High level Content-based image retrieval

Silvério Pereira, 76505
Pedro Santos, 76532

Introduction

Development of an application to do image retrieval by extracting high level information from the images.

The application is divided in two parts:

- Indexing set of images
- Search by image /test to get the results
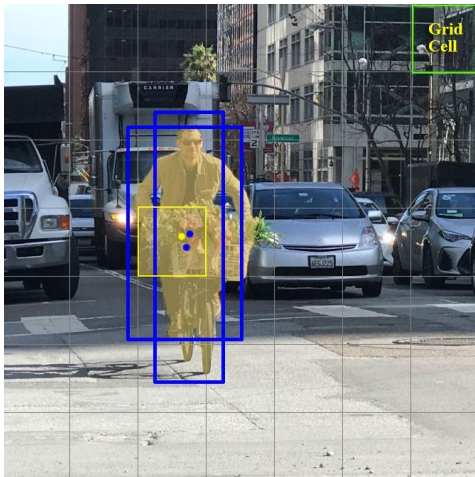- Rank the results

What do we use to score images?

- Object detection
- Visual Saliency
- Face recognition
- Object Character recognition

Yolov3

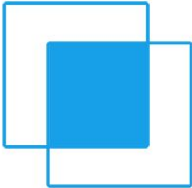YOLO (You Only Look Once) is a state of the art object recognition algorithm.

- Divides the image in grid cells and applies a CNN classifier to each, each grid cell has a maximum of objects it can predict.
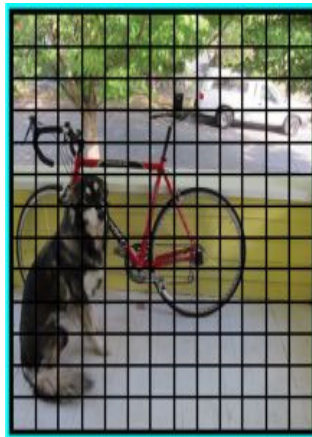
$$\widehat{y} = \begin{bmatrix} P_c \\ b_x \\ b_y \\ b_y \\ c_1 \\ c_2 \\ \dots \\ c_n \\ P_c \\ b_x \\ b_y \\ b_y \\ c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix}$$

