Nikhil Nagraj Rao
PID: n3621940

Date: 10-Nov-14
Subject: CAP 5415

Programming Assignment 3 – Part C

Implementation Details:

1. The entire implementation is in the file 'detectPedestriansHOG.m'.
2. To run the script, the trained data is needed from SVM. This was generated from the previous script, which was submitted in Part B. This has to be in the directory 'Results/Wb.mat'. The entire directory containing the file 'Wb.mat' is attached with the submission.
3. The script reads the input image, generates a scale space and using a sliding window approach calculates the HOG descriptors using the function 'findHistBlock()'.
4. The SVM weights are then multiplied and added with the bias, to obtain the scores of the all window positions in the image over all the entire scale space.
5. All the positive scores undergo non maxima suppression and bounding box overlap analysis to individually detect all pedestrians (humans) in the input images.
6. Output is shown as a figure with a bounding box around each detected pedestrian/human.

Note: *For all the images shown, a step size of 10 has been used for the sliding window, an overlap ratio of 0.9 is employed.*

Results:

Positive Examples:

Trial-1: The results for a simple image are shown in the first trial. The detection is quite straightforward and is detected in the first scale itself. Non Maxima suppression was performed on the bounding boxes with an overlap threshold of 90% meaning that all bounding boxes having an overlap of greater than 90% were removed. Notice that all the overlapping bounding boxes were removed in this case due to the fact that there is only one person.

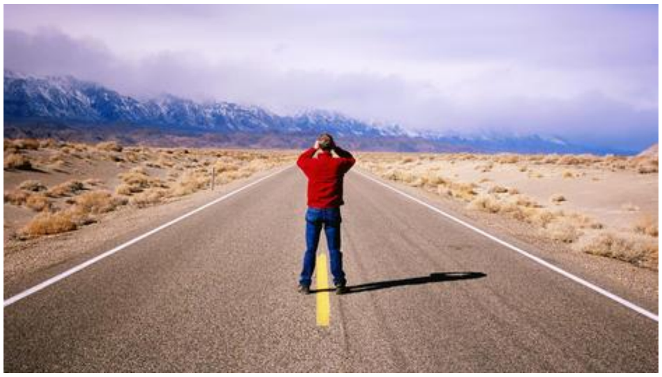

Figure 1: Input Image

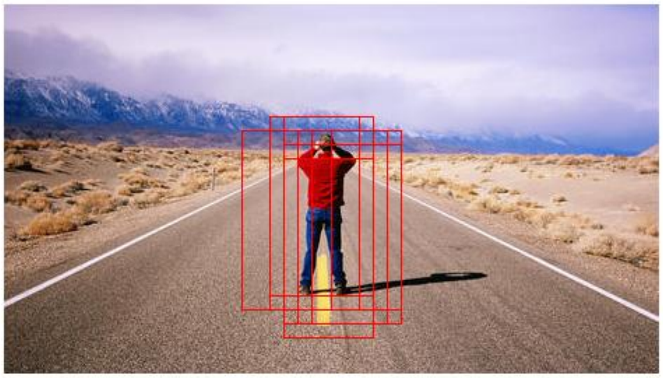

Figure 2: Output without Suppression

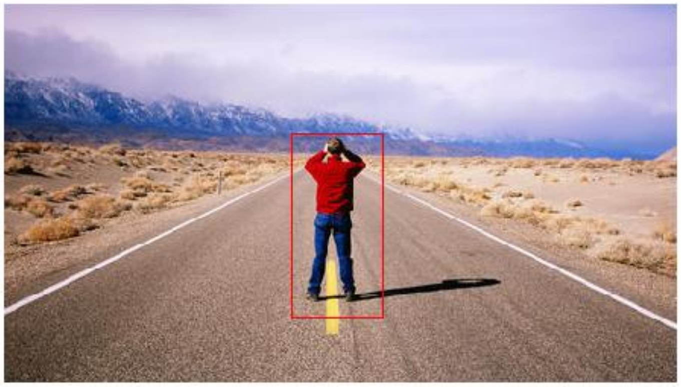

Figure 3: Output with Non Maxima Suppression

Trial 2: In this trial a slightly challenging image was chosen. If observed, we see two persons, one in the background and one in the foreground. The foreground person is too large to be detected confidently using a 64x128 window. Scale space has to be used in this case to satisfactorily detect the foreground person. It can however be observed in Figure 5 that there is some positive detection of the foreground person in case of lower pyramid levels (indicated by the smaller rectangles). But the maximum confidence score was obtained at a higher pyramid level as shown by Figure 6(Observe the bigger rectangle). The background person was however detected in the lower scale.

