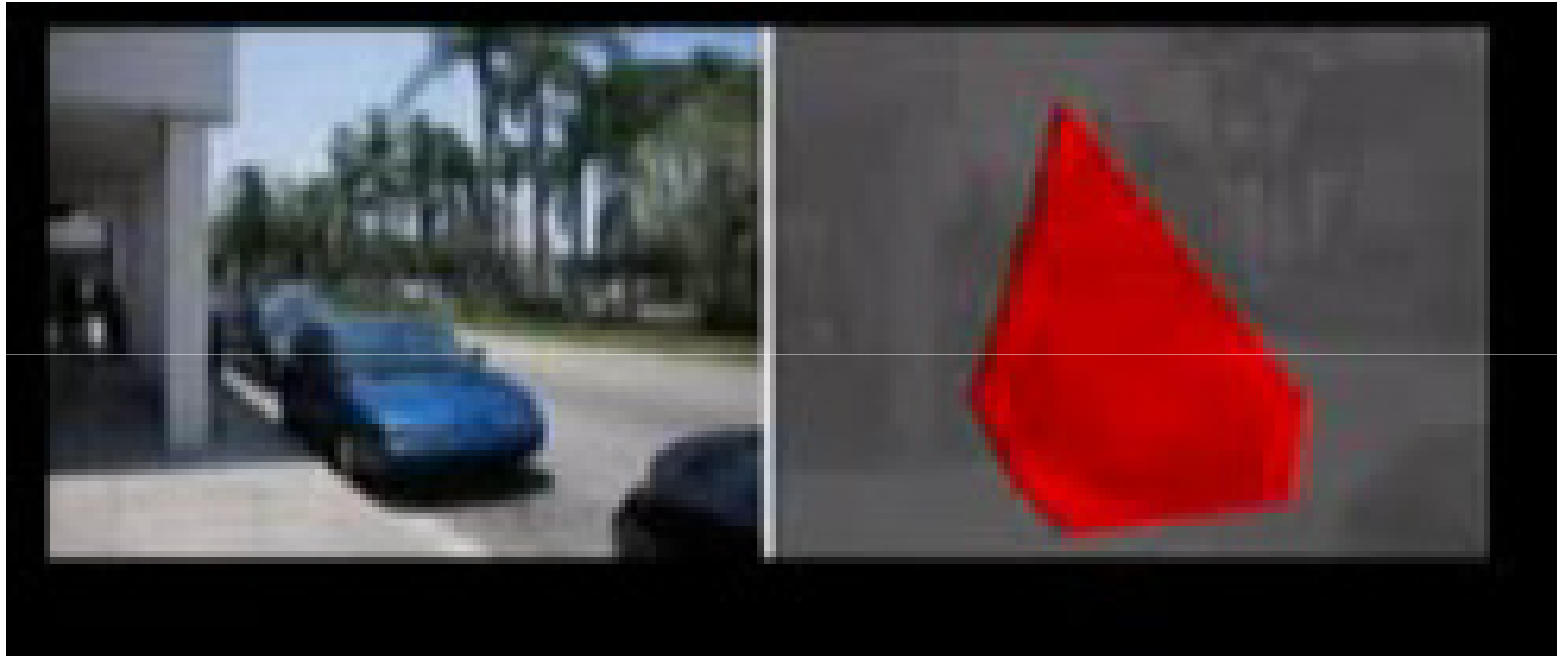
VARIANCE OF COLOR DISTRIBUTION

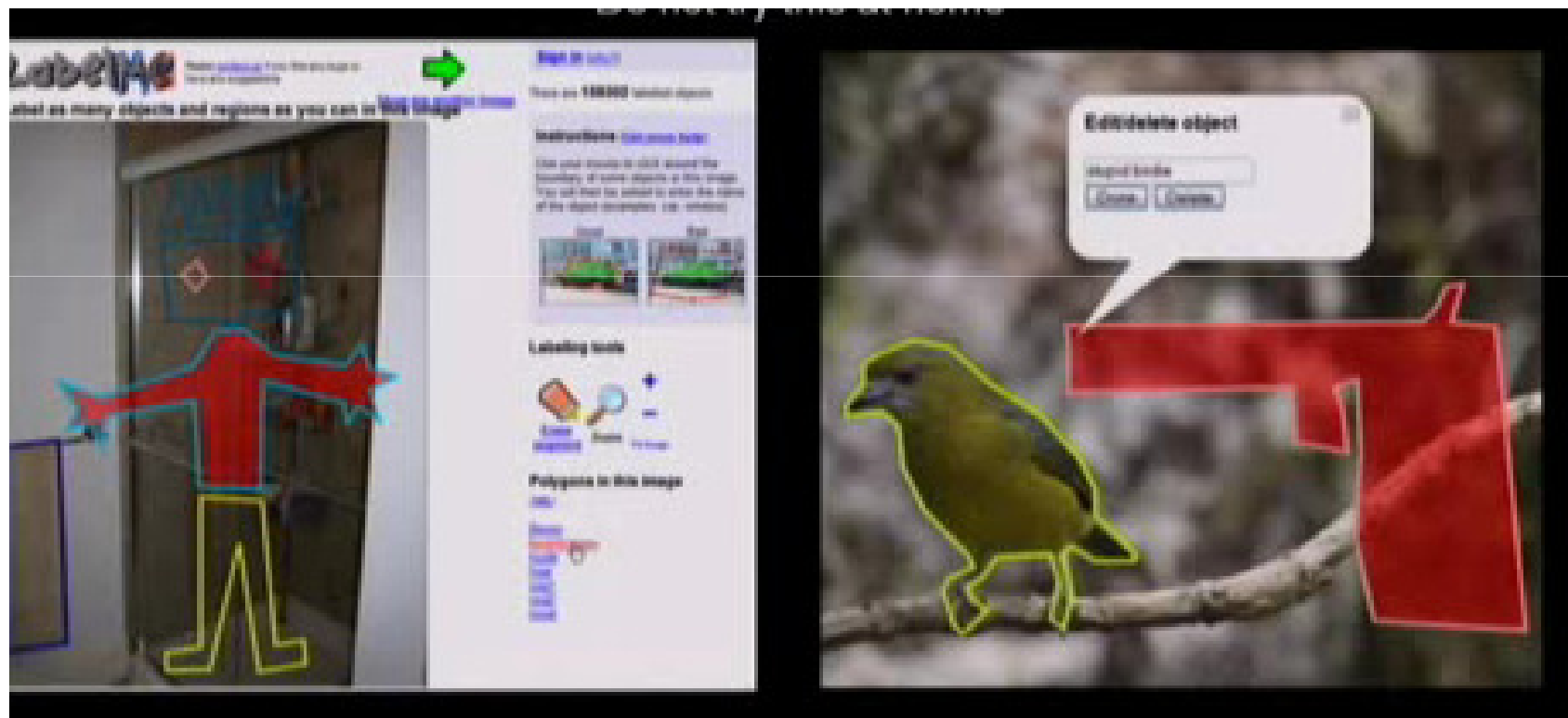
Variance	
apple	0.2918
arm	0.3133
tree	0.3261
cherry	0.3714
sky	0.4532
person	0.526
book	0.5471
car	0.5848
window	0.6992
animal	0.7399
fruit	0.7819

