16-720 Computer Vision: Homework 5
Image Understanding

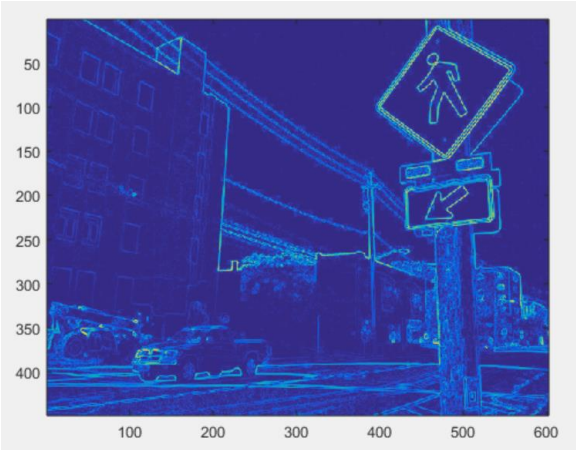
Abhishek Bhatia (abhatia1@andrew.cmu.edu)

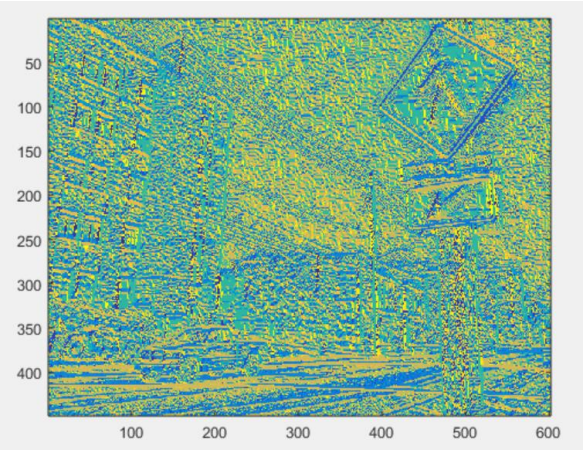
4/14/2016

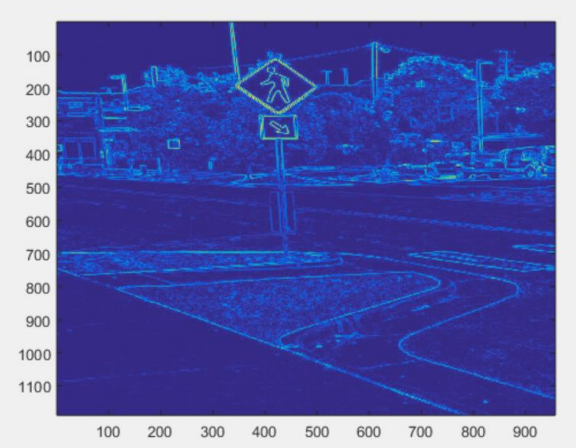
1) Image detection using Histogram of Gradients:

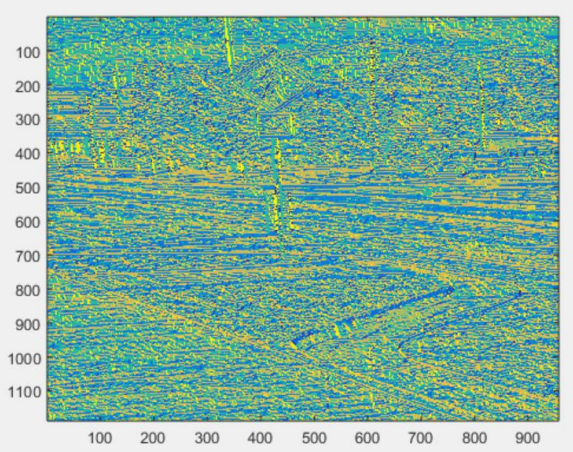
1.1) Image Gradients:

Imagesc visualization of magnitude and orientation of the gradient of two test images:



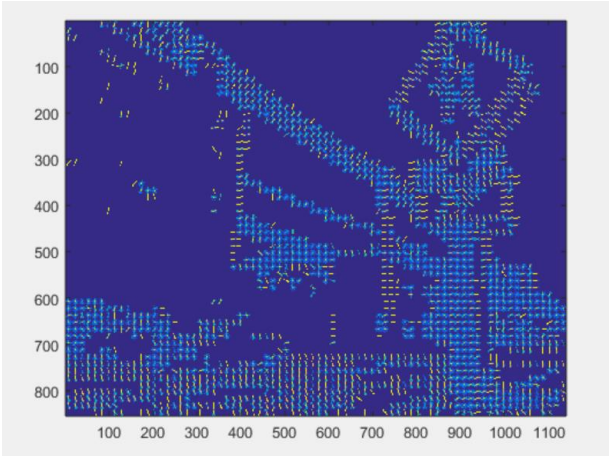


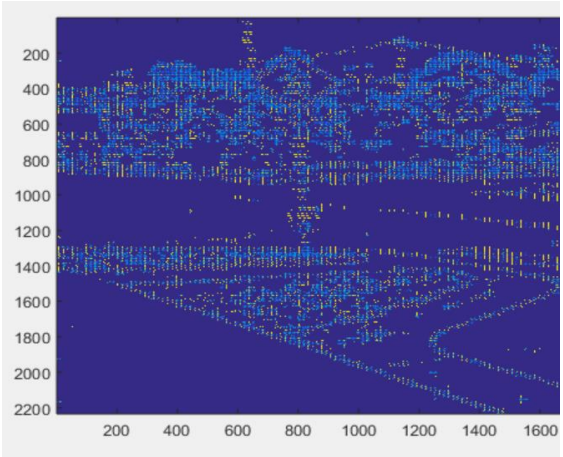




1.2) Histogram of Gradient Orientations:

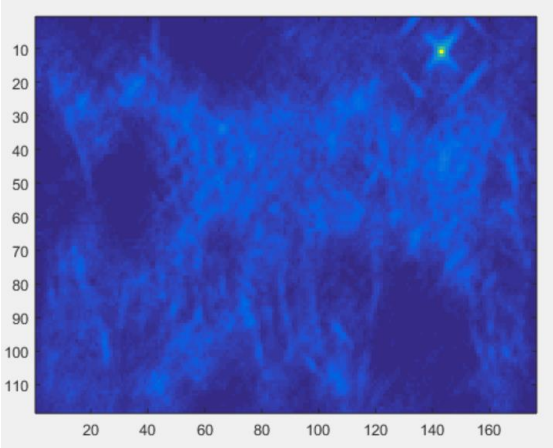
HOG feature visualization of the test images used in previous sub-section:

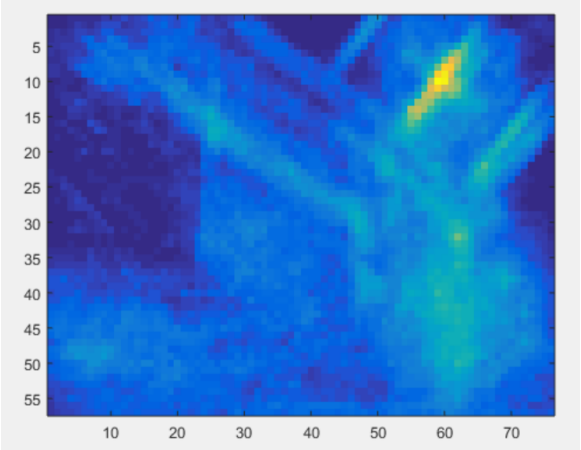


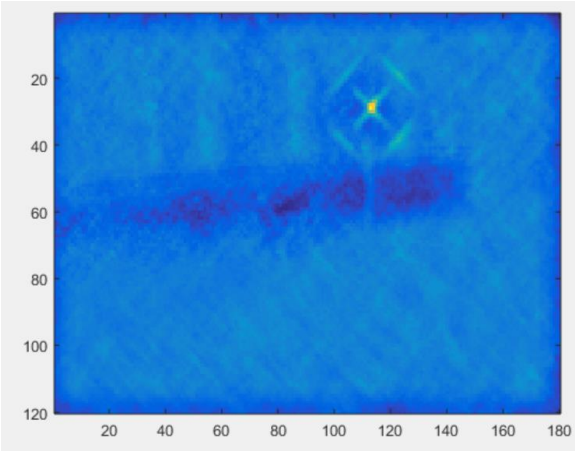


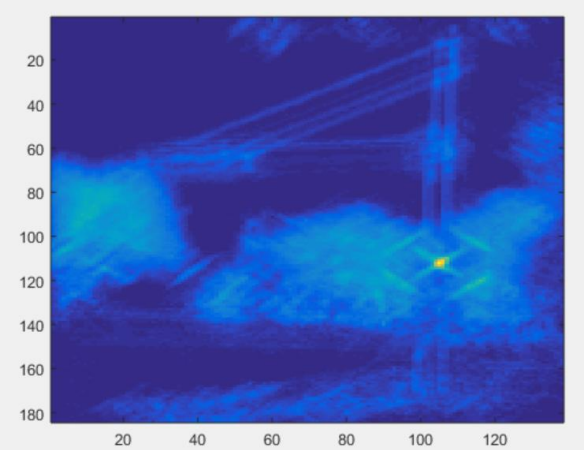
1.3) Detection:

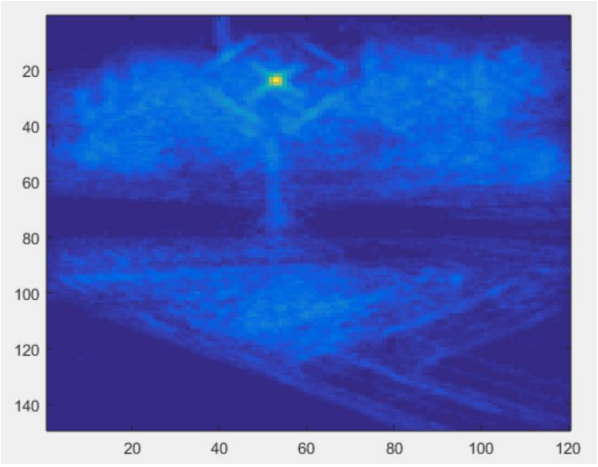
Heatmap visualization of correlation and detected window of all the test images:

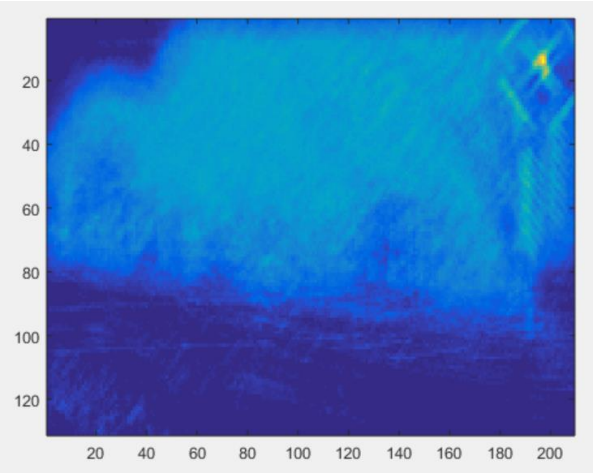



5
10
15
20
25
30
35
40
45
50
55
10
20
30
40
50
60
70

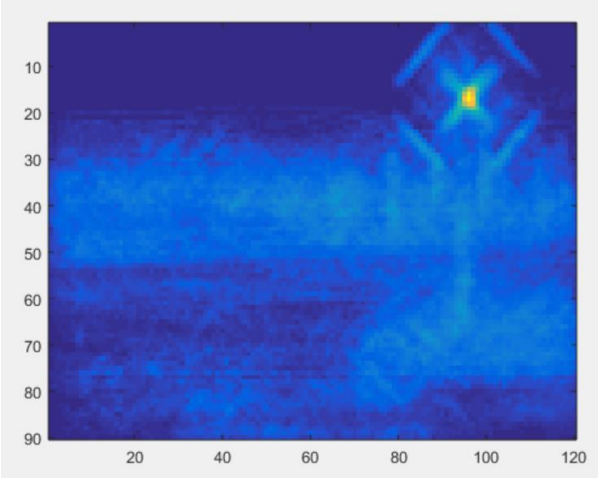

20
40
60
80
100
120
20
40
60
80
100
120
140
160
180


20
40
60
80
100
120
140
160
180
20
40
60
80
100
120



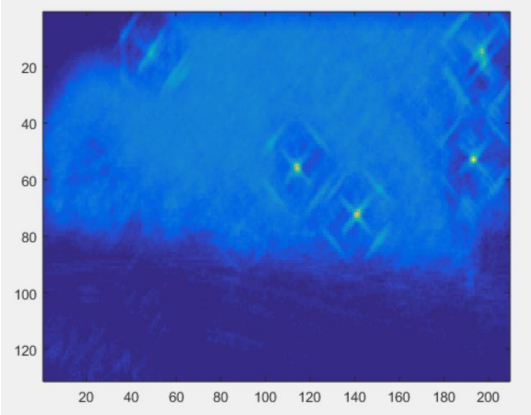


Sign Post Reflector in typical crosswalk application

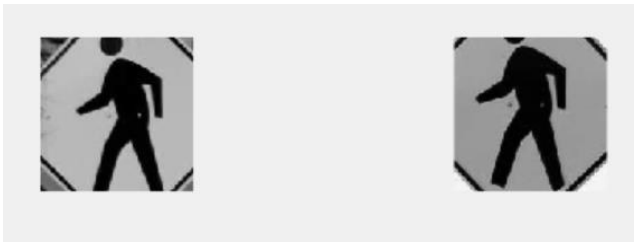


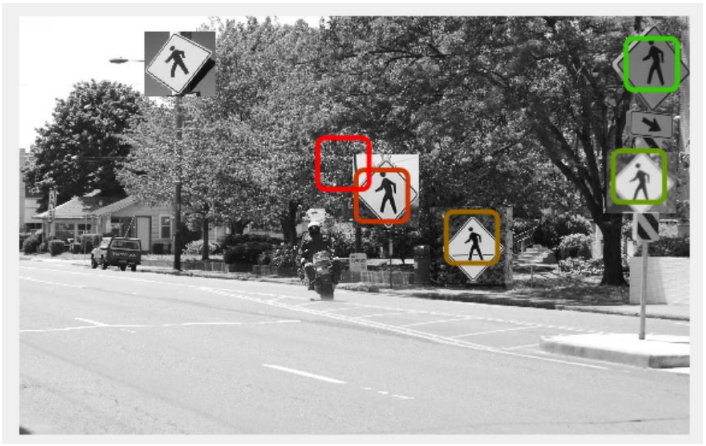
1.4) Extra Credit – Multiple Detections

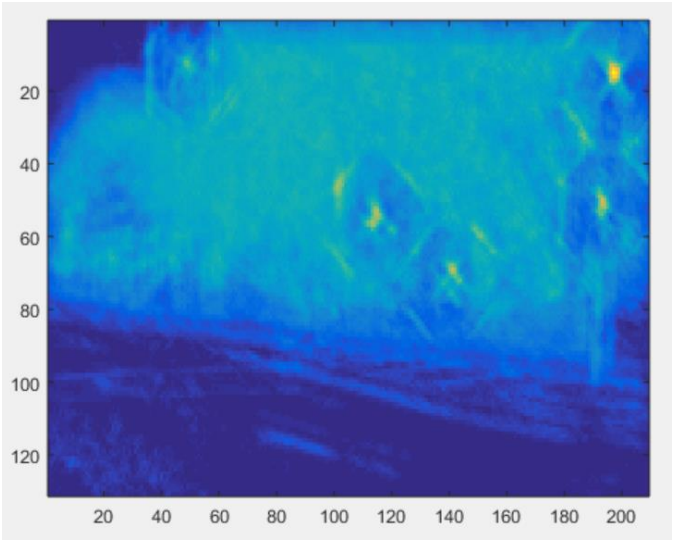
Example of multiple detection (around 2-5 distinct in-stances) of the template in the same image, also Heatmap visualization of correlation below:

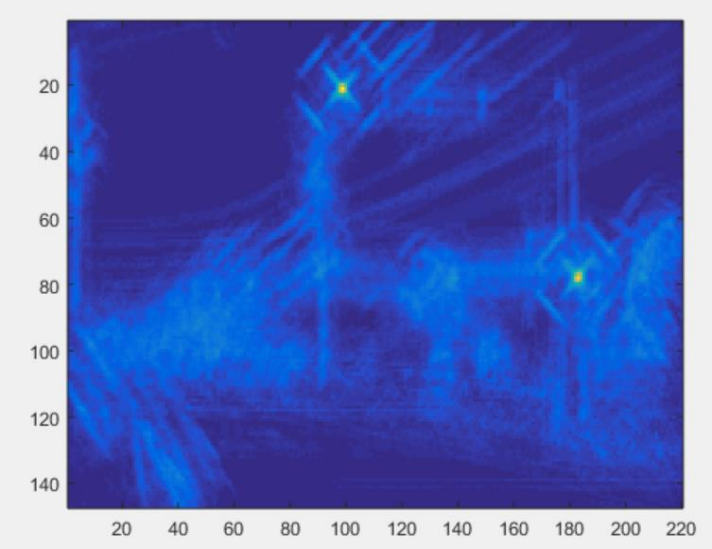


Examples with multiple templates as input and multiple detections Heatmap visualization of correlation below:





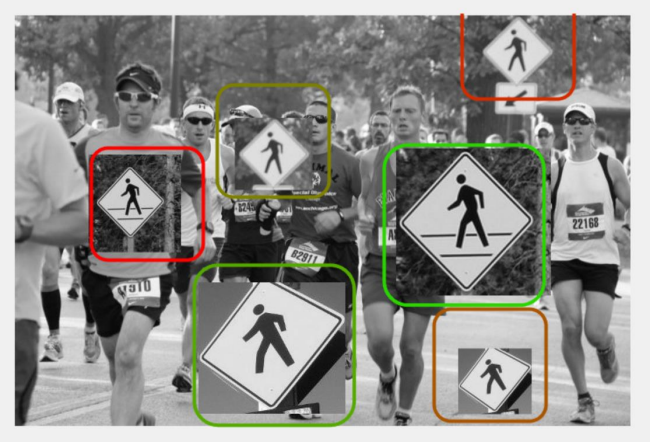

20
40
60
80
100
120
20
40
60
80
100
120
140
160
180
200



2) Learning Templates

2.5) Multi-Scale Detection:

Multiscale detection output showing the detector output on images that contains multiple instances of an object at several different scales. The results show appropriately scaled bounding boxes around each object detected.



Discussion:

1) Of all the different types of learning approaches I tried, the training with LDA worked best for me. The training based on just positive examples performed bad, and was also expected, as we need to remove some information about the negatives when we train the templates. The average of 5 positive and 100 negative performed well, but the LDA approach worked best.
2) I tried different values of the regularization parameter Lambda, mostly between 0 and 1. I got the best results for the value: 0.35.
3) Overall the HOG (Histogram of Gradients) based template matching approach worked well on the limited dataset. The benefits with this approach compared to the SSD (Sum of Squared Distances) we used in the Lucas-Kanade assignment were the light invariance. In Lucas Kanade assignment we witnessed that the template matching approach failed with lighting variations and we had to introduce more information 'appearance basis' to tackle those variations. HOG on the other hand utilizes the gradient information and edge detection techniques and hence performs better in scenarios with lighting changes. SSD relies heavily on closeness of pixels within the template and actual scene, which is good in some scenarios like detecting something in a continuous video feed, where large changes to the scene around the object of interest are not expected. HOG should perform better in more robust scenarios, but it can also get confused with similar shapes/geometries.