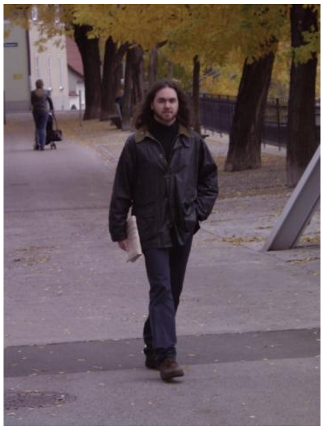

Figure 4: Input Image-2

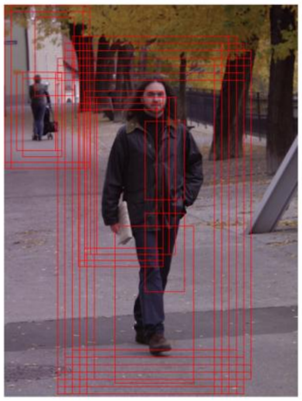

Figure 5: Output without Suppression

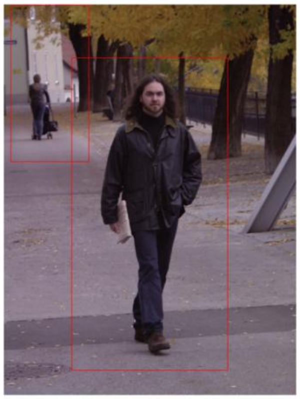

Figure 6: Output with Non Maxima Suppression

Trial 3: In this trial an even more challenging image was chosen. The input had multiple persons and was from a different viewpoint. The method satisfactorily detected all persons with reasonable accuracy. Figure 7 shows the input image, Figure8 shows output without suppression and Figure9 shoes output with suppression. The working of the scale space based peak confidence detection is clearly visible in this example.