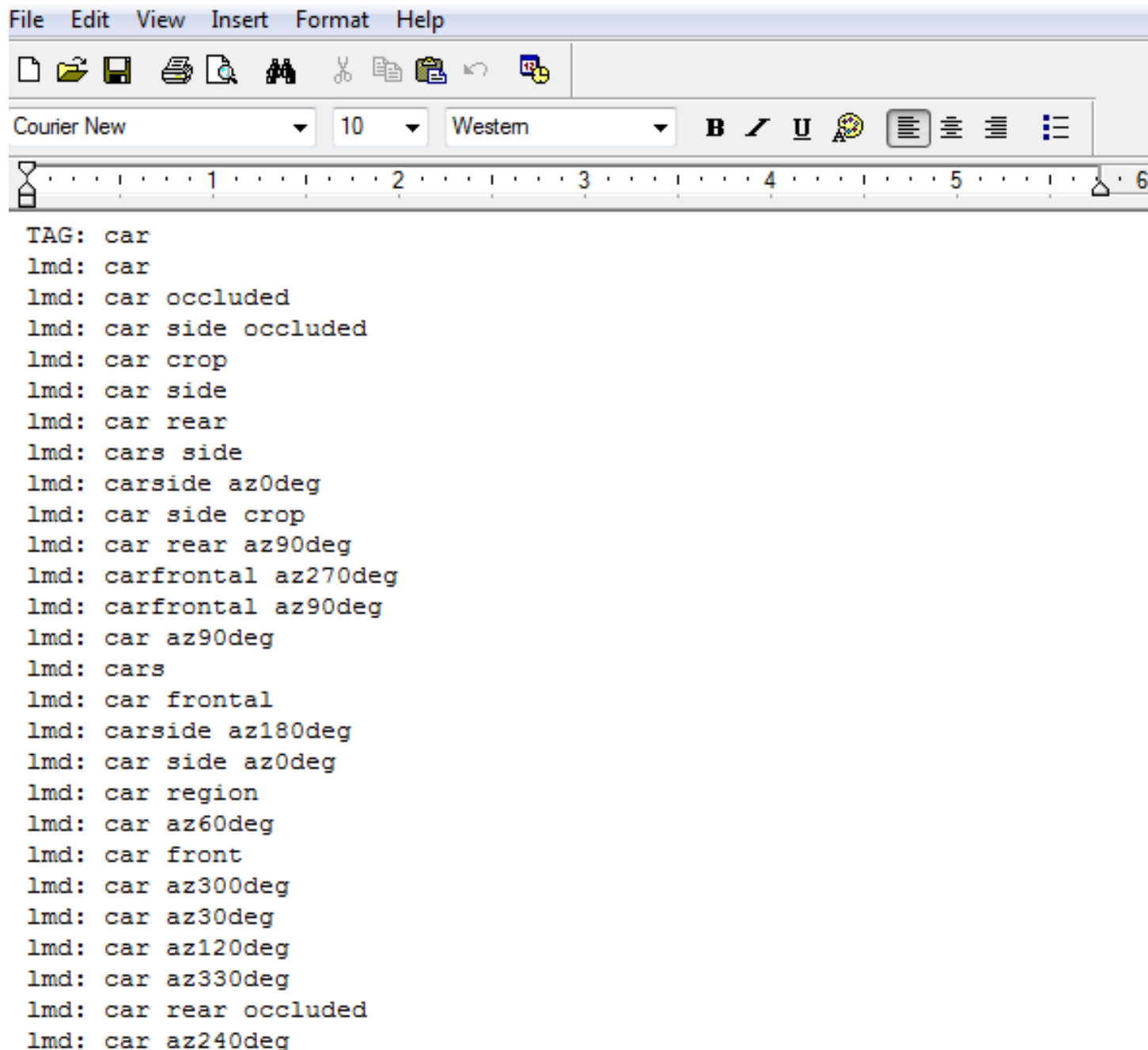
DISCUSSION

- Quality control
- Text label itself
- Statics pictures & Sequence pictures
- dataset









The image shows a screenshot of a text editor window. The menu bar at the top includes 'File', 'Edit', 'View', 'Insert', 'Format', and 'Help'. Below the menu bar is a toolbar with icons for file operations (new, open, save, print, find, replace, copy, paste, undo, redo) and text formatting (bold, italic, underline, strikethrough, bulleted list, numbered list, indent, outdent). The font settings are set to 'Courier New', size '10', and style 'Western'. A horizontal ruler is visible below the font settings, with markers from 1 to 6. The main text area contains a list of tags and lmd values, each on a new line. The tags are: TAG: car, lmd: car, lmd: car occluded, lmd: car side occluded, lmd: car crop, lmd: car side, lmd: car rear, lmd: cars side, lmd: carside az0deg, lmd: car side crop, lmd: car rear az90deg, lmd: carfrontal az270deg, lmd: carfrontal az90deg, lmd: car az90deg, lmd: cars, lmd: car frontal, lmd: carside az180deg, lmd: car side az0deg, lmd: car region, lmd: car az60deg, lmd: car front, lmd: car az300deg, lmd: car az30deg, lmd: car az120deg, lmd: car az330deg, lmd: car rear occluded, and lmd: car az240deg.

```
TAG: car
lmd: car
lmd: car occluded
lmd: car side occluded
lmd: car crop
lmd: car side
lmd: car rear
lmd: cars side
lmd: carside az0deg
lmd: car side crop
lmd: car rear az90deg
lmd: carfrontal az270deg
lmd: carfrontal az90deg
lmd: car az90deg
lmd: cars
lmd: car frontal
lmd: carside az180deg
lmd: car side az0deg
lmd: car region
lmd: car az60deg
lmd: car front
lmd: car az300deg
lmd: car az30deg
lmd: car az120deg
lmd: car az330deg
lmd: car rear occluded
lmd: car az240deg
```

