Q 1.1 Image gradients

The ‘mygradient’ function implements computation of the image gradients using $[-1,0,1]$ filter[1]. The visualization of the magnitude and orientation of test images ‘test0’ and ‘test1’ is shown in Fig. 1 and Fig. 2 respectively.

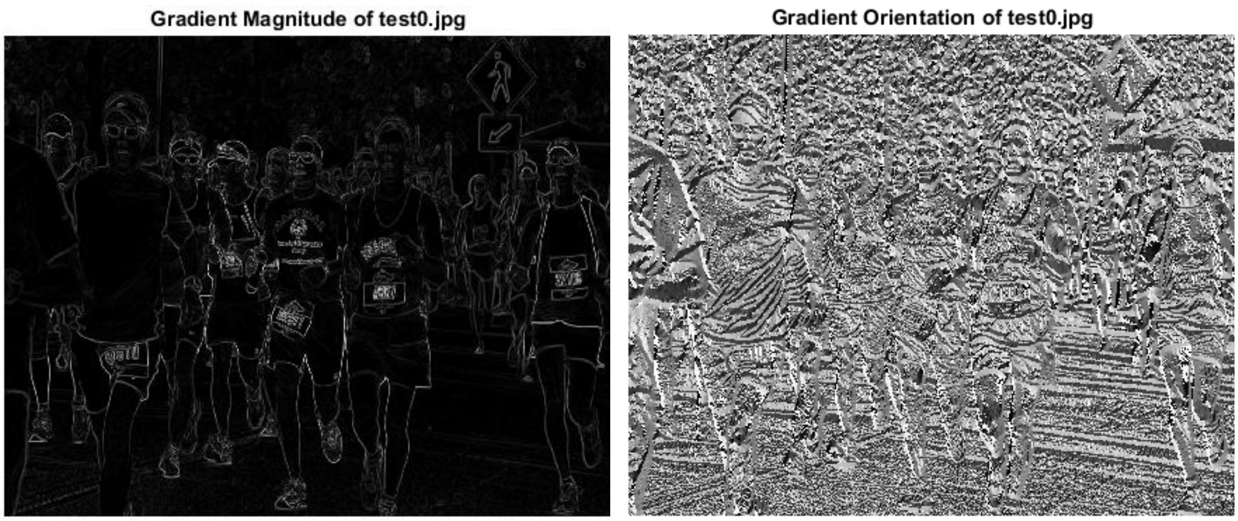

Figure 1. Visualization of the magnitude (left image) and orientation (right image) of test image ‘test0.jpg’.

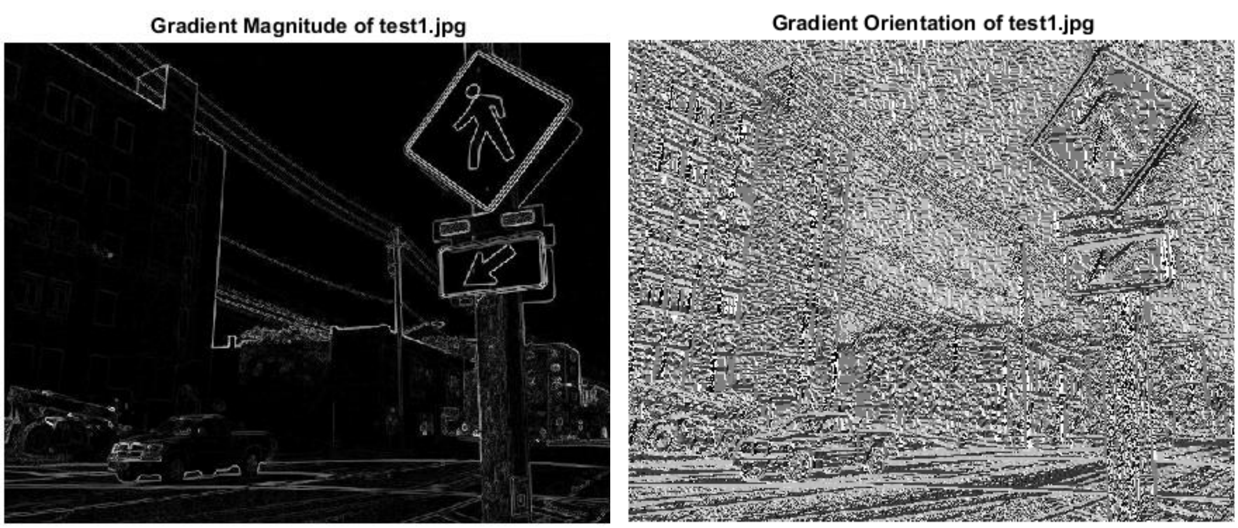

Figure 2. Visualization of the magnitude (left image) and orientation (right image) of test image ‘test1.jpg’.

[1] As suggested in:
N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, San Diego, CA, USA, 2005, pp. 886-893 vol. 1.

Q 1.2 Histograms of gradient orientation

The visualization of the histograms of gradient orientations of test images ‘test0’ and ‘test1’ is shown in Fig. 3. The case of 0 orientations is handled to improve the results.