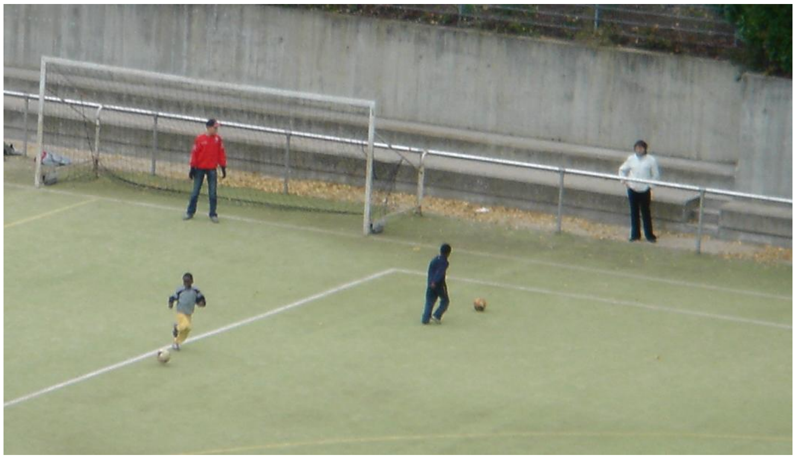

Figure 7: Input Image

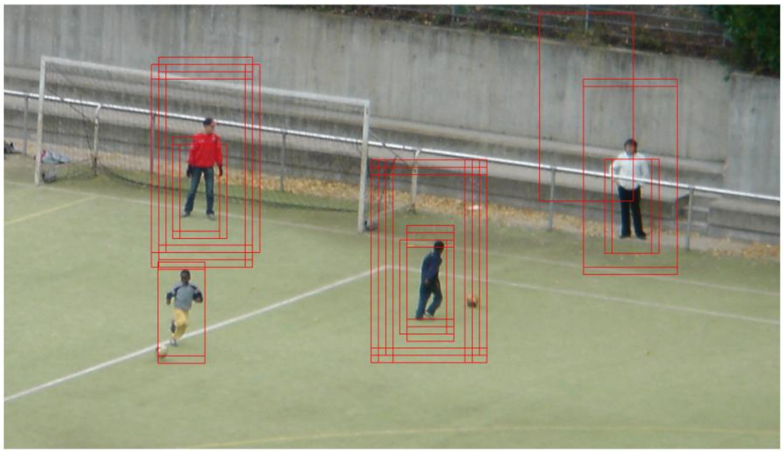

Figure 8: Output without suppression

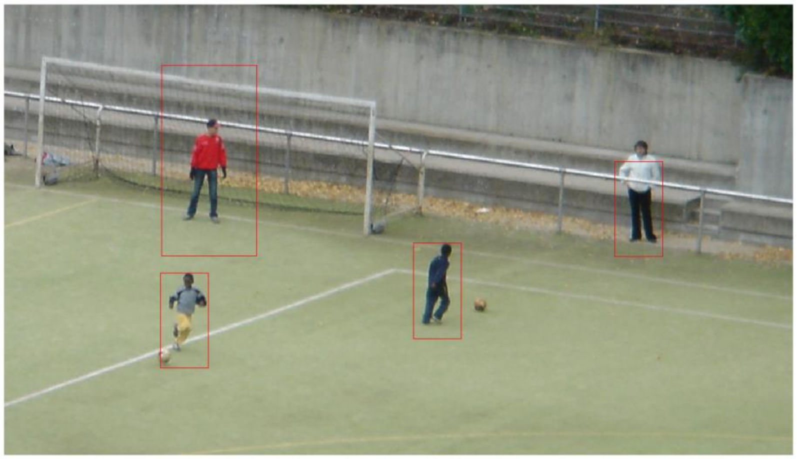

Figure 9: Output with Non Maxima Suppression

Trial 4: The image shown in Figure10 was chosen for this trial. This was to test the algorithm in the presence of varying background illuminations. Notice the illumination changes in the background due to sunlight variations. The method was not sensitive to illumination changes in the background.

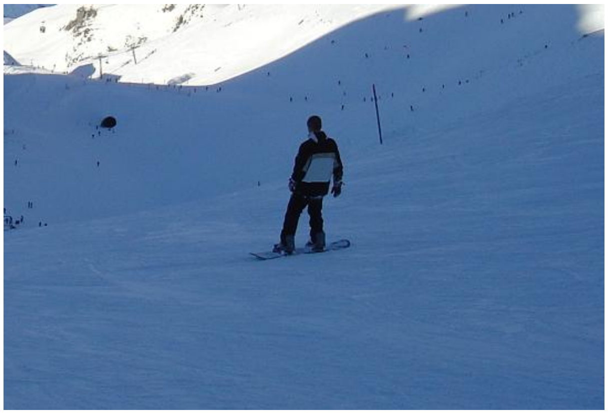

Figure 10: Input Image

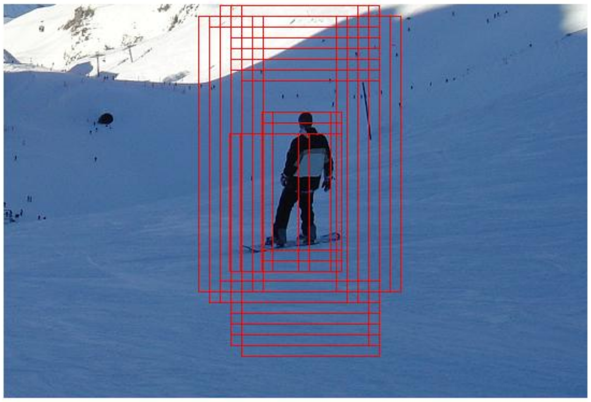

Figure 11: Output without Suppression

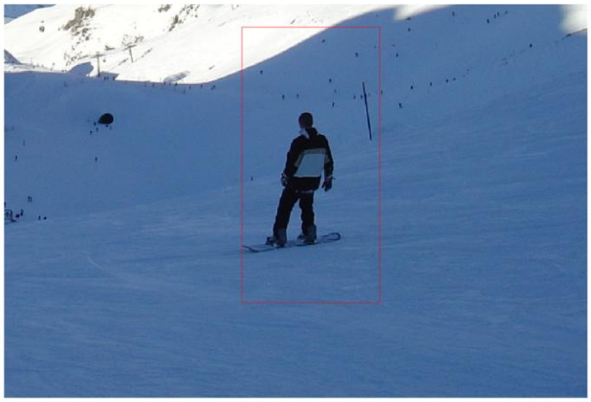

Figure 12: Output with Non Maxima Suppression

Trial 5: In this trial, the image is of a person facing back with two bags in both hands. This was particularly chosen to test the algorithm with variations in the appearance of a person especially when holding a bag and facing backwards. The background in this image is not very simple. The results are shown in the following figures.

