

## **Transfer-learning with object detection**, summary statistics and snapshots. by Nichole King

Single-shot detector (SSD) trained on images of Android lawn statues using the bounding box detector and feature detector from pre-trained model configuration `ssd_resnet50_v1_fpn_640x640_coco17_tpu`.

There are 7 images containing a mix of android lawn statues.

class id class name #of images w/ class in it

0	cupcake :	4
1	euclair:	4
2	icecream:	3
3	gingerbread_man:	5
4	icecream_sandwich:	3
5	honeycomb:	2
6	kitkat:	2
7	jellybean:	2
8	donut:	2

see notebook at:

[https://github.com/nking/curvature-scale-space-corners-and-transformations/blob/master/src/duckies\\_and\\_statues%2C\\_interactive\\_eager\\_few\\_shot\\_obj\\_det\\_training\\_colab.ipynb](https://github.com/nking/curvature-scale-space-corners-and-transformations/blob/master/src/duckies_and_statues%2C_interactive_eager_few_shot_obj_det_training_colab.ipynb)

## Transfer-learning with object detection (cont.)

14 test frames were taken from a YouTube video.

```
ground_truth_labels_all=[3, 9, 2, 2, 3, 9, 5, 2, 2, 2, 2, 2, 3, 2, 8, 8, 2, 8, 8, 3, 8, 9, 9, 3, 6, 3, 5, 8, 9, 8, 0, 0, 9, 4, 9, 9, 9, 9, 8, 0, 0, 4, 0, 4, 4, 9, 9, 0, 9, 0, 3, 0, 3, 7, 1, 6, 9, 1, 9, 6]
```

```
pred_all=[3, 7, 0, 0, 3, 7, 5, 1, 0, 2, 0, 2, 3, 1, 8, 8, 2, 1, 8, 3, 1, 4, 7, 5, 4, 3, 5, 8, 7, 1, 0, 0, 0, 7, 0, 0, 7, 0, 0, 7, 7, 7, 0, 8, 7, 5, 0, 0, 7, 0, 4, 7, 3, 7, 8, 2, 1, 7, 1, 1, 7, 6]
```

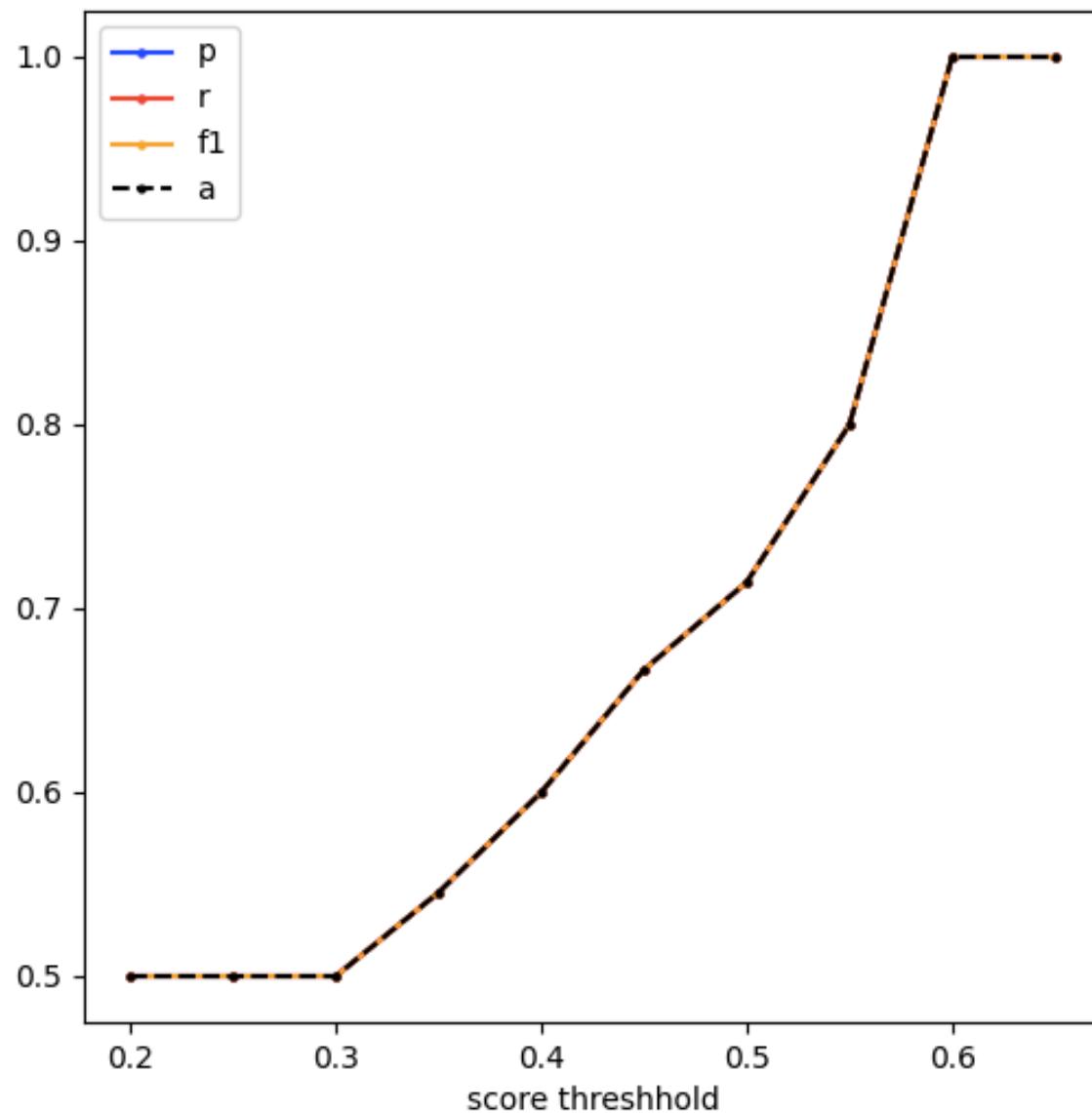
```
scores_all= ['0.29', '0.25', '0.55', '0.45', '0.39', '0.33', '0.32', '0.32', '0.30', '0.29', '0.25', '0.25', '0.53', '0.38', '0.29', '0.62', '0.32', '0.31', '0.66', '0.45', '0.43', '0.29', '0.28', '0.27', '0.32', '0.63', '0.47', '0.44', '0.42', '0.32', '0.29', '0.28', '0.26', '0.52', '0.48', '0.45', '0.37', '0.33', '0.30', '0.28', '0.28', '0.44', '0.34', '0.39', '0.36', '0.35', '0.33', '0.31', '0.31', '0.28', '0.27', '0.27', '0.61', '0.48', '0.35', '0.28', '0.35', '0.28', '0.50', '0.47', '0.38', '0.35']
```

