Face Detection Algorithm using Histograms of Oriented Gradients

Luisa Fernanda Borbón R.
Universidad de los Andes
Departamento de Ingeniera Biomdica
lf.borbon@uniandes.edu.co

Abstract

Detection problems are an important computer vision are that consist in finding the the instances of an specific category inside an image. In this case, a PHOG algorithm was implemented in a sample of the Caltech WebFaces dataset in order to learn to detect human faces. Using multi-scale HOG algorithm with the best hyper parameters, an average precision of 81.16% was obtained for the test set. Finally a more expressive descriptor is recommended to increase even more the accuracy of the algorithm, for example including colour.

1. Introduction

Detection is an important task of high-level computer vision. This area of study has the objective of finding the positions in an image that correspond to the instances of an specific category [5]. The algorithms developed on this matter are used in applications such as image retrieval, security, surveillance and advanced driver assistance systems. Due to the range variety of applications in this topic, there have been several models developed for detection like the Viola-Jones algorithm, image segmentation and blob analysis, gradient and derivative based approaches, deep learning object detection, feature based object detection and template matching [4].

One common approach to detection is using an sliding window to scan patches of the image individually and decide if they contain the object of interest as an image classification problem [6]. In order to discriminate from the different instances of the image, a robust feature descriptor should be used, which differentiates the object of interest from others. One of the most famous descriptors in this matter is the histogram of oriented gradients (HOG). This approach is based on the fact that an objects appearance and shape can be characterized by the distribution of local gradients and edge directions in order to detect it [2].

In this sense, the present work intends to apply a multi-scale HOG approach to detect human faces from a subset of the Caltech WebFaces dataset. This was done by using the PHOG baseline prepared by James Hays for computer science 143 class at Brown University [3].

2. Materials and Methods

In order to detect the human faces form the subset of Caltech WebFaces images and extra test images, as mentioned before a PHOG baseline was used. On the following sections, the dataset and the methodology followed are described in detail.

2.1. Dataset

The Caltech Web Fasces dataset contains 10,524 human faces images taken from the web by typing common names into Google Image Search. In this dataset, annotations are provided with the coordinates of the eyes, nose and center of the mouth for each frontal face are provided in a ground truth file [1]. In this work, a subset of 6,713 cropped faces images of resolution 36x36 were used as training and 130 images were used initially for testing. Finally, in order to test the algorithm in more natural pictures, 20 images taken by the author were also used; a sample of each set of images is shown in figure 7.

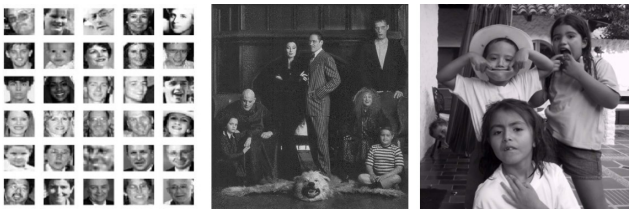

Figure 1. On the left, a sample of some of the 36x36 training images used, in the middle one of the initial test scenes and in the right a sample of the extra scenes added by the author.

2.2. Multi-scale HOG

The detection strategy was based on James Hays Face Detection with a Sliding Window baseline, created for computer science 143 class at Brown University, which consists

on implementing a PHOG algorithm, originally developed by Dalal and Triggs [2].

First, the positive and negative instances are used to train a HOG descriptors that are going to identify objects that have similar or different shapes to it. This is done by dividing each image into cells, making the convolution of each patch with different border filters and normalizing them to get the histograms of oriented gradients descriptors for each one cell. This is done with both positive and negative instances and the different HOG descriptors are classified using an SVM classifier that will discriminate between face or not, according to their shape [2].

The library used in order to run the different features like the hog strategy was the open source library VLFeat, that has different computer vision algorithms to implement. Also, this process is combined with spatial pyramid images in order to provide the multi-scale classification and find different size human faces. This is expected to work on detection challenges because objects tend to have similar shapes despite their illumination, position and sizes.

In this case, the main parameters that can be taken into account in the PHOG detection algorithm are the number of positive and negative instances used to train the model, the weigths and confidence of the SVM classifier, the scales, step and sizes taken into account for the sliding windows, the spatial pyramid and the threshold value after which a possible detection is considered a real detection. This hyper parameters need to be tuned and defined according to the images and their relative sizes.

2.3. Evaluation

In order to evaluate a general detection problem, different evaluation approaches can be taken like Precision and Recall curves, Average Precision (AP) and the F-value. In this case, the evaluation was done with a function inside Hay's algorithm which calculates the false positives, false negatives and true positives in order to build the associated PR curve and get a precision measure.

3. Results

The pyramidal HOG descriptors were created

After this, taking into account the corresponding labels and tuning the hyper parameters of interest, the PR curve shown in 3 was obtained, with an overall average precision of 81.6%.