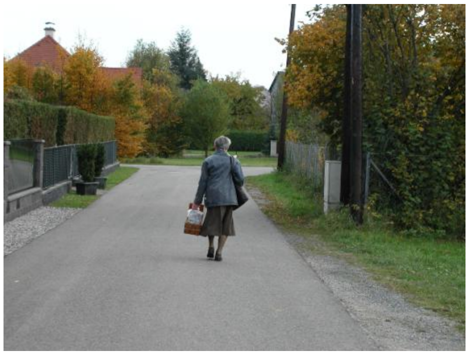

Figure 13: Input Image

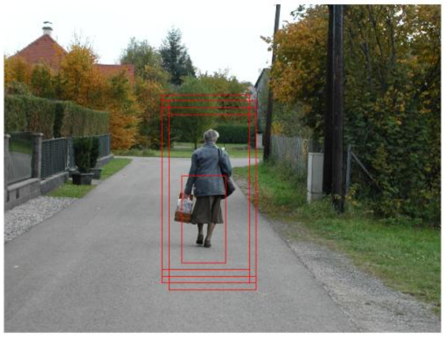

Figure 14: Output without suppression

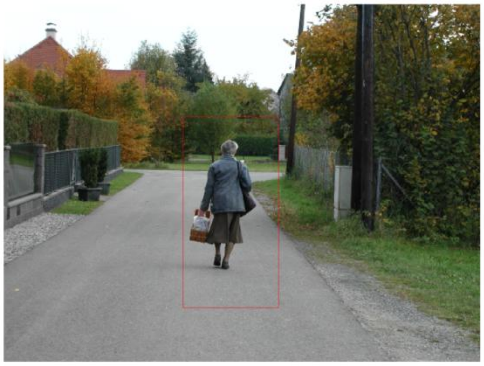

Figure 15: Output with Non Maxima   Suppression

Negative Examples

The following figures show the output for negative testcases.