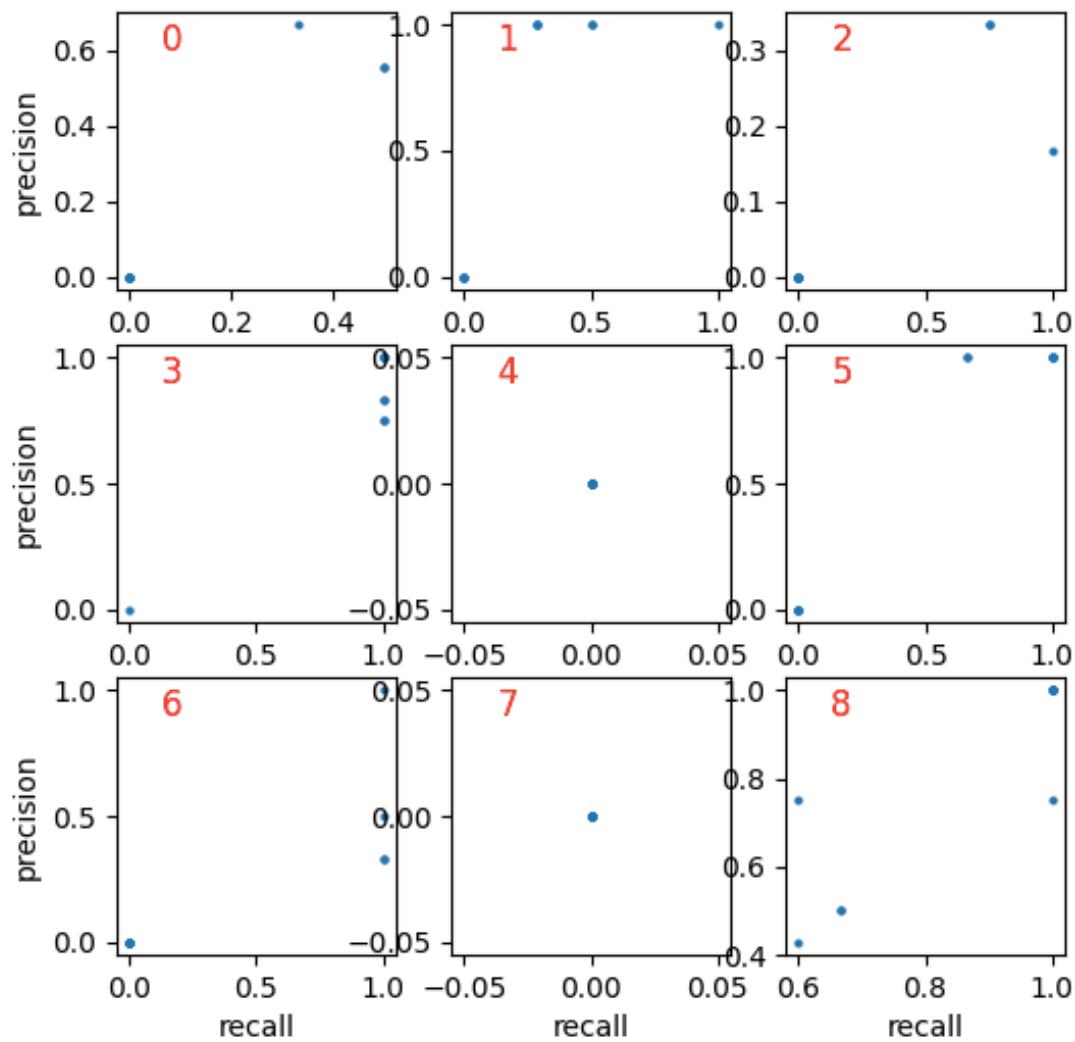
statistics were calculated for precision, recall, accuracy and f1 score and are plotted on the next 2 pages.

One can see that for identification threshold, we should set the score **threshold to 0.5 or 0.55**.

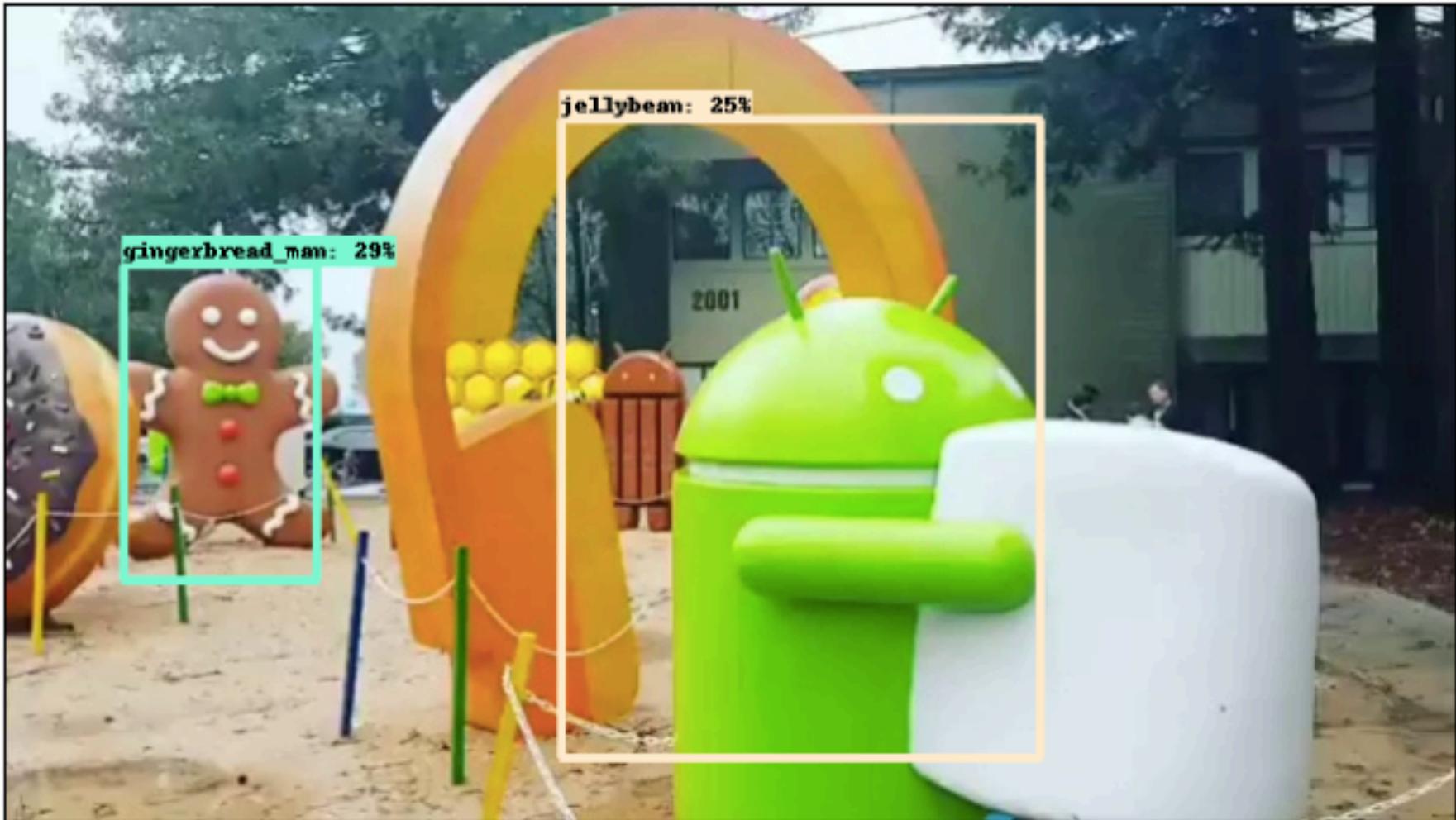
figures after the 2 plots are of the detection bounding boxes on 14 test frames labeled by their detection scores and labels. The detections plotted have scores from 20% to 100%.