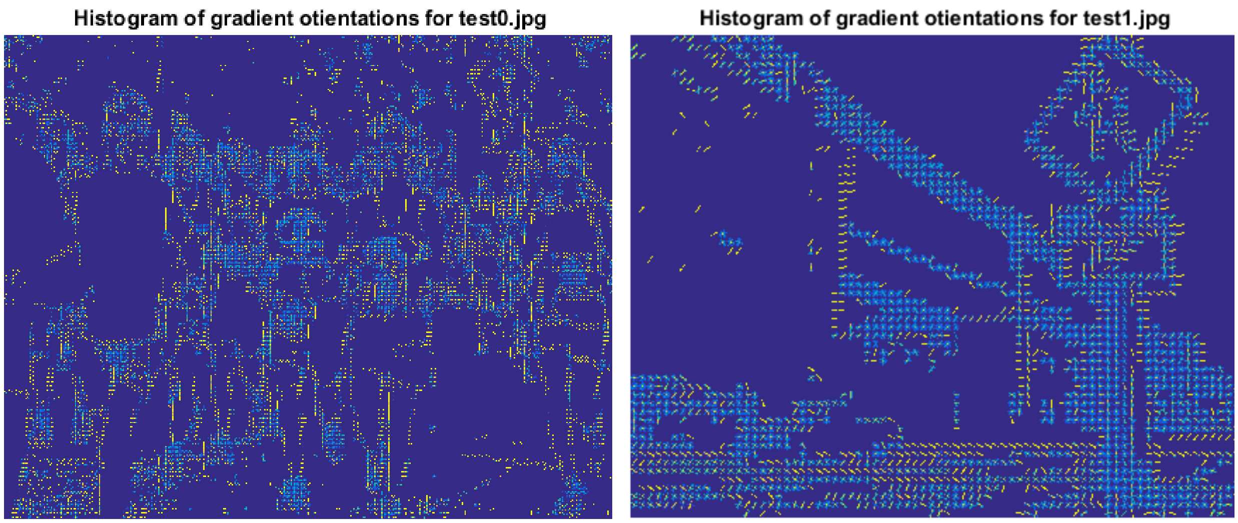

Figure 3. Visualization of histograms of gradient orientation of test images ‘test0.jpg’ (left) and ‘test1.jpg’ (right).

Q 1.3 Detection

MATLAB does not support cross correlation output of the same size as one of the inputs, therefore it is easier to use conv2 function with one of the inputs (the smaller) rotated for 180 degrees instead of xcorr2.

I defined a function called ‘detectwh’ which does the detection (like ‘detect’ function) and additionally returns the heat map. The ‘detect_script’ is slightly modified, so that in order to show the heat map as a result in the ‘detect_script’ script, it is enough to uncomment the call to ‘detectwh’ function (where the ‘detect’ function is called).

The results of the detection function on all the test images along with their heat maps (cross-correlation response map) are shown in Fig. 4 – 10.

Figure 4. The result of the detection function (left) and the heat map (right) of ‘test0.jpg’ image.

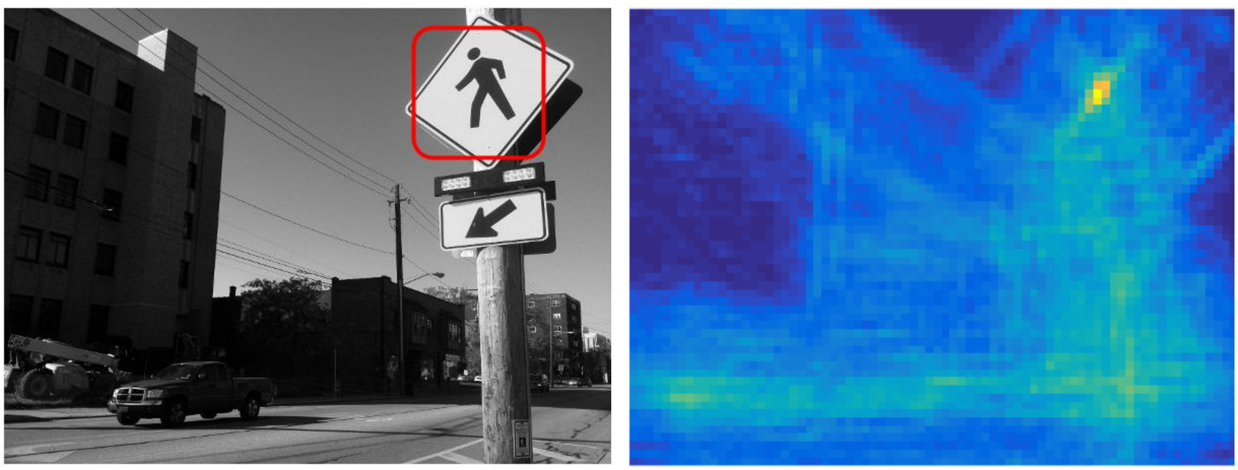

Figure 5. The result of the detection function (left) and the heat map (right) of ‘test1.jpg’ image.

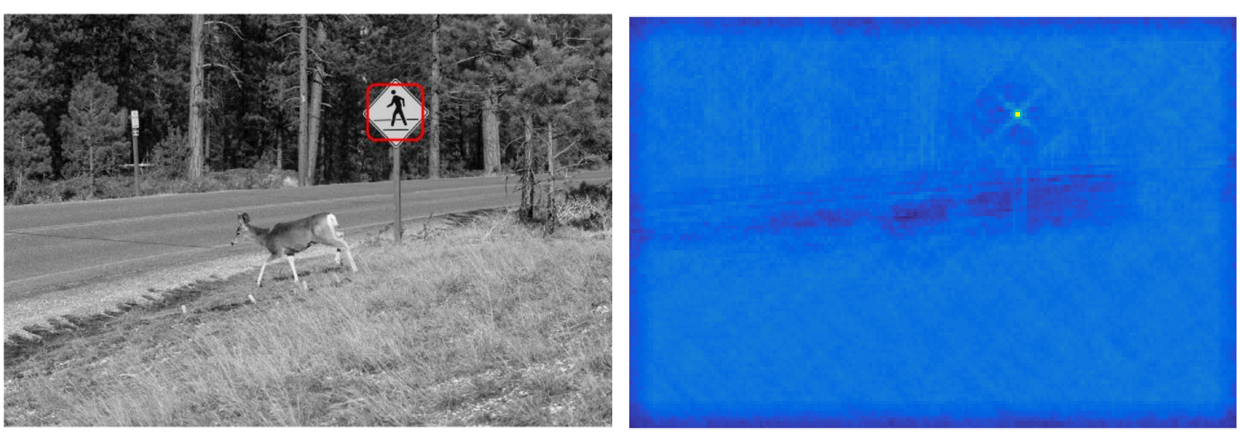

Figure 6. The result of the detection function (left) and the heat map (right) of ‘test2.jpg’ image.