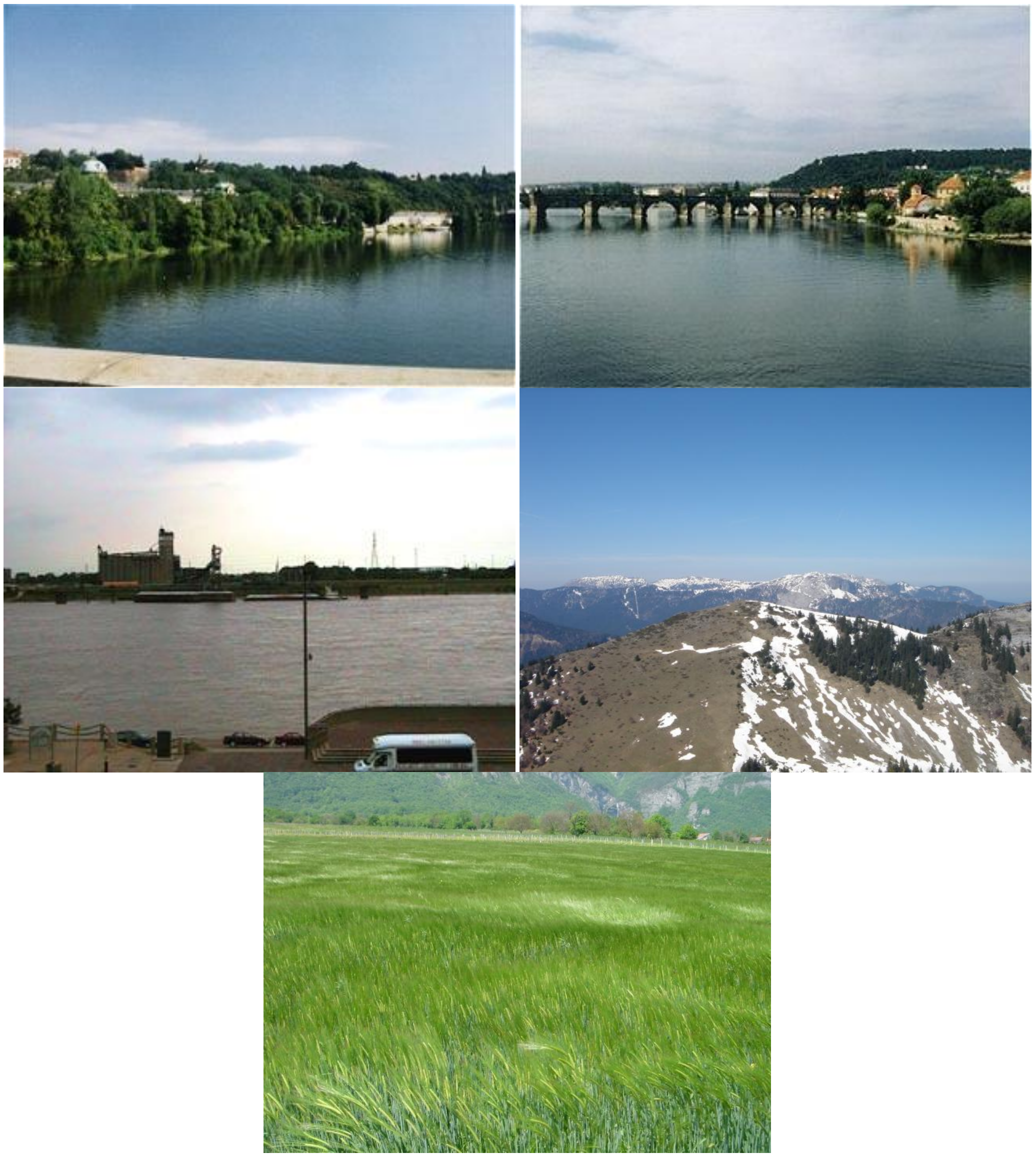

Figure 16: Input images for negative testcases

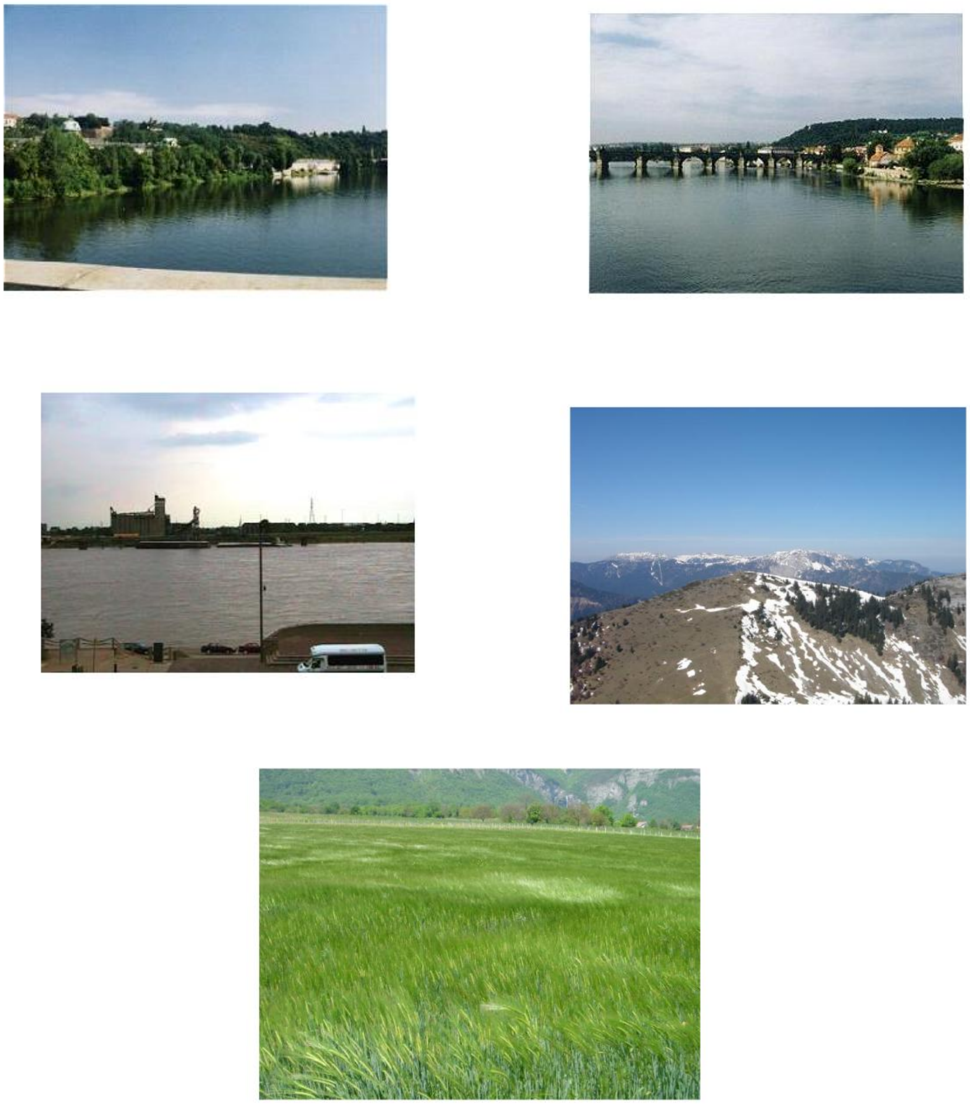

Figure 17: Output images for negative testcases. No bounding box is seen

Failure cases:

False Positives:

Figure 18 shows a case of false positive detection from the algorithm when the threshold is made zero for the second input image.