these are the  
same for  
multiclass classification

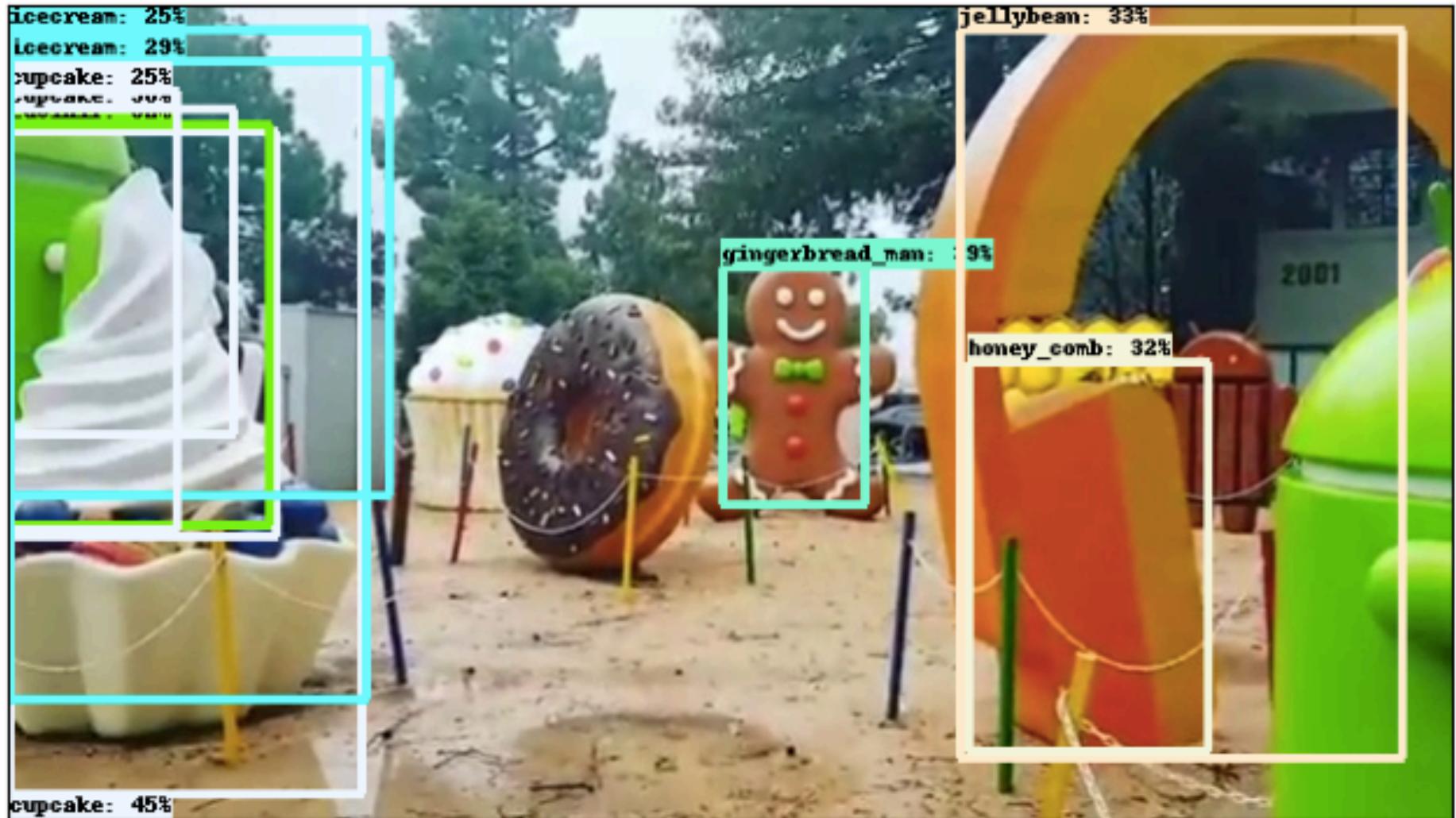




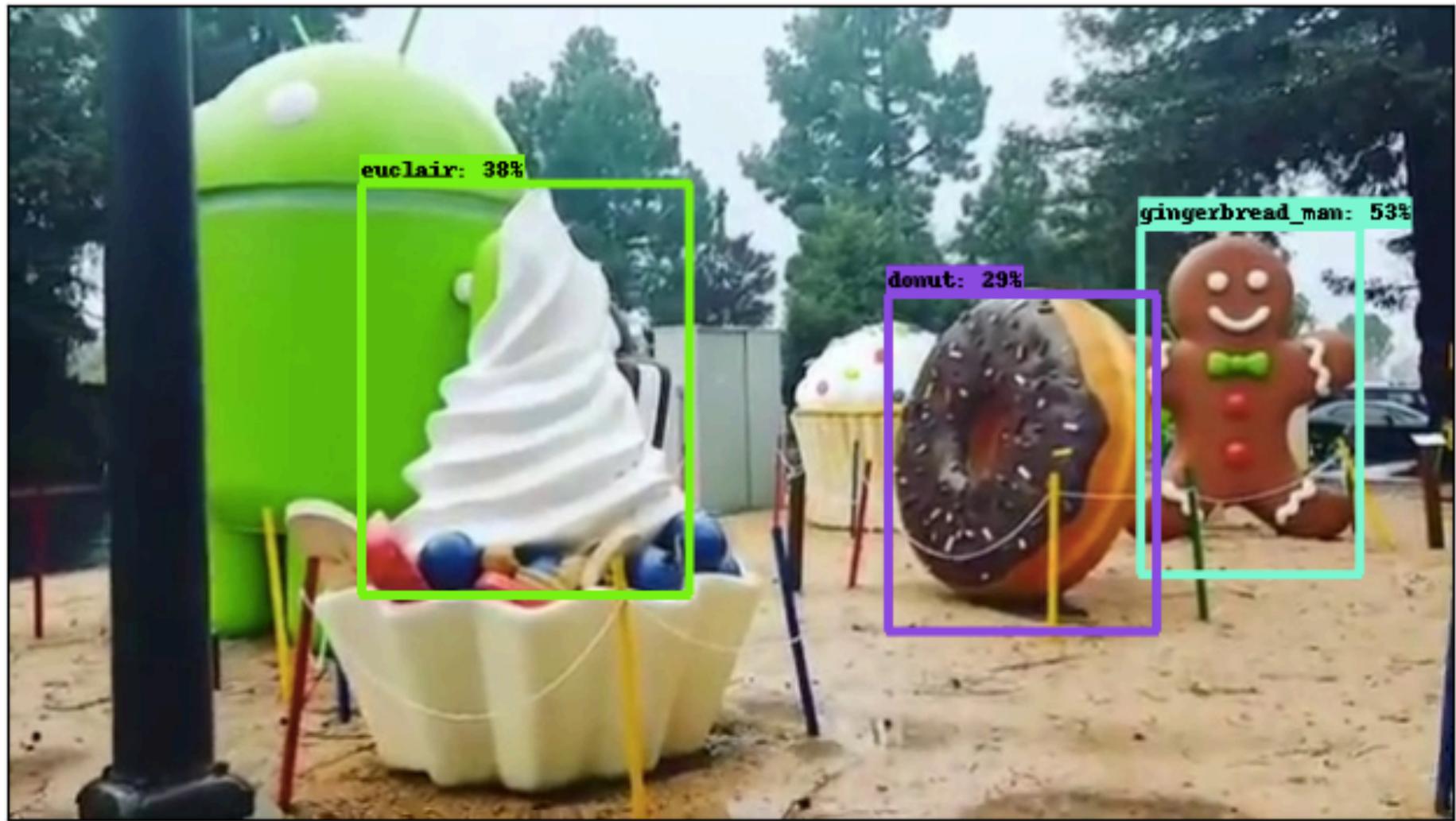
writing to /tmp/tmpveptyde1/1.jpg



writing to /tmp/tmpveptyde1/2.jpg



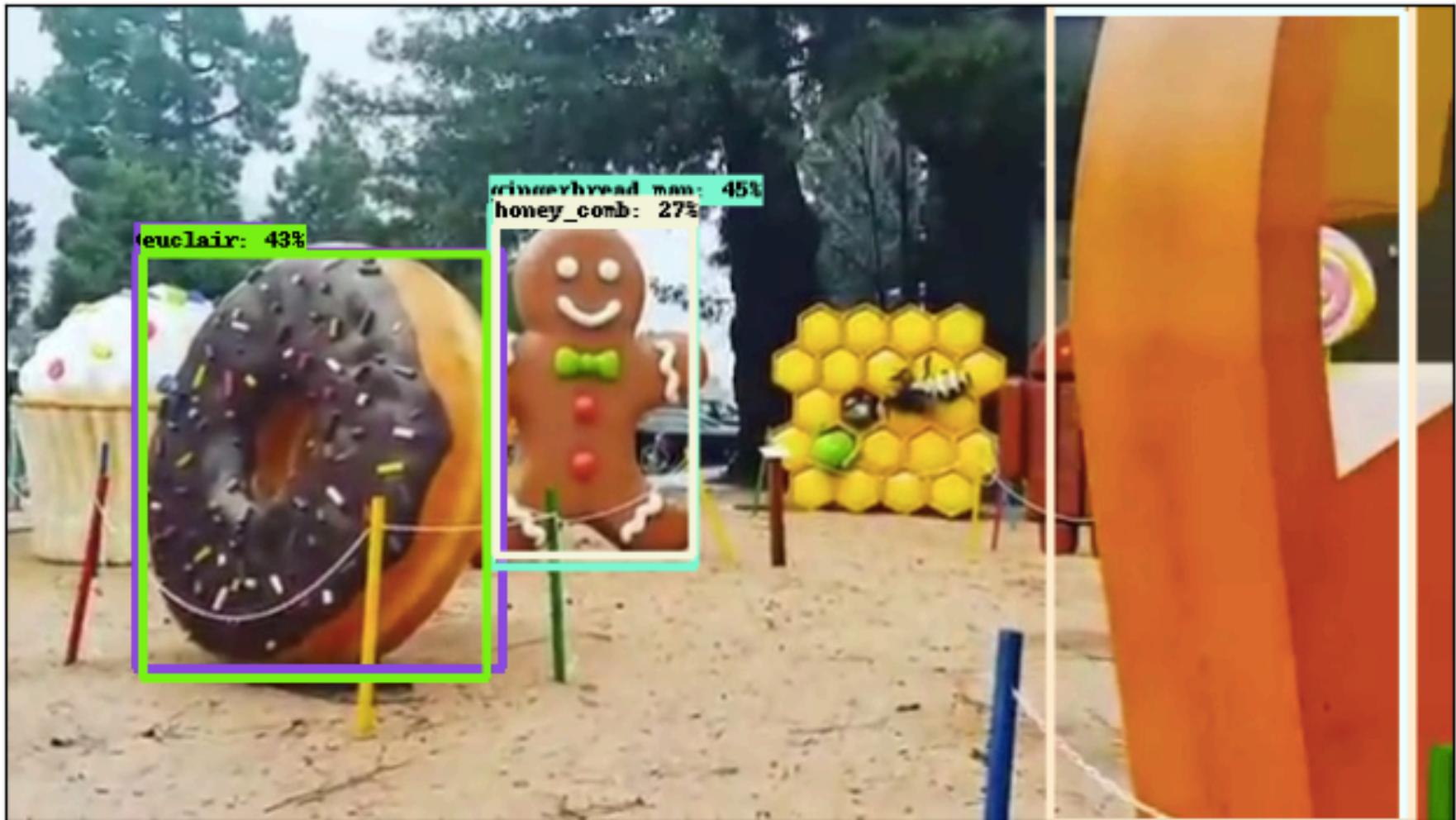
writing to /tmp/tmpveptyde1/3.jpg



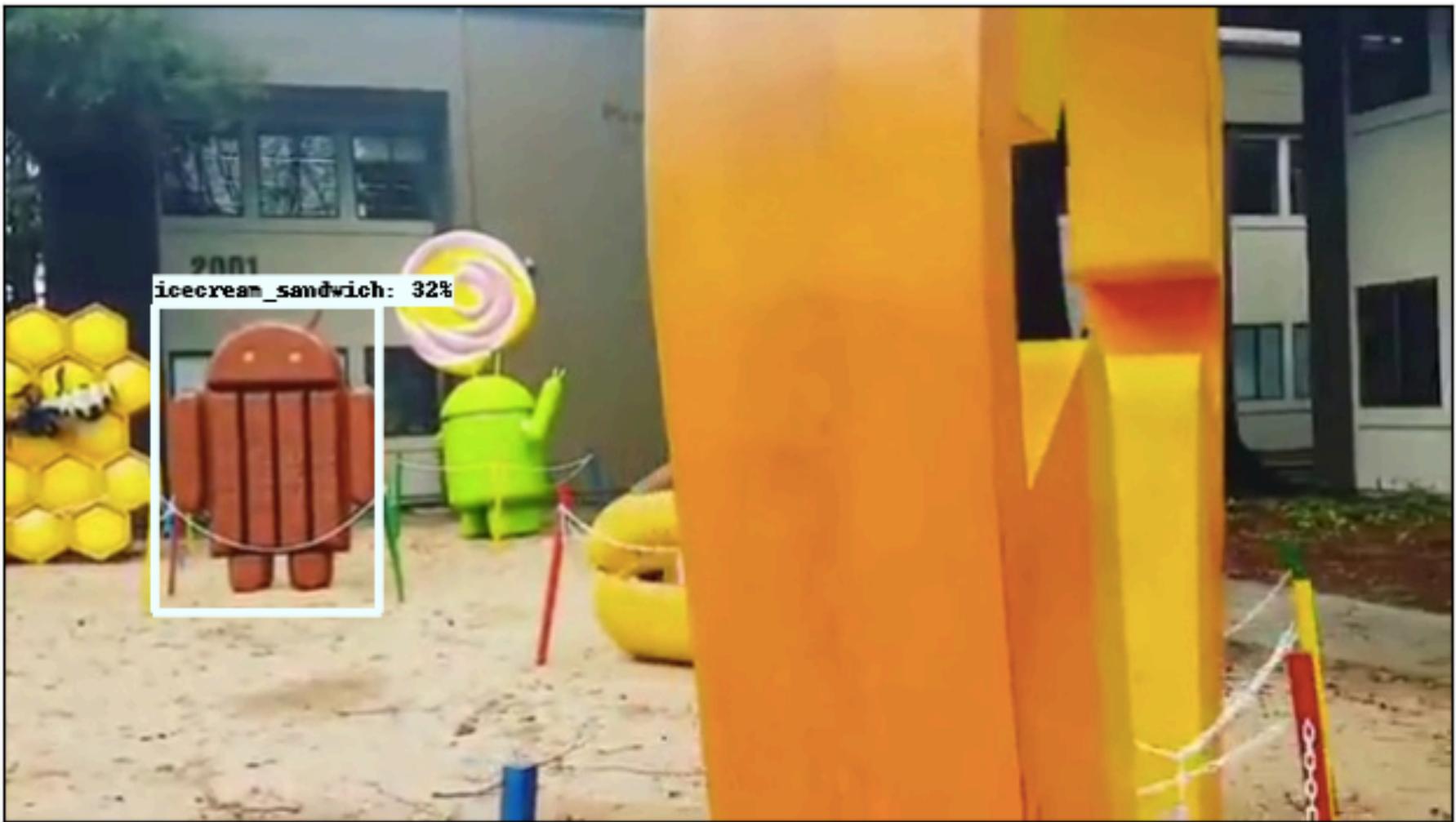
writing to /tmp/tmpveptyde1/4.jpg