ESP

PLAYER 1



GUESSING: CAR

GUESSING: HAT

GUESSING: KID

SUCCESS!
YOU AGREE ON CAR

PLAYER 2



GUESSING: BOY

GUESSING: CAR

SUCCESS!
YOU AGREE ON CAR

GOOGLE IMAGE LABELER

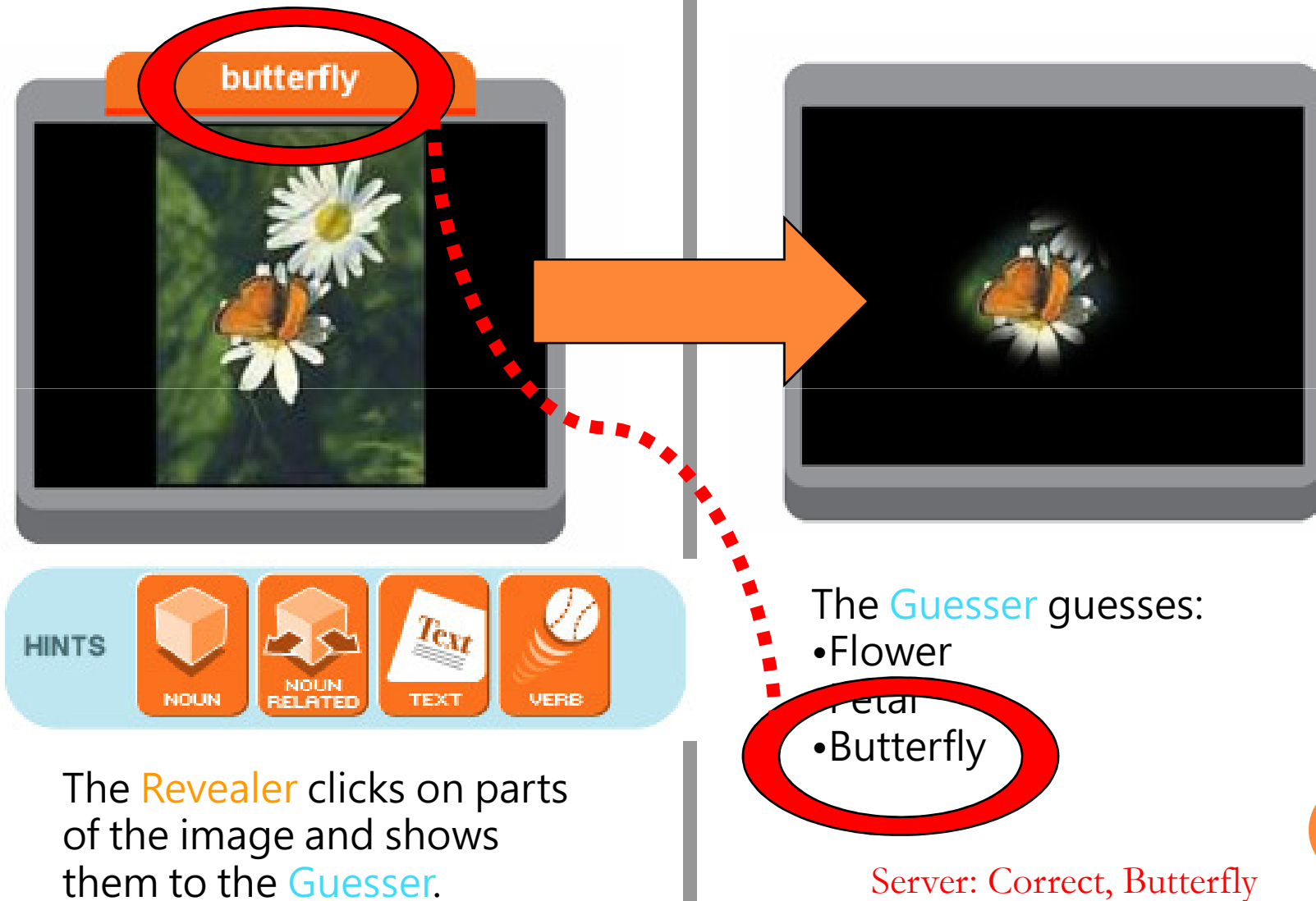
- The ESP Game has been licensed by Google.