Note:

1) Some of the images used for testing were provided to me by my batchmates Tushar Chugh and Bikram Jot Hanjra, I was having trouble finding images of the internet with multiple instances of same templates and we had minor discussions about generating images with replication of the same template with different scales and placing them randomly on the images provided as part of the assignment, this worked really well for.

Below are the image outputs for template learning (2 different images with multiple instance of templates, showing the top 5 detections) in the format expected.

Image 1:

Single scale detector output:

Trained with 1 positive example

Trained with an average of 5 positive examples

Trained with an average of 5 positive examples + 100 negative examples

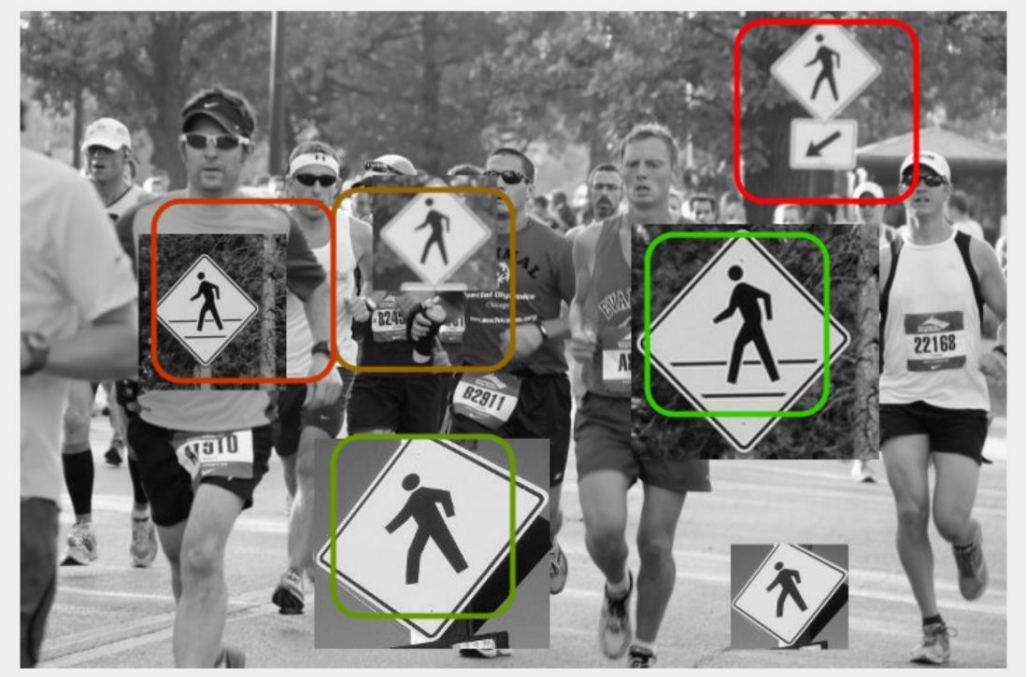
Trained with LDA using 5 positive examples + 100 negative examples