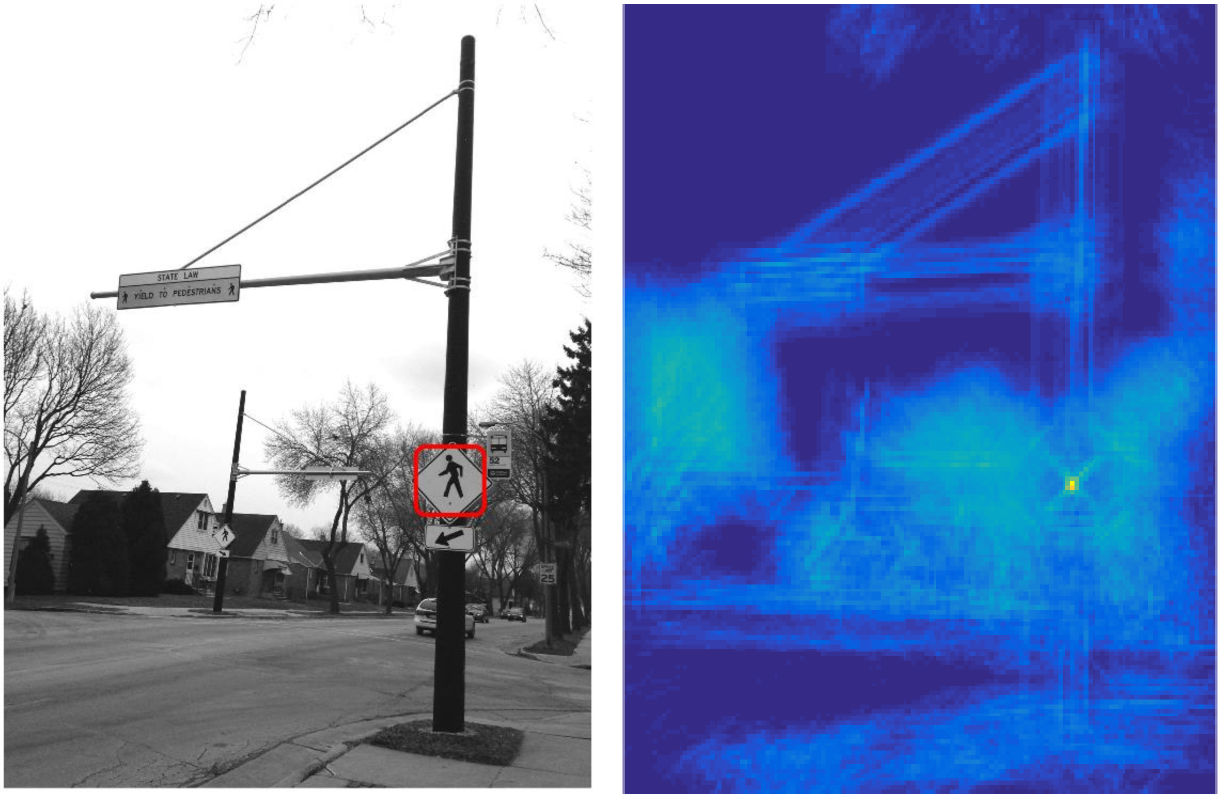

Figure 7. The result of the detection function (left) and the heat map (right) of ‘test3.jpg’ image.

