Activity Report on

Interacting with software using gestures

Submitted by

Aditya Jain, 10200114003
Pallab Kumar Ganguly, 10200114025
Sauradip Nag, 10200114044
Swati Ghosh Hazra, 10200114056

Under the guidance of

Mr. Tamojit Chatterjee, Software Engineer, indeed.com
Dr. Kousik Dasgupta, Asst. Prof., Kalyani Government Engineering College

February 12, 2018

Abstract

Our work on object detection involved detection of open hand from images with as much accuracy as possible, and as little false positives as possible. Our work revealed that smaller window sizes were more promising in terms of accuracy, so we tested classifiers of smaller window sizes, with varying number of stages to obtain a working classifier model. This report deals with the description of the evaluation perfomed with smaller window sizes, its results and interpretation.

I. Introduction

For recognition of gestures, it is essential to be able to first identify the object that will execute the gesture, in this case, the hand. The first step of the project is therefore, to detect the hand successfully. For this purpose, we built a classifier using the Viola-Jones object detection framework [1].

II. Work Done

1. From our series of tests, it was found that with the increase in window size, the number of positive samples successfully detected from a set of 200 images decreased.

2. We therefore proceeded with smaller window size. In particular, we trained three classifiers corresponding to three window sizes:

- C1: Trained on 20x20 positive images
- C2: Trained on 30x30 positive images
- C3: Trained on 40x40 positive images

In addition to this a 50x50 classifier was used for benchmarking [2]. Each of the above HAAR type cascade classifier was trained for 15 stages, using GAB algorithm.

i. Test dataset & performance metric

The test dataset was created as below:

1. Raw positive images of hands (both L and R) grayscale, resized to 20x20

2. Superimposed[1] on random 100x100 grayscale images, created 500 superimposed images.

We also point out to the reader that the in dataset described above, each image consists exactly one positive image. Thus there are 500 true positives in our dataset. The location, height and width of the positives are available in the filename, are used as labels. The parameters varied and their ranges are described in Table 1, & the algoithm in Table 2.

[1] Using opencv_createsamples

Table 1: *Parameters [3]*

No.	Parameter	Range
1	Scale Factor	$[1.01, 1.02, ..., 2.2]$
2	Min. Neighbors	$[1, 2, ..., 50]$
3	Window Size	$[20, 30, 40, 50]$

Table 2: *Algorithm*

1. size $\leftarrow$ 500
2. params $\leftarrow$ scale, neighbours, stages
3. vals $\leftarrow$ range of values of params
4. load cascade file
5. for each val in vals do:
6. count $\leftarrow$ 0
7. for each img in training-set do:
8. detectMultiScale(img, params)
9. get predicted region for img, compare with actual labels
10. if predicted region covers $\geq$ 70% actual region, count it as hit.
11. accuracy=hits/size
12. plot accuracy vs. params.

ii. Results

Refer to figures 1, 2 & 3 for results of tests performed.

III. Future Work

1. We pipeline the results of this stage into a Convolutional Neural Network to further weed out false positives and distinguish between various gestures.
2. We analyse the results from this stage to get an idea of the data that is needed for the CNN.

IV. Conclusion

1. The above evidence suggests the 20x20 classifier seems to be the most promising for future work.
2. We should use parameters in the range: 1.01-1.10 for scale and 2-8 for neighbours

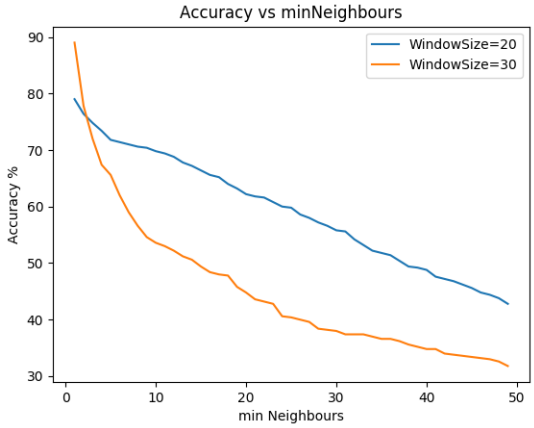

Figure 1: *Accuracy vs. minNeighbours, scale=1.03, 15 stage classifiers*

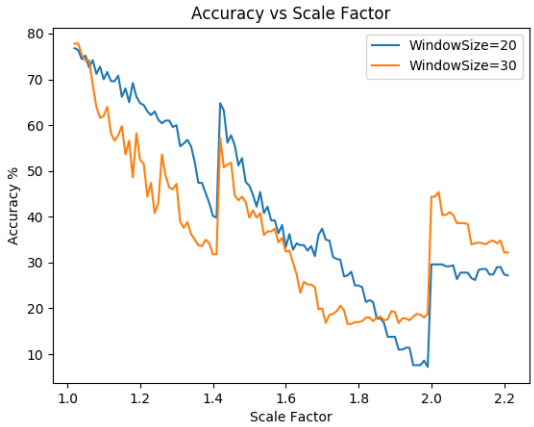

Figure 2: *Accuracy vs. scaleFactor, minNeighbours=2, 15 stage classifiers*

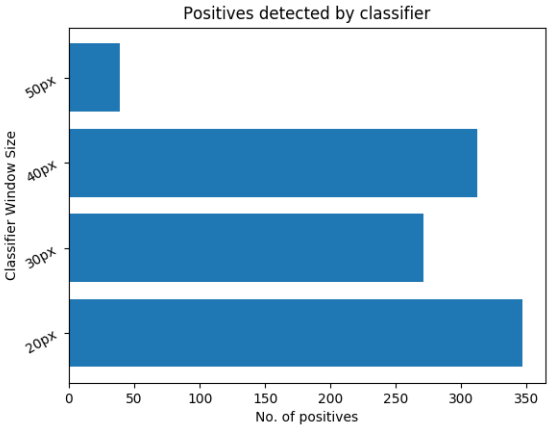

Figure 3: *Positives detected vs. windowSize, scale=1.08, minNeighbours=9, 15 stage classifiers*

References

[1] Viola, Paul and Jones, Michael J., *Robust Real-Time Face Detection*. International Journal of Computer Vision 57(2), 137−154, 2004.

[2] Lefkovits, Szidonia, *Teaching Improvements on HAAR Classifiers*. Petru Maior University Targa-Mures, Department of Mathematics & Computer Science.

[3] OpenCV 2.4.13.5 Documentation/API Reference/Object Detection

Name with Signature

1. .....
   Aditya Jain
2. .....
   Pallab Kumar Ganguly
3. .....
   Sauradip Nag
4. .....
   Swati Ghosh Hazra

Counter signed by guides

1. .....
   Dr. Kousik Dasgupta
2. .....
   Tamojit Chatterjee