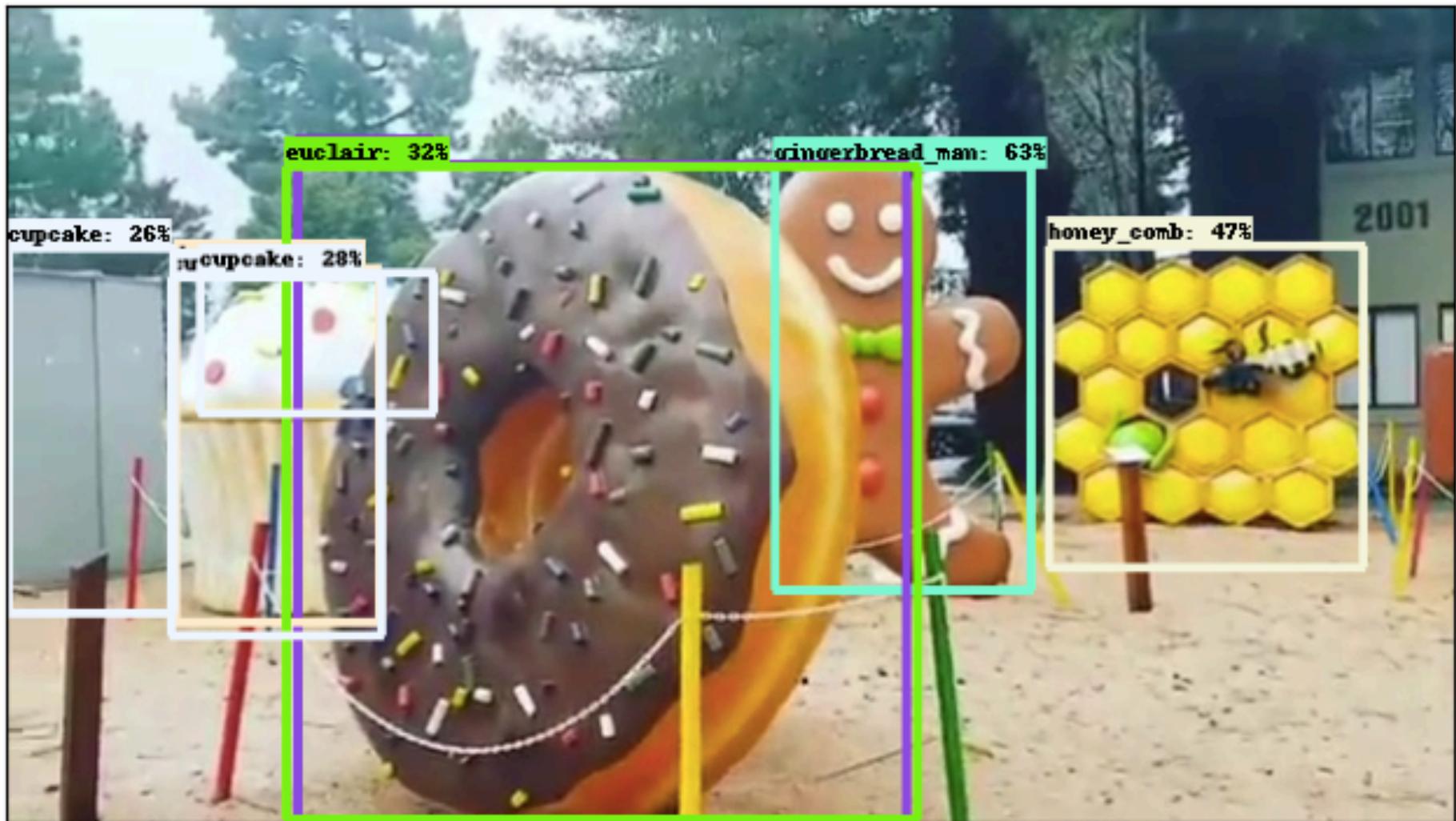
writing to /tmp/tmpveptyde1/5.jpg



writing to /tmp/tmpveptyde1/6.jpg



writing to /tmp/tmpveptyde1/7.jpg



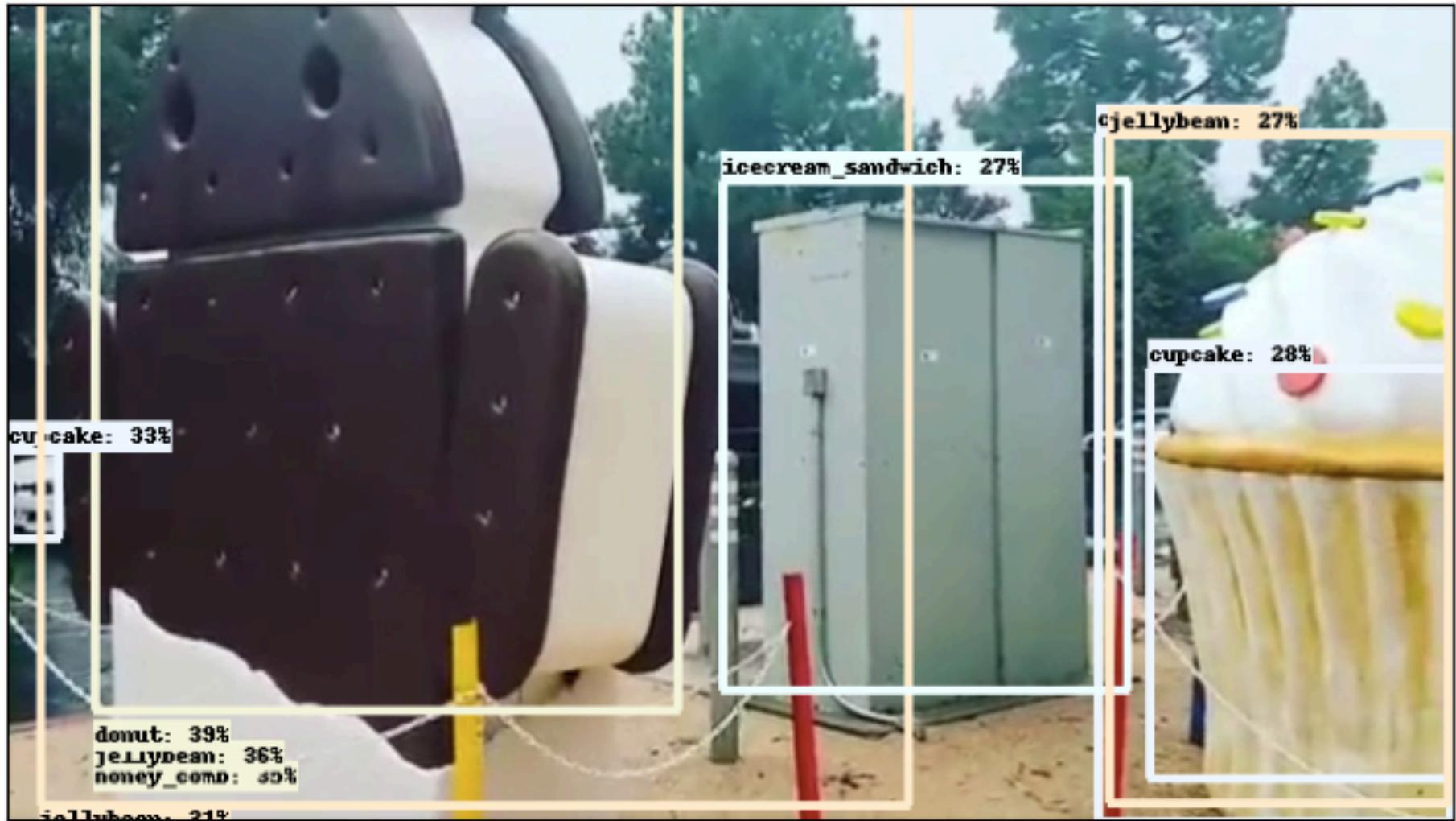
writing to /tmp/tmpveptyde1/8.jpg



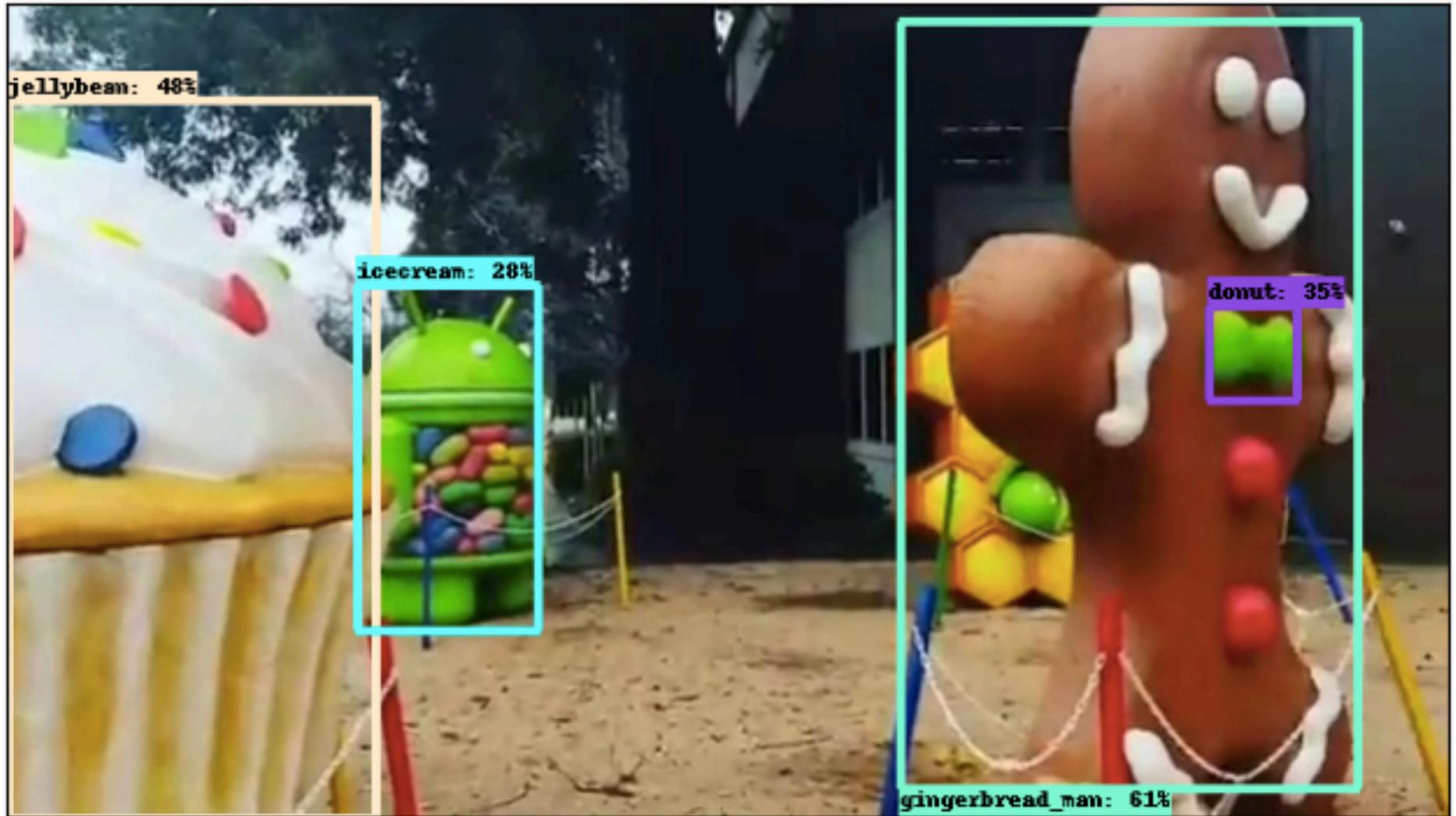
writing to /tmp/tmpveptyde1/9.jpg



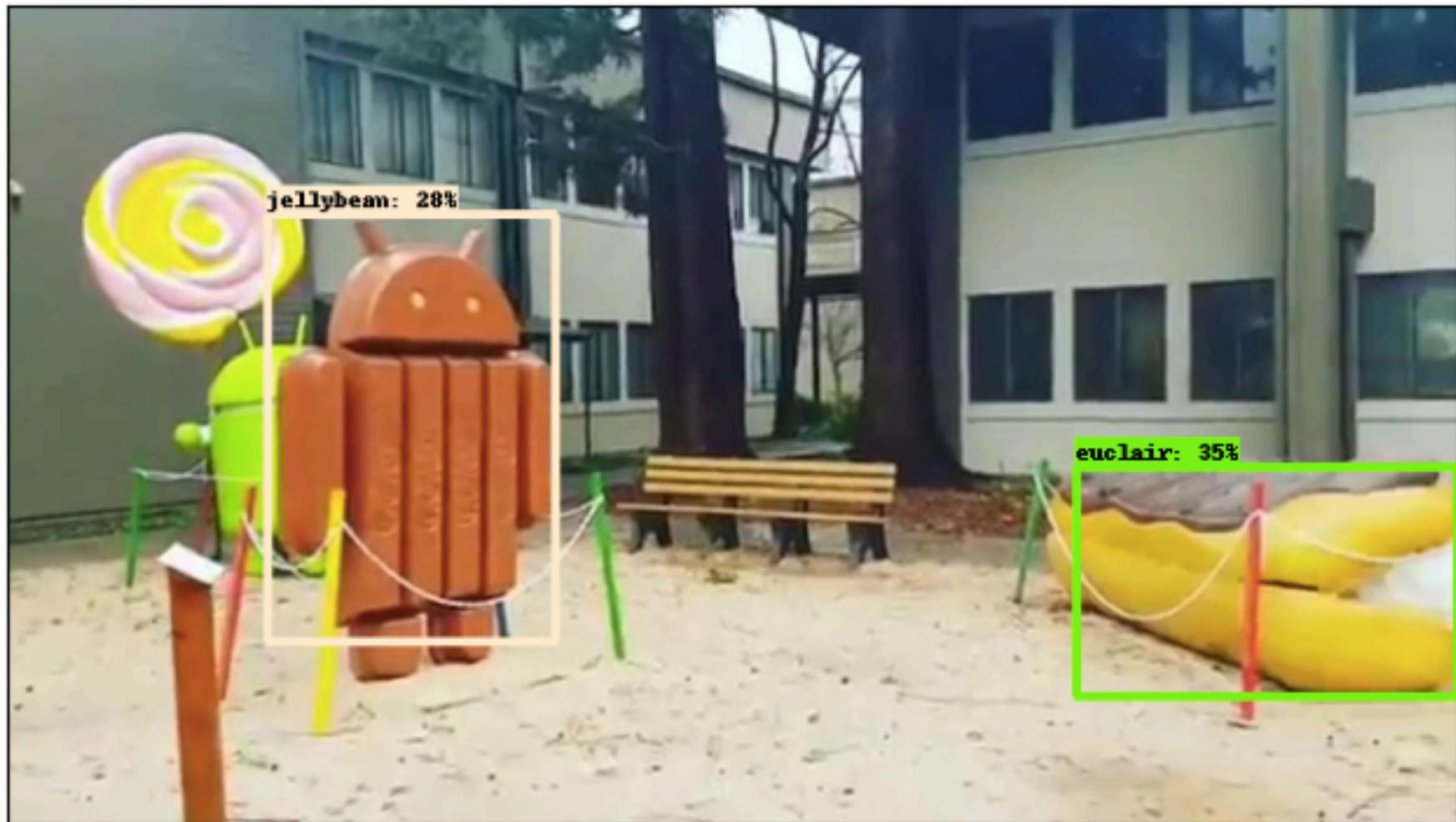