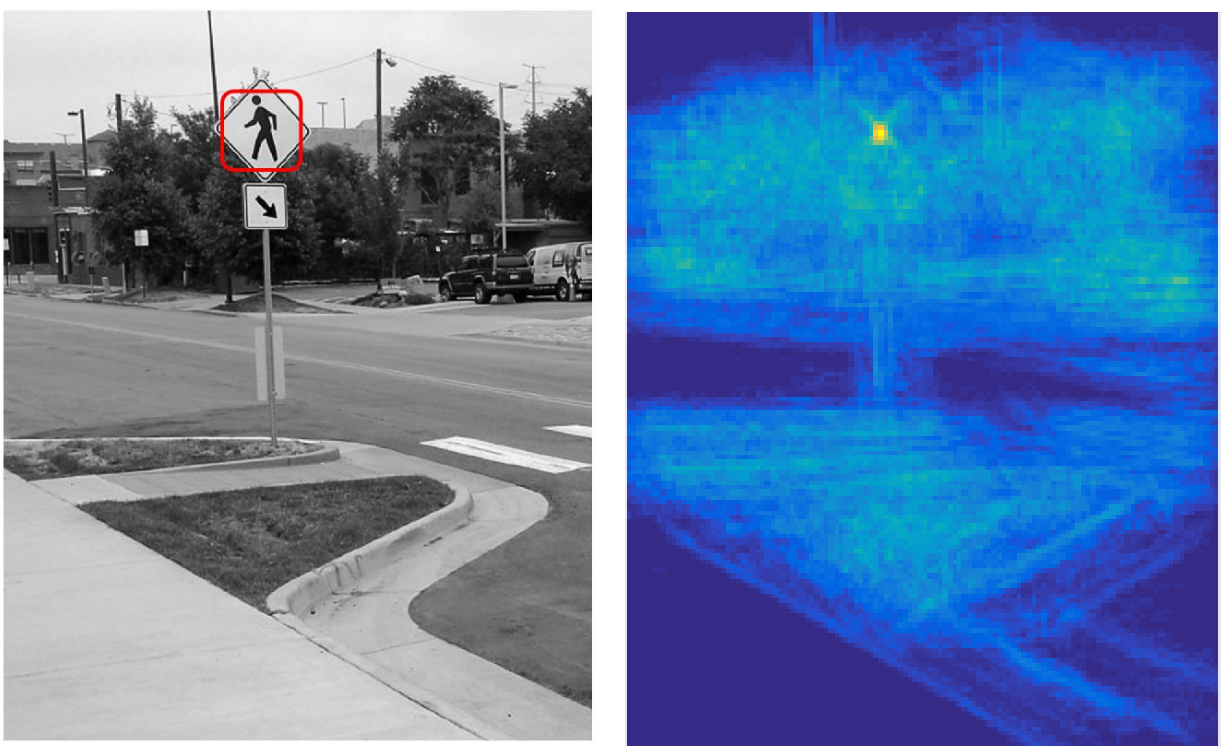

Figure 8. The result of the detection function (left) and the heat map (right) of ‘test4.jpg’ image.

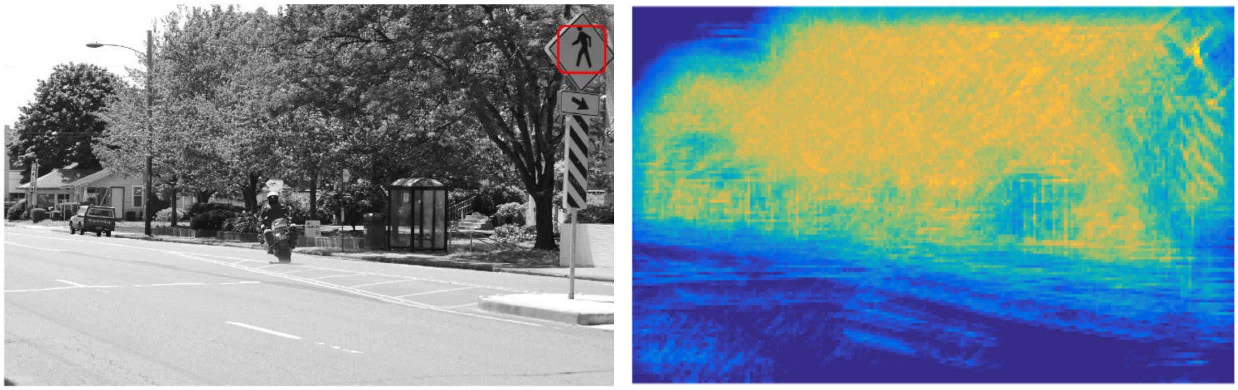

Figure 9. The result of the detection function (left) and the heat map (right) of ‘test5.jpg’ image.

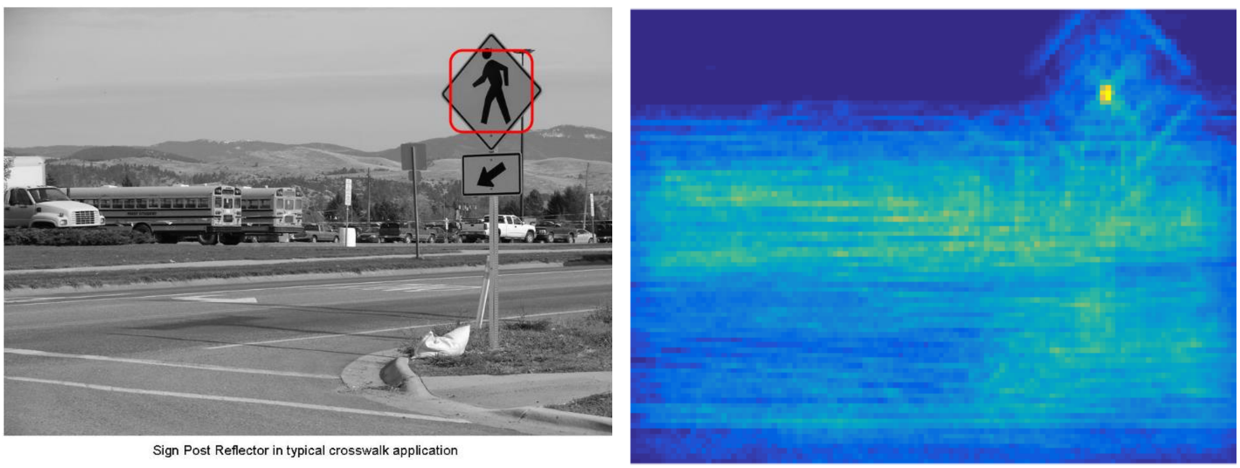

Figure 10. The result of the detection function (left) and the heat map (right) of ‘test6.jpg’ image.

Q 1.4 Multiple detections

The modified script for multiple detections is in ‘detect_script_custom’. The only change is the names of the test files. The test image files used in this section are in the ‘custom’ folder.

The ‘test100.jpg’ image is used to select the template and the ‘test101.jpg’ is used for detection. Test image ‘test100.jpg’ and the selected template are shown in the top of Fig. 11. The two detections on the ‘test101.jpg’ and the associated heat map are shown in the bottom of the Fig. 11.

