Subverting Face Detection with Haar Cascade Detector

Pooja Vijayakumar
University of Southampton
pv1n21@soton.ac.uk

1. Introduction

Haar Cascade detector makes use of Haar-like features and cascade classifiers for object detection. However, this algorithm is not able to detect object with complete accuracy and is prone to false positives and false negatives. This coursework focuses on how this algorithm can be subverted in the task of face detection. Various camouflage techniques were implemented in order to subvert the algorithm from detecting face on images.

2. Viola-Jones face detection algorithm

The Viola-Jones algorithm first consists of Haar-like features. These works in similar way to convolution kernels except that the values in convolution kernels will be learnt during training while in rectangle features, they are fixed. Using each rectangle feature on an image result in one value that will be obtained by subtracting sum of black pixels and sum of white pixels. Figure 1 shows the example of the rectangle features shown relative to the enclosing detection window.

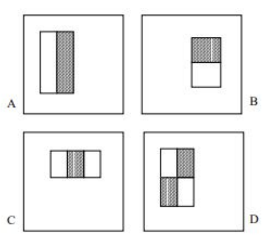


Figure 1. Adapted from [1]

Obtaining subtraction of sum of pixels especially with a stride of 1 used by the rectangle features may result in too many features being present and some of the rectangle features used may not even be relevant to a particular window. Hence Adaboost is used to find the relevant features. Weak learners are made after each of the feature are trained on the image, Each feature will be the weighted combination of all the feature. In order to determine if an object is present in a image window, the weighted combination of all the features will be used. This result in reduction of 160000+ features to 6000 features when training the algorithm to detect faces [2].

Computing the sum when each feature is slid over the image in each window can be time consuming as there are too many features present. Hence, the algorithm also makes use of the integral image. Using this technique the sum over any window will be calculated using only 4 array references [1]. The integral image image at location xy contains the sum of the pixels above and to the left of x,y inclusive [1].

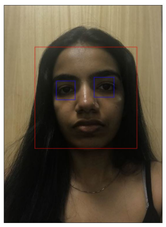
Finally, the algorithm also make use of the attentional cascade. In an image, there may be many regions where an object being detected will not be present. It will be time-consuming to pass each feature over a window that may not contain the object being detected. If the first feature detects an object at a window, then a second feature is evaluated on the window and so on [1]. If negative result was present with any of the feature evaluated, then the region will be rejected as potentially containing the object.

3. Subverting the Algorithm

The following section describes 3 methods used to subvert the face detector. Changes were made to the brightness and contrast of images and different camouflages were applied to the face.

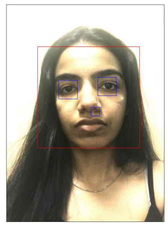
3.1. Adjusting the brightness of image

Figure 2 shows the effect adjusting image brightness. Image 1 shows the original image. The face and eyes were detected accurately. In image 2, the exposure and brightness of image were decreased until the face appeared darker. This caused the region around the cheeks below eyes, while visible appear darker. The region around the eyes and cheeks now has darker pixels. Hence, there may be no edge detection between the area of eyes and cheeks causing the eyes to not be detected. There is very little to no difference in pixel intensities between the area of eyes and cheeks. In image 3, the brightness and exposure were increased until the background washed out. There is false positive detection of nostril as eyes in the image. This may be because high exposure caused nostril to be appear darker and more prominent due to shadows around the upper lips as the surrounding pixel brightness were increased. This may have been mistakenly detected by haar-like features such as WBW rectangle feature as region separating the eyes and cheeks. In image 4, the brightness and exposure were increased further. The face was unable to be detected by the detection algorithm as the darker pixels that were present in the image becomes lighter. Hence, similar to the eyes in image 2, the rectangle features may not be able to detect facial features due to lack difference in pixel intensities in the image.

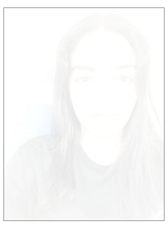


(a) image 1

(b) image 2



(c) image 3



(d) image 4

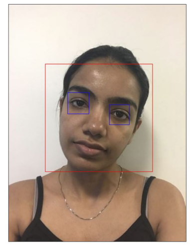
Figure 2. Effect of adjusting image brightness

3.2. Camouflage with face mask and shower cap

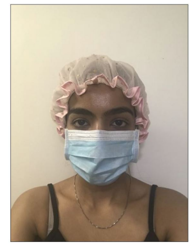
In Figure 3, image 1 is the original image without any camouflage. The face and eyes were detected correctly. In image 2 and 3, a shower cap and face mask were added to camouflage the face at different angles. The face was not detected at both the angles. This may be because all prominent features that makes up the face such as the lighter cheeks regions were covered and hence the haar-like features are not able to detect any changes in pixel intensities usually found on facial features. The covered regions which form sub-windows on the image may be identified as non-face regions by the first rectangle feature. Due to the mechanism of attentional cascade employed, this may cause the sub-windows to be be rejected as containing a face. Since most of the regions of the face are covered, the detection algorithm rejects most of the sub-windows of the images as containing face outright causing the overall face to not be detected.

3.3. Camouflage with googles, scarf and cap

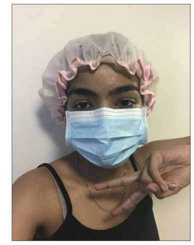
Figure 4 shows 3 images of face camouflage using a cap and googles. A scarf was wrapped around the neck. The pictures of the camouflage with the face were taken at different angles. The face was not detected in all 3 images. This may be because prominent features such as the forehead, the eyes and the bridge of the nose was completely covered by the darker cap and googles. Hence, there were no difference in pixel intensities between the darker and lighter regions at these sub-windows to be detected by the haar-like features. The neck area around the chin being covered by the scarf may also be preventing edge detection around the jawline by the rectangle features. For example, the BWB rectangle feature may not be able to detect the area on the nose bridge between the two eyebrows due to the presence of googles.



(a) image 1



(b) image 2



(c) image 3

Figure 3. Effect of Camouflage with face mask and shower cap

(a) image 1

(b) image 2

(c) image 3

Figure 4. Effect of Camouflage with googles and cap

4. Conclusion

There are many different methods used to camouflage the face or manipulate images in order to trick the algorithm from detecting eyes and faces. From the above methods, it can be seen that the algorithm can be subverted easily with occlusions on the face even though the face is still detectable by another human. Furthermore, the algorithm may also wrongly detect other areas of the face as eyes. This shows that the Viola-Jones algorithm while effective, does not produce results with perfect accuracy. Hence, other algorithms such as You Only Look Once (YOLO) or Single Shot Detector (SSD) algorithm can be used that makes use of convolutional neural networks for object detection with high accuracy.

References

[1] P. Viola and M. Jones, ”Rapid object detection using a boosted cascade of simple features,” Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

[2] “OpenCV: Face Detection using Haar Cascades,” docs.opencv.org. `https://docs.opencv.org/4.7.0/d2/d99/tutorial_js_face_detection.html` (accessed Mar. 21, 2023).