Multi-Scale detector output using LDA

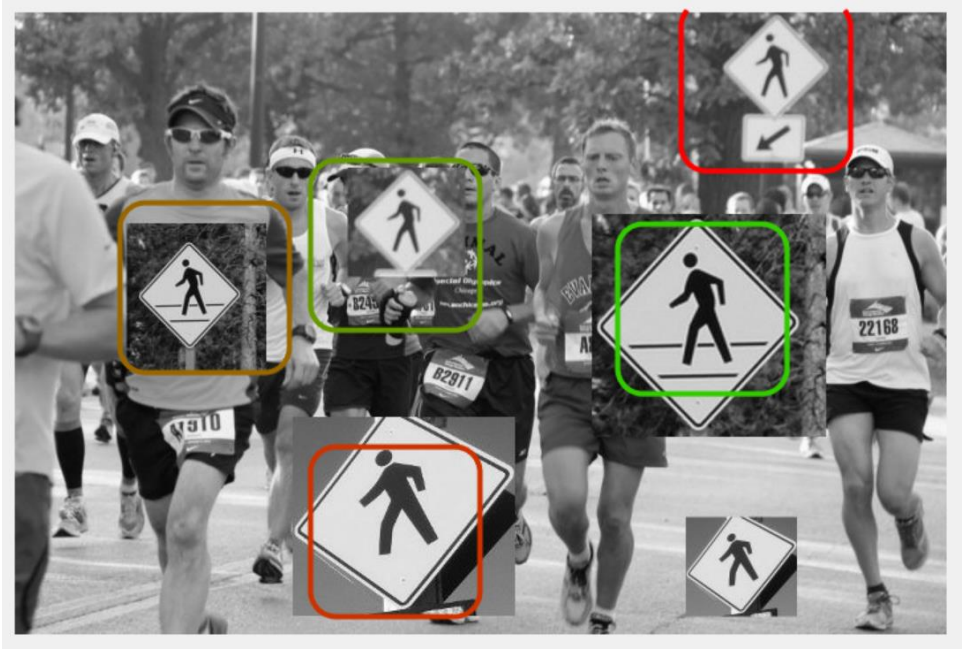
Image 2:

Single scale detector output:

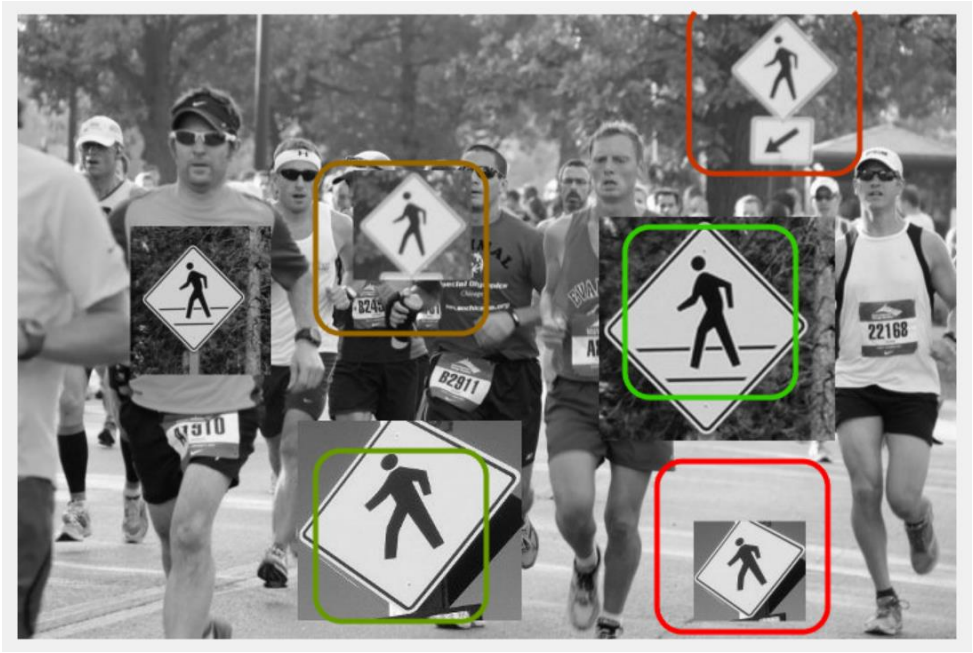
Trained with 1 positive example



Trained with an average of 5 positive examples



Trained with an average of 5 positive examples + 100 negative examples



Trained with LDA using 5 positive examples + 100 negative examples

Multi-Scale detector output using LDA