The ‘test111.jpg’ image is used to select two different templates and the ‘test112.jpg’ is used for detection. Test image ‘test111.jpg’ and the selected templates are shown in the top of Fig. 12. The two detections on the ‘test112.jpg’ and the associated heat map are shown in the bottom of the Fig. 12.

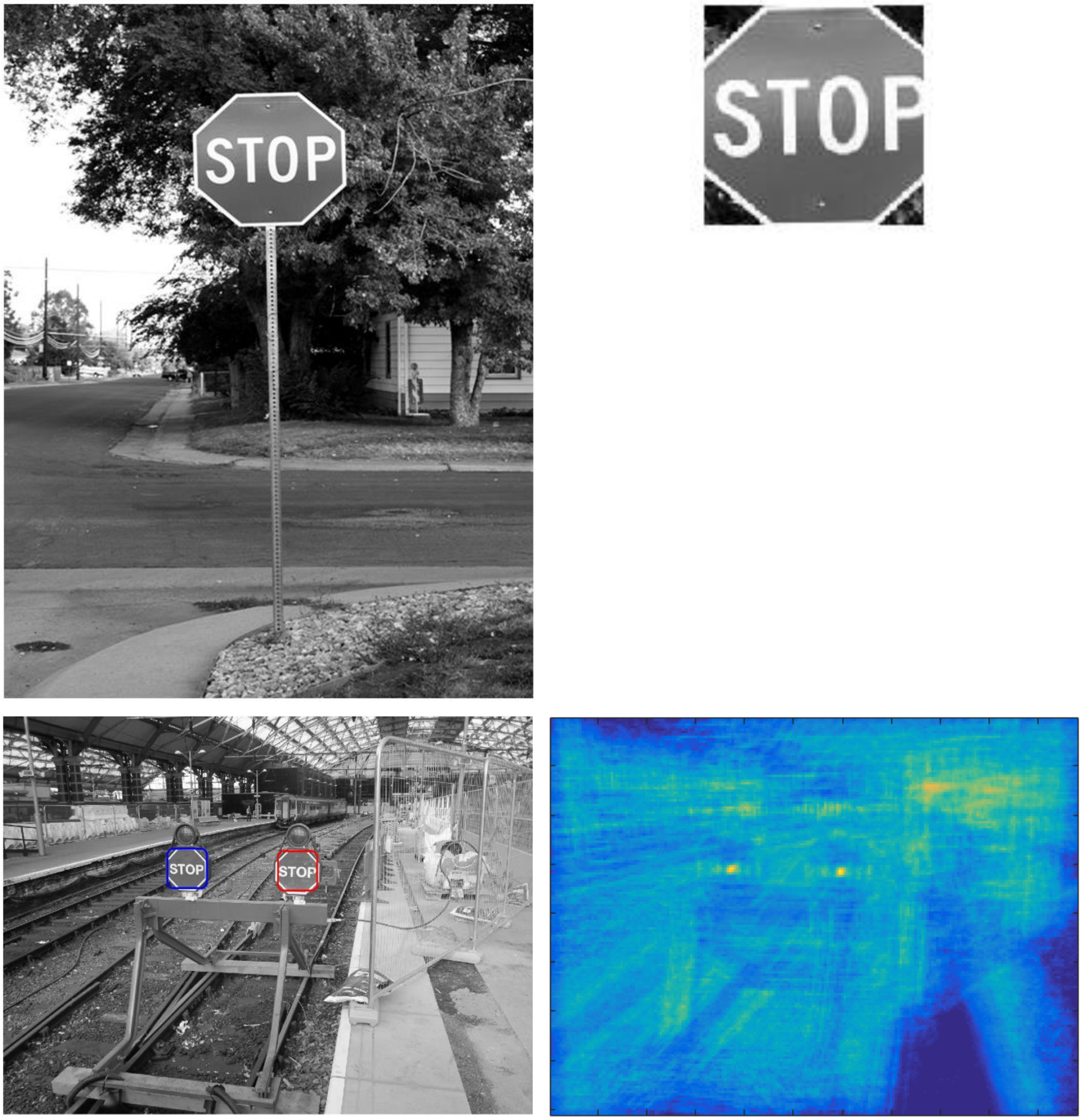

Figure 11. The original image ‘test100.jpg’ (top left) and the selected template (top right). The result of 2 detections (bottom left) and the heat map (bottom right) of ‘test101.jpg’ image.