writing to /tmp/tmpveptyde1/10.jpg



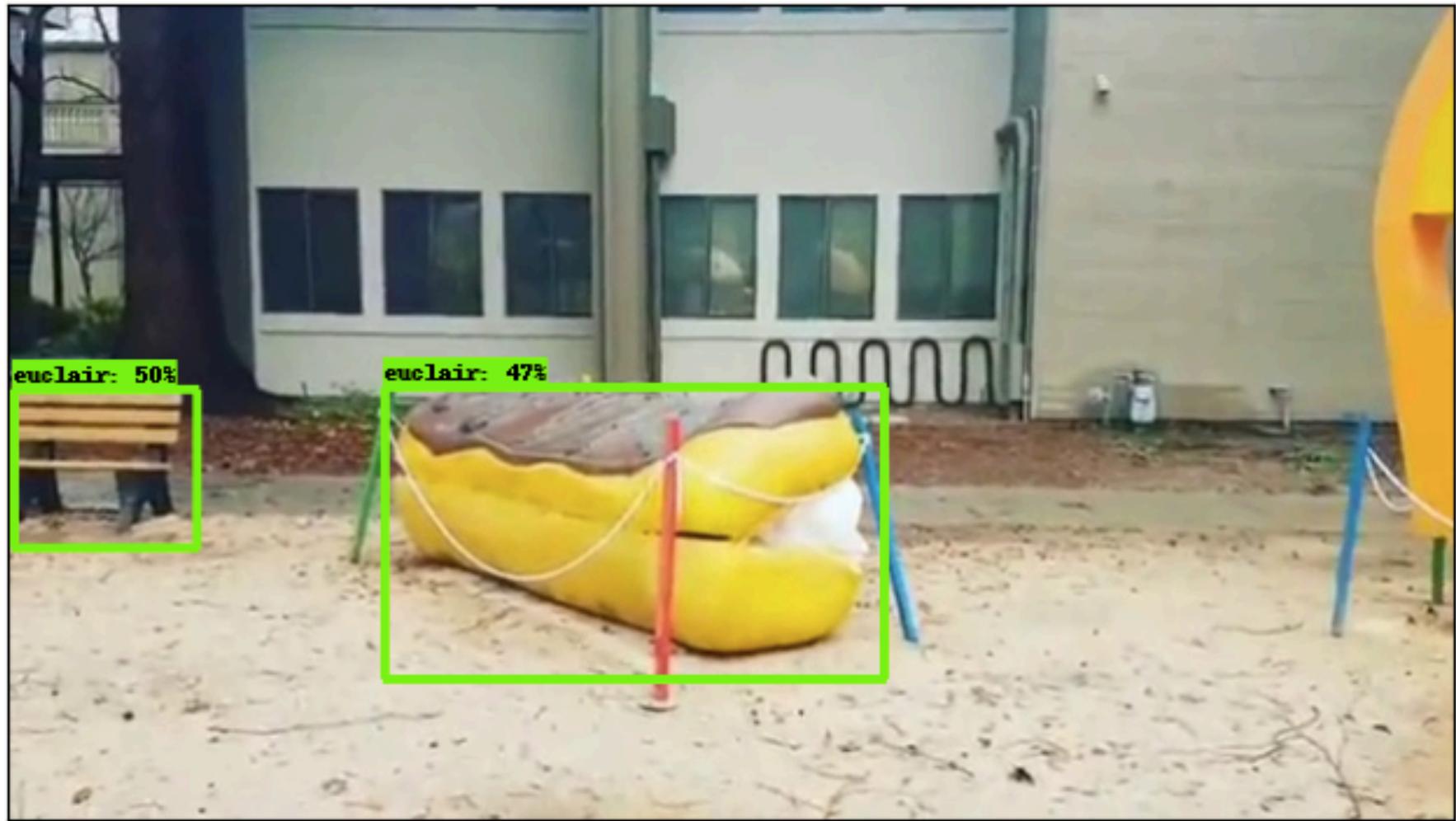
writing to /tmp/tmpveptyde1/11.jpg



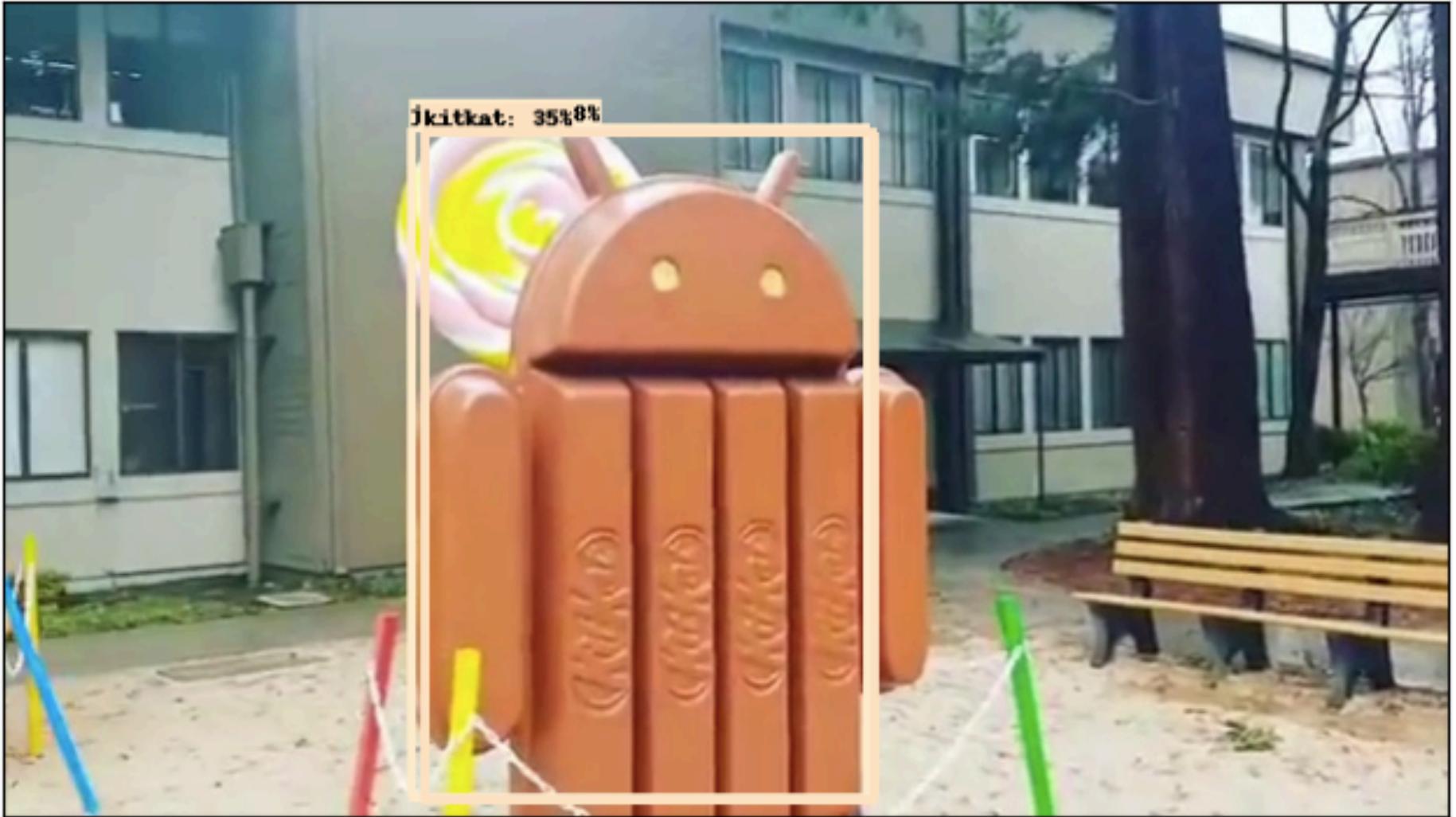
writing to /tmp/tmpveptyde1/12.jpg



writing to /tmp/tmpveptyde1/13.jpg



writing to /tmp/tmpveptyde1/14.jpg

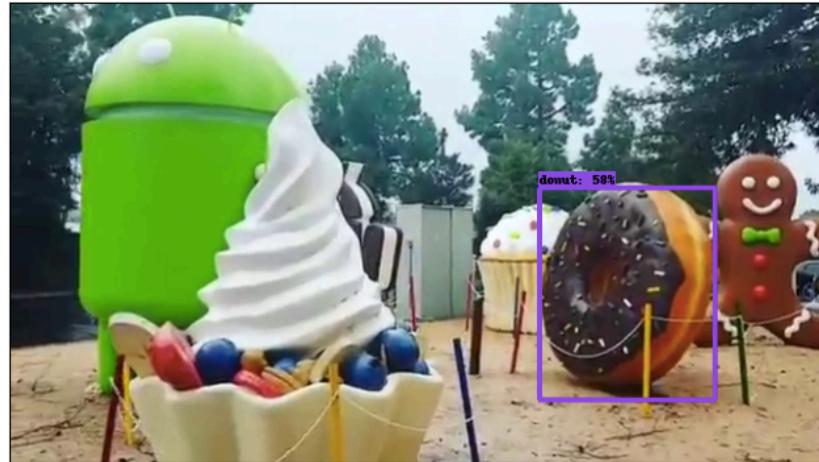


**detections with scores >= 0.5 follow.**

writing to /tmp/tmpaq5c70s4/3.jpg



writing to /tmp/tmpaq5c70s4/4.jpg



writing to /tmp/tmpaq5c70s4/5.jpg



writing to /tmp/tmpaq5c70s4/7.jpg



writing to /tmp/tmpaq5c70s4/8.jpg



writing to /tmp/tmpaq5c70s4/11.jpg



writing to /tmp/tmpaq5c70s4/14.jpg



writing to /tmp/tmpaq5c70s4/1.jpg



writing to /tmp/tmpaq5c70s4/2.jpg



writing to /tmp/tmpaq5c70s4/6.jpg



writing to /tmp/tmpaq5c70s4/9.jpg



writing to /tmp/tmpaq5c70s4/10.jpg



writing to /tmp/tmpaq5c70s4/13.jpg



writing to /tmp/tmpaq5c70s4/12.jpg



## **Next to consider**

- add more training data
- try other pre-trained models
- try other variants of transfer-learning with object detection models.
- for the test frames from the video, could consider applying HMM inference to the best scoring identifications for “smoothing” inference where missing data. can use “filtering” queries for the next observations... can use particle filtering for similar reasons. can even build a better training set by adding the inferred missing identifications (bounding boxes and new ground truth labels) to the existing training dataset and then get a new test dataset.