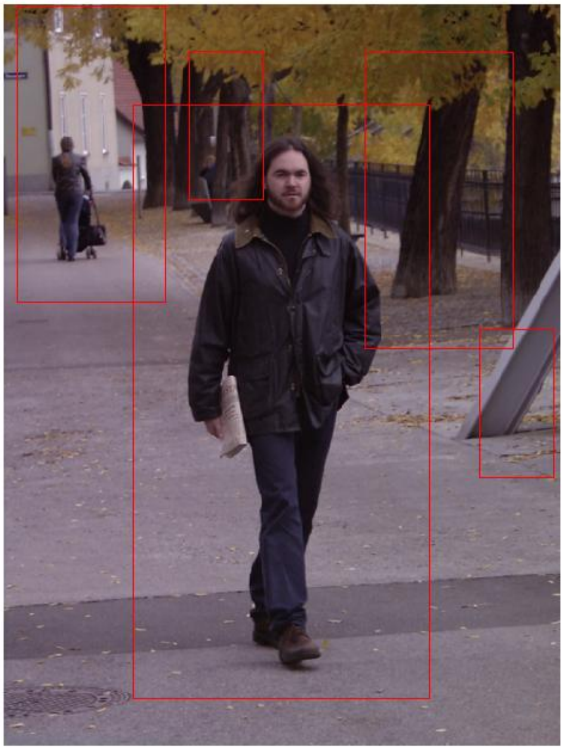

Figure 18: False Positive Example

Notice that the false positives are near the tree trunks which appear upright similar to human beings. This may be probably due to the fact that Negative samples used for training the SVM did not capture this minor anomaly of upright trees as negatives. To improve the generalization, the negative training data must be updated with such images of upright objects such as trees and poles, which would generalize the classifier to humans and pedestrians only. This was not done in the current implementation due to memory constraints of the system being used.

False Negatives:

Figure 19 shows a case of false negative.

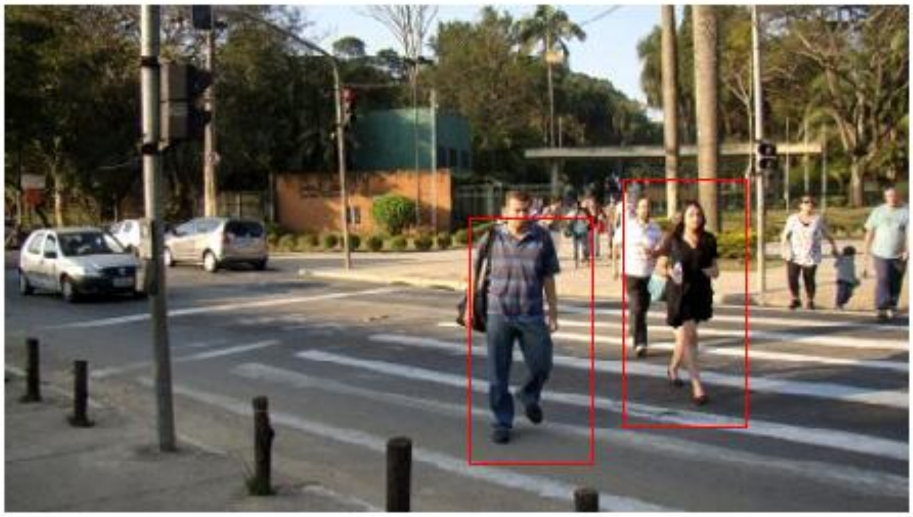

Figure 19: False Negative Example

In this case, the humans in the background were not detected leading to a case of False Negative. This is probably due to the fact that the resolution of the image is not sufficiently large. There was some amount of detection (the three people on the right side of the image) however, when the image was resized to a larger resolution, although there was some blurring observed due to resizing. The algorithm was not able to detect people who were further deep in the background. This could be due to loss of details due to blurring when resizing.