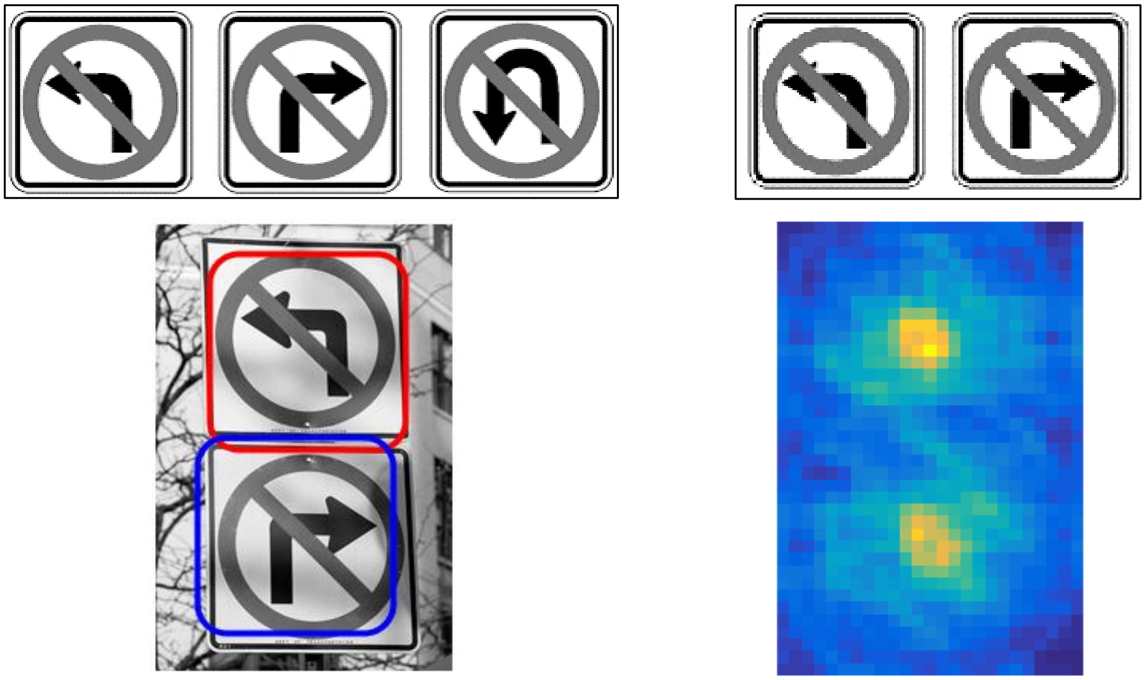

Figure 12. The original image ‘test111.jpg’ (top left) and the selected template (top right). The result of 2 detections (bottom left) and the heat map (bottom right) of ‘test112.jpg’ image.

Important: The results in the sections 2.1-2.5 are only to show the progress and ideas. The actual requested results (as specified on page 6 of the description PDF) are after section 2.5 in the “Template learning results” section.

Q 2.1 Select Patches

The function asks interactively from the user to input the number (and names) of images and the number of positive and negative examples for each image. The function makes the user selections of the positive examples square. Then it randomly selects the intended number of negative examples. The result of the function on the test images ‘test0.jpg’ and ‘test1.jpg’ with the selected areas on the image is shown in Fig. 13. The number of the positive examples (selected by user) and negative examples (selected randomly without overlaps) is determined by the user for each image in the collection.

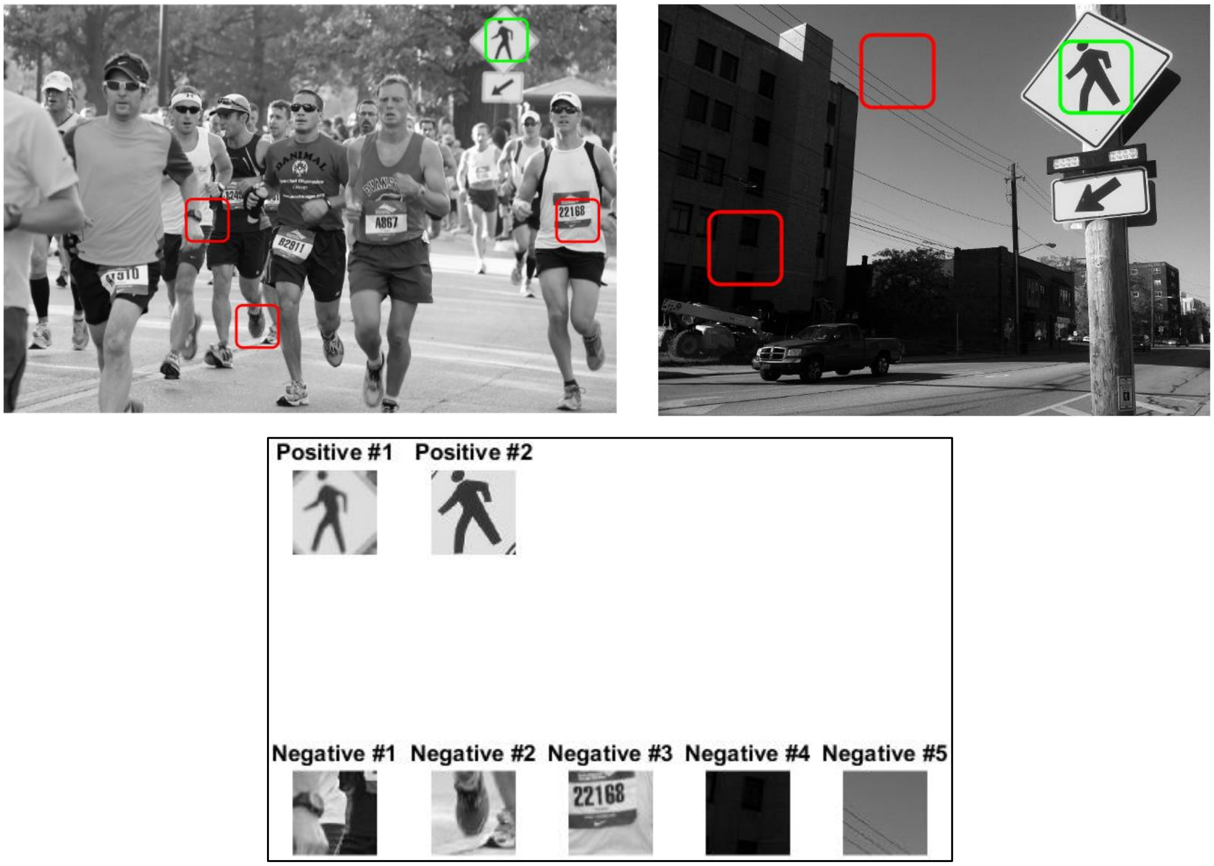

Figure 13. The selected positive (green) and negative (red) areas on the test images ‘test0.jpg’ (top left) and ‘test1.jpg’ (top right) with the resulted positive and negative resized examples (bottom).