1 anchor box

Non-max suppression

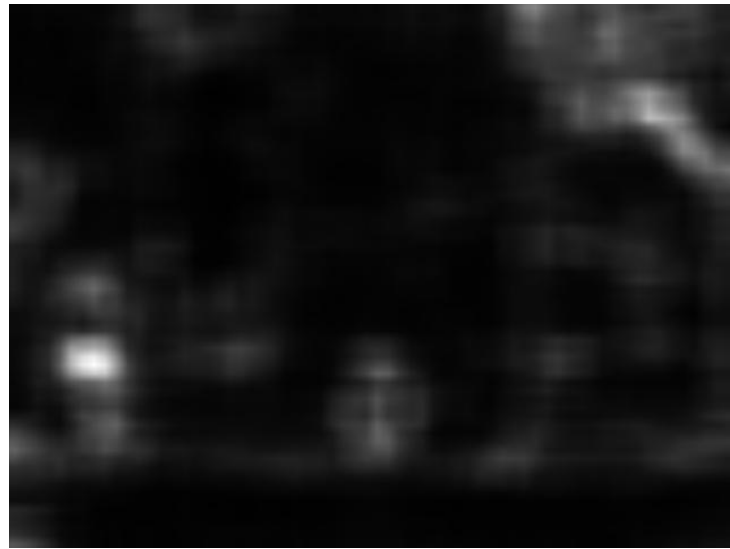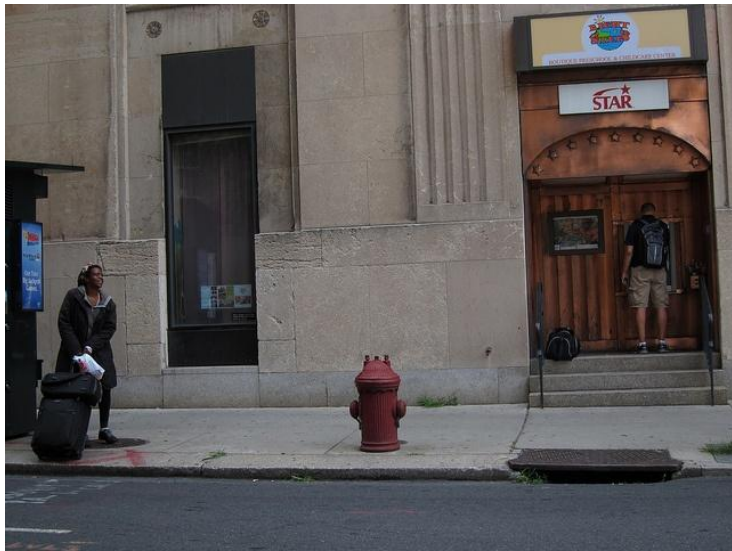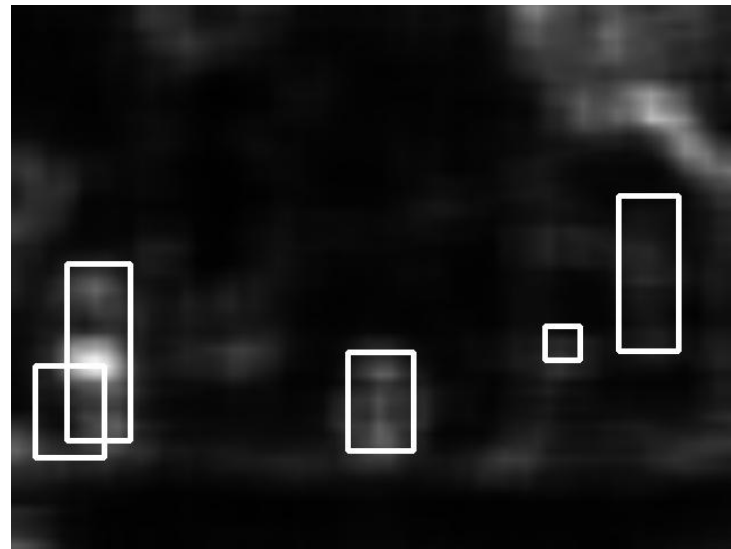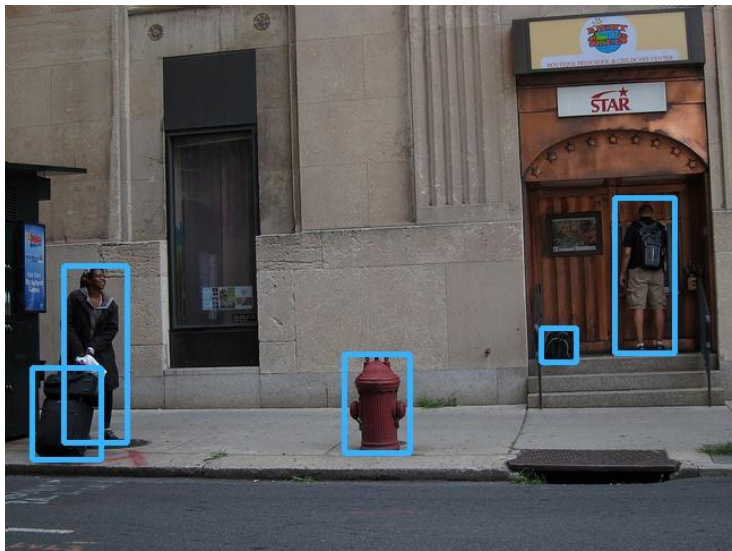$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

Tunable parameters

- Confidence threshold
- Non-max suppression threshold (IoU)
- Max number of objects in a cell grid