Finally, the algorithm was tested on the additional testing images as its shown in figures 4, 5 and 6,

4. Conclusions

The multi scale HOG strategy presents as an effective method for human face detection, due to the fact that they have really similar shapes and edges. Because of this, the results obtained with the algorithm were significantly good as expected. On the other hand, the basic limitation of the

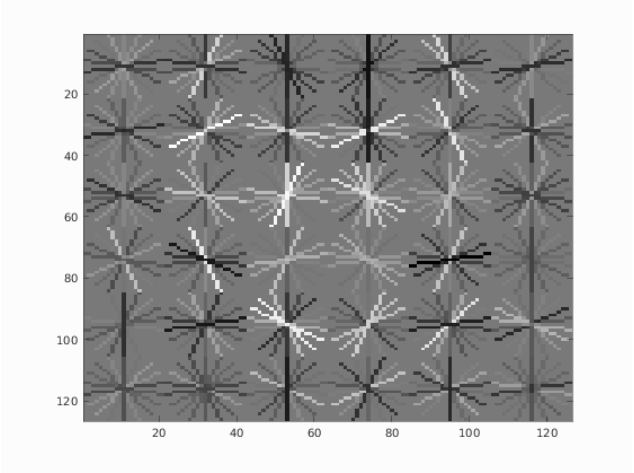

Figure 2. Sample of a human face HOG descriptor.

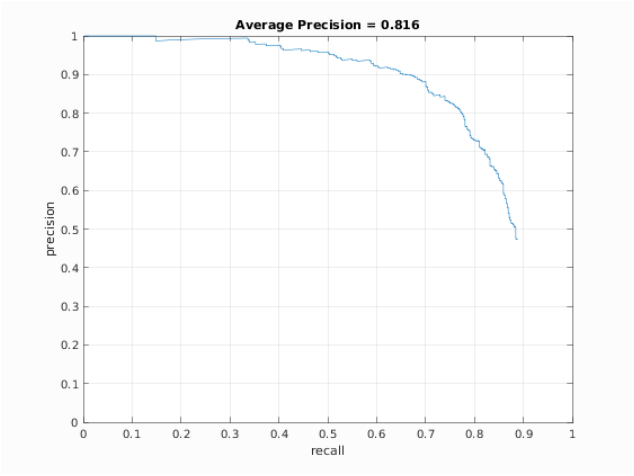

Figure 3. Precision and Recall Curve obtained in the test images with the best hyper parameters.

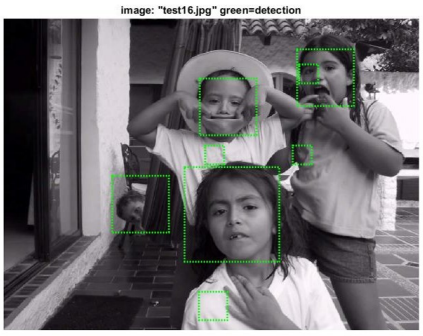

Figure 4. Extra Test result for image 'test16'

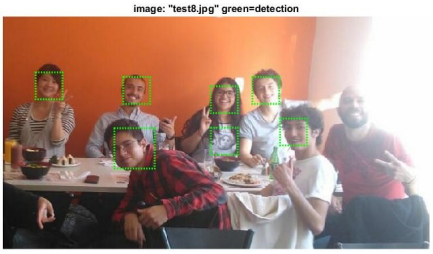

Figure 5. Extra Test result for image 'test8'

Figure 6. Extra Test result for image 'test4'

multi-scale HOG algorithm are objects with similar shapes, because specially when working with small HOG descriptor sizes the shape can be pretty similar between a face and circular objects.

For further work, a wider descriptor could be implemented using color. In the case of human faces, as human skin is pretty easy to recognize it could give valuable information between faces and objects with similar shapes. The other limitation are the changes in the direction of viewing of the face, which can lead to a different shape and could be solved by training different types of HOG descriptors with faces that have different viewing positions.

References

[1] CaltechVisionLab. Caltech 10, 000 web faces, 2003. [Online] http://www.vision.caltech.edu/ImageDatasets/Caltech10KWebFaces/Description.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *international Conference on computer vision & Pattern Recognition (CVPR'05)*, volume 1, pages 886–893. IEEE Computer Society, 2005.

[3] J. Hays. Project 4: Face detection with a sliding window, 2013. [Online] http://cs.brown.edu/courses/cs143/2013/proj4/.

[4] MathWorks. Object detection in computer vsion, 2019. [Online] https://la.mathworks.com/discovery/object-detection.html.

[5] Y. Pang, Y. Yuan, X. Li, and J. Pan. Efficient hog human detection. *Signal Processing*, 91(4):773–781, 2011.

[6] W. Zhang, G. Zelinsky, and D. Samaras. Real-time accurate object detection using multiple resolutions. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8, Oct 2007.

5. Extra

Figure 7. Waldo was found on the image *13_Interview_Interview_On_Location_13_558*