Fig. 14 shows the 4 user-selected positive examples with the 79 randomly-selected non-overlapping negative examples on the ‘test0.jpg’ image. The templates in the following sections are calculated using the patches in the Fig. 13 above.

Figure 14. 4 user-selected positive examples with the 79 randomly-selected non-overlapping negative examples on the ‘test0.jpg’ image.

Q 2.2 Positive template learning

The result of the detection using the positive templates of the Fig. 14 on the test image 'test0.jpg' is shown in the Fig. 15.

Figure 15. The result of the detection function using positive examples on the 'test6.jpg' image.

Q 2.3 Positive negative template learning

The result of the detection using the positive and negative examples of the Fig. 14 on the test image 'test0.jpg' is shown in the Fig. 16. There is an improvement over the detection using only positive examples.

Figure 16. The result of the detection function using positive and negative examples on the 'test6.jpg' image.

Q 2.4 Detection using LDA template

The result of the detection using the LDA template of the Fig. 14 on the test image 'test0.jpg' is shown in the Fig. 17. I used $\lambda = 0.01$ as specified in the reference of this section; larger values were giving slightly worse detection results. Also, choosing a small value in order to preserve the actual data and only avoid singularity in the covariance matrix has more meaning than largely changing the matrix.

There is an improvement over the detection using only positive examples; however, there is not an obvious improvement over the detection using both positive and negative examples.

Figure 17. The result of the detection function using the LDA template on the ‘test6.jpg’ image.

Q 2.5 Multi-scale detection

The result of the multi-scale detection using the LDA template of the Fig. 13 on the test image ‘test102.jpg’ is shown in the Fig. 18.

Figure 17. The result of the multi-scale detection using the LDA template on the 'test102.jpg' image.

Template learning results

Set 1 is used to test the detection of traffic lights and consists of two images 'test10.jpg' and 'test11.jpg' for training and 'test12.jpg' for the testing. These images are shown in the Fig. 18.[2]

[2] Source for all images: Google Images Search

Figure 18. The images used for traffic light detection. Images "test10" (top left) and "test11" (top right) are used for the training and image "test12" (bottom) is used for the testing.

The template and the result of the top 5 detections using 1 positive example is shown in the Fig. 19.

Figure 19. The template (right) and the result of detection on test image 'test12" (left) using 1 positive example.

The templates from images “test10” and “test11” and the result of the top 5 detections using the average of 5 positive examples is shown in the Fig. 20. It shows slightly worse detection compared with only using 1 good positive example.

Figure 20. The templates (right) and the top 5 detections on test image ‘test12” (left) using 5 positive examples from the training images “test10” and “test11”.

The result of the top 5 detections with the template calculated using average of 5 positive examples and the average of near to 100 negative examples is shown in the Fig. 21. It shows slightly better detection compared with only using the positive example.

Figure 21. Top 5 detections in the test image “test12” with a template calculated using the average of 5 positive and near to 100 negative examples from the training images “test10” and “test11”.

The result of the top 5 detections with a template calculated by LDA algorithm using 5 positive examples and near to 100 negative examples is shown in the Fig. 22. It shows better detection compared with other previous detectors and can detect 5 traffic lights in the image.

Figure 22. Top 5 detections in the test image “test12” with a template calculated using LDA using 5 positive and near to 100 negative examples from the training images “test10” and “test11”.

The result of the top 6 multi-scale detections with a template calculated by LDA algorithm using 5 positive examples and near to 100 negative examples is shown in the Fig. 23. It shows the detection of the nearest light at a different scale. Since the appearance of the nearest traffic light was different from the others, I photoshopped it to make the appearance almost similar to the others but with a different scale.

Figure 23. Top 6 multi-scale detections in the test image “test12” with a template calculated using LDA using 5 positive and near to 100 negative examples from the training images “test10” and “test11”.

Set 2 is used to test the detection of human faces and consists of an image ‘test21.jpg’ for training and ‘test22.jpg’ for the testing. These images are shown in the Fig. 24.[3]

Figure 24. The images used for human face detection. Image “test21” (left) is used for the training and image “test22” (right) is used for the testing.

The template from “test21” image and the result of the top 5 detections in “test22” image using 1 positive example is shown in the Fig. 25.

Figure 25. The template (right) and the result of detection on test image ‘test22” (left) using 1 positive example.

The templates from image “test20” and the result of the top 5 detections using the average of 5 positive examples is shown in the Fig. 26. Like set 1, the detection using the average of 5 positive examples shows slightly worse detection compared with only using 1 good positive example.

[3] Source for all images: Google Images Search

Figure 26. The templates (right) and the top 5 detections on test image ‘test22” (left) using 5 positive examples from the training image “test21”.

The result of the top 5 detections with the template calculated using average of 5 positive examples and the average of 100 negative examples is shown in the Fig. 27. It shows better detection compared with only using the positive examples.

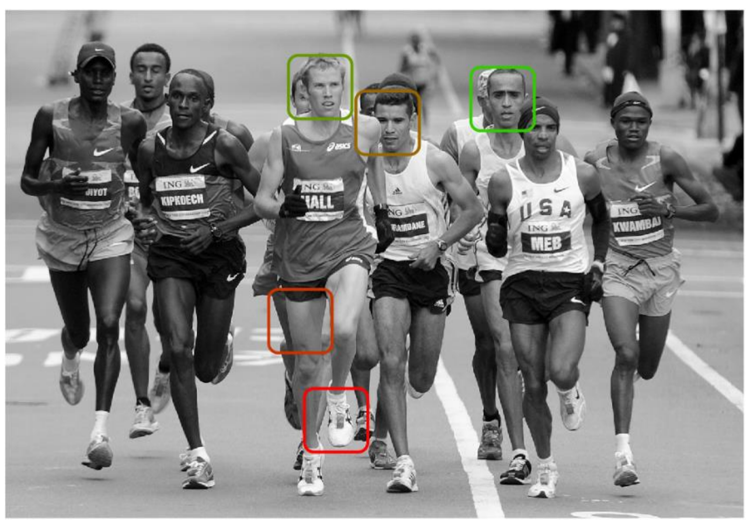

Figure 27. Top 5 detections in the test image “test22” with a template calculated using the average of 5 positive and 100 negative examples from the training image “test21”.