Visual Saliency
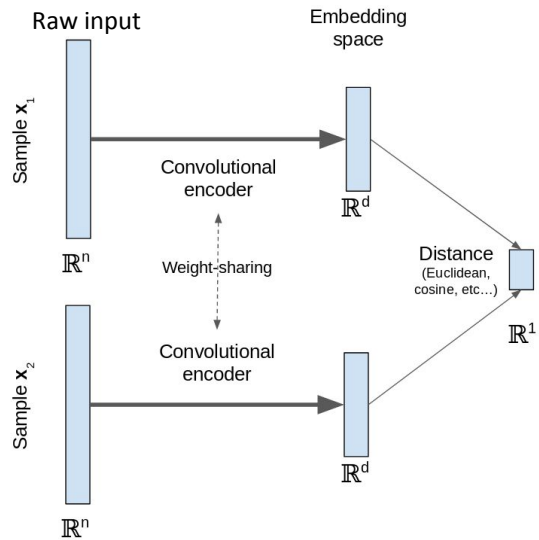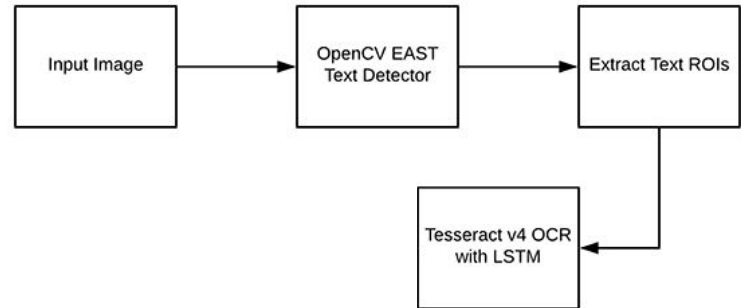
Relevance Calculation

One Shoot Face Recognition

OCR (optical character recognition)

- The function verifyText() is called and it returns a list with the words detected in the given image;

- Verification if it was detected any text in the image;

- Removing Special Characters;

- Performing Stemming;

Transfer Learning

- It is not easy… it is necessary to have some installed dependencies and to do and manipulate configurations files.
- To be able to see the results of the configuration of the network is necessary to wait several hours.
- Collect the labeled data
- The Darknet implementation is used to train the data using the GPU
- It is necessary to have two files, one with the paths to the images of the training set and another to the test set.

Transfer Learning

- Somo lines in file detector.c were changed to save the intermediate weights more frequently.

- Download the pre-trained model from darknet

- Create some files to configure Darknet and set the architecture of the network

```
rute@rute-X550JK:~/Documents/cadeiras/vc/progs/darknet$ ./darknet detector train
 obj.data cfg/yolov3-tiny.cfg backup/yolov3-tiny_5000.weights > /home/rute/Docum
ents/cadeiras/vc/vc1819-76505-76532/project/proj1/training/log.txt
```
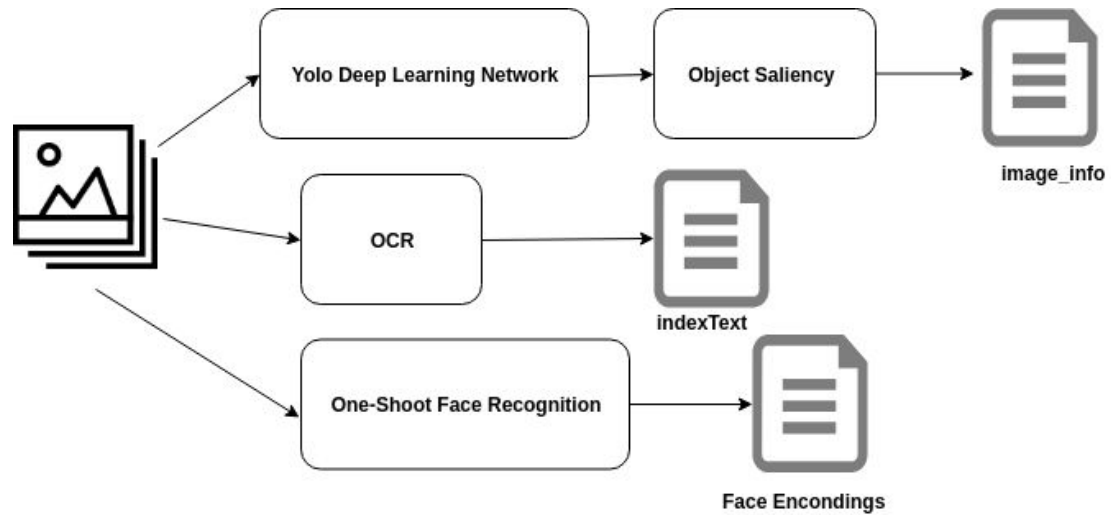