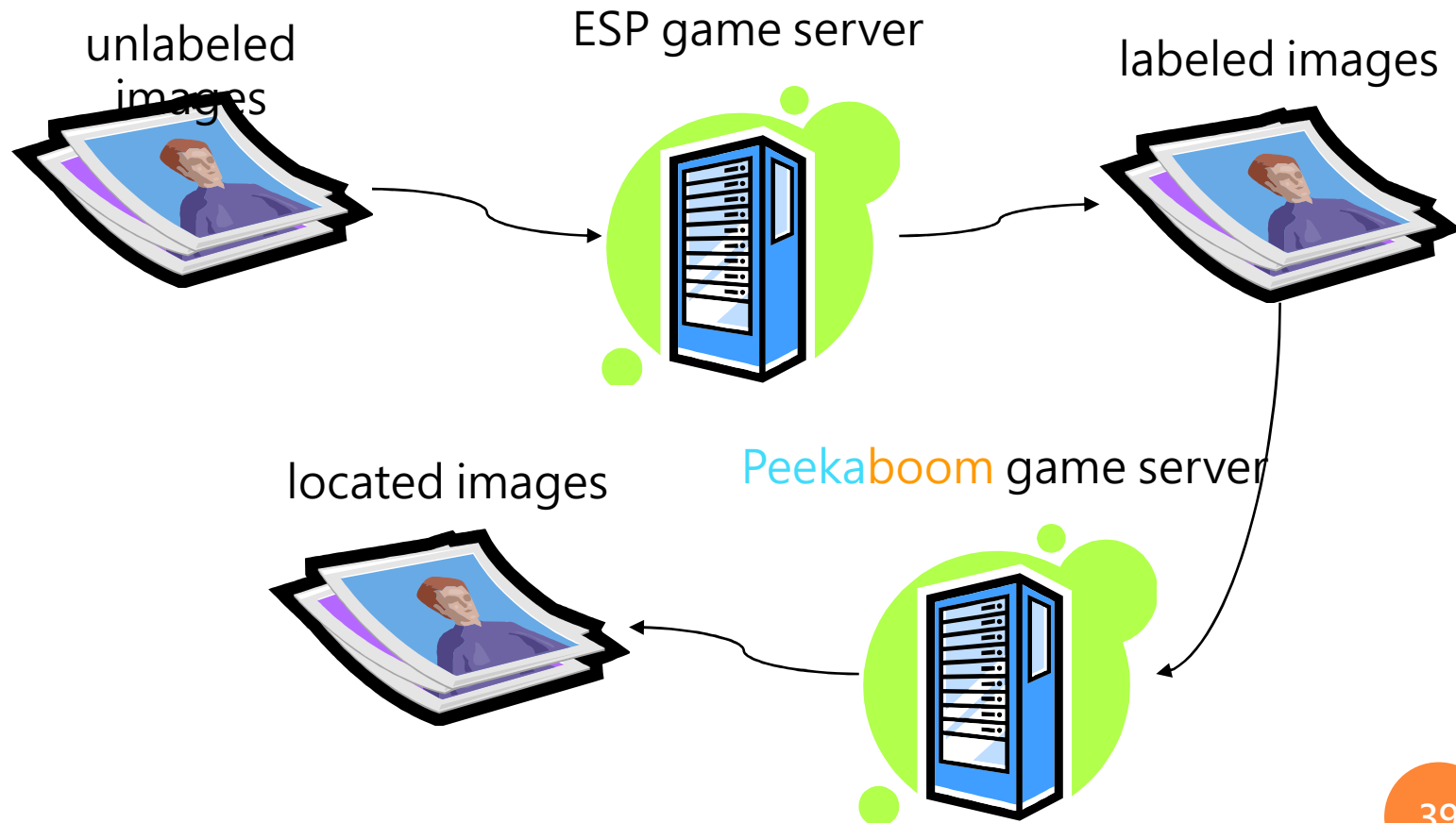
THE LIMITATIONS OF ESP

- The ESP Game can label images, but it cannot:
 - Find the objects being labeled.
 - Determine the way in which the object appears – does the label “car” refer to the text “car” or an actual car in the image?



PEEKABOOM



COMPLETING THE IMAGE CYCLE



HINTS

man	baseball	man	swing
			
 NOUN	 RELATED NOUN	 TEXT	 VERB

HINTS