The result of the top 5 detections with a template calculated by LDA algorithm using 5 positive examples and 100 negative examples is shown in the Fig. 28. It shows much better detection compared with other previous detectors and can detect 4 faces in the image.

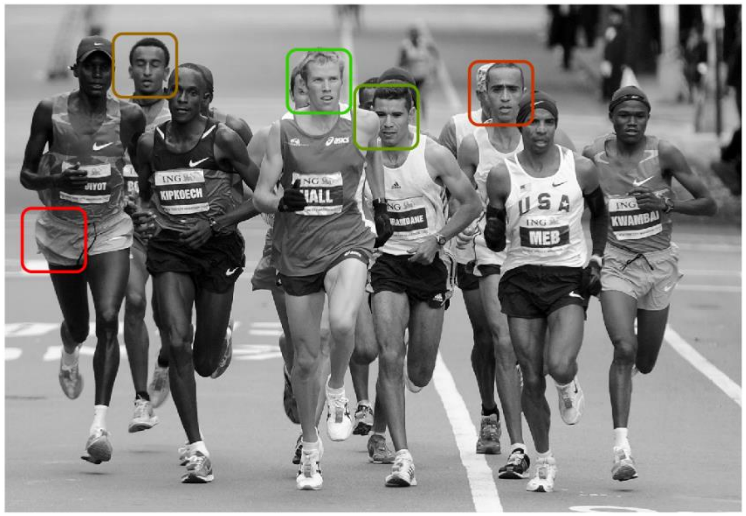

Figure 28. Top 5 detections in the test image “test22” with a template calculated using LDA using 5 positive and 100 negative examples from the training image “test21”.

The result of the top 6 multi-scale detections with a template calculated by LDA algorithm using 5 positive examples and 100 negative examples is shown in the Fig. 29. It shows the detection of 5 faces in 3 different detection scales.

Figure 29. Top 6 multi-scale detections in the test image “test22” with a template calculated using LDA using 5 positive and near to 100 negative examples from the training image “test21”.

Conclusion

In both datasets I used in the previous section, the detection result using only 1 good positive example is slightly better than the result of using an average of 5 positive examples. After checking the average template, in both cases the average positive template is affected with the backgrounds of all 5 positive templates, reducing the sharpness of the object gradients and adding additional noise to it, while the single positive example is selected much more carefully and has better gradients and less noise in overall. However, when the average of negative examples is subtracted from the average of the positive examples, the background noise is reduced and the gradients of the object are enhanced. Therefore, the detection is better than using only single or 5 positive templates. Using the linear discriminant analysis (LDA), the covariance of the negative examples enhances the template and the final detection algorithm is able to detect the objects with worse backgrounds than the previous algorithms. Therefore, the detection with the LDA trained template is much better.

The reason that the detection using the histogram of gradients (HOG) works better than the detection with the pixel values is that the detection using the pixel values is extremely sensitive to the illumination and color changes. For example, a detection algorithm using the pixel values calculates a large difference between a traffic light with a green light on and a traffic light with a red light on; also it calculates a large difference between the dark and light skin colors. On the other hand, we saw in the previous section that the HOG-based algorithms could easily detect both this cases.

Q 2.6 Mixture of templates

The code is implemented in the “tl_detect_script_custom.m” file. A result of using two templates calculated from “test21” and “test24” images to detect faces in the “test22” image file is shown in Fig. 30.

To be able to compare the scores from different templates, I used the result of the cross-correlation which was used in the previous sections (returned by detect function).

I couldn’t find a very good image with the same object from different angle of views in the image to differentiate the mixture of templates with a simple approach.

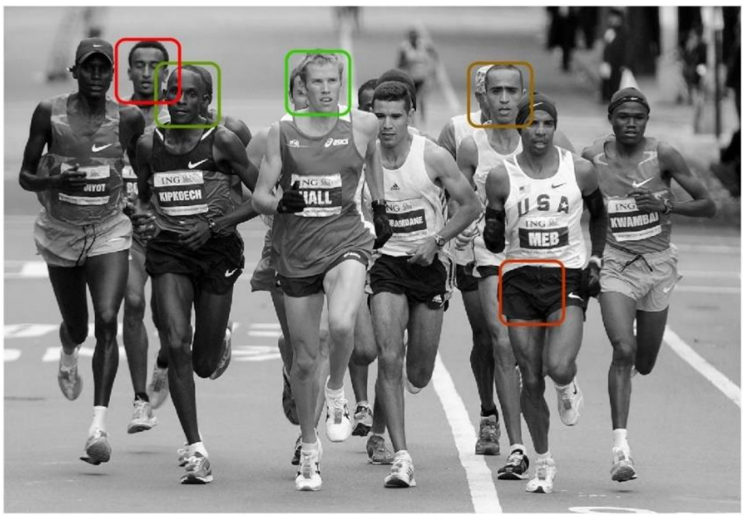

Figure 30. Top 5 detections in the test image “test22” with a mixture of templates calculated using LDA using 5 positive and near to 100 negative examples from each of the training images “test21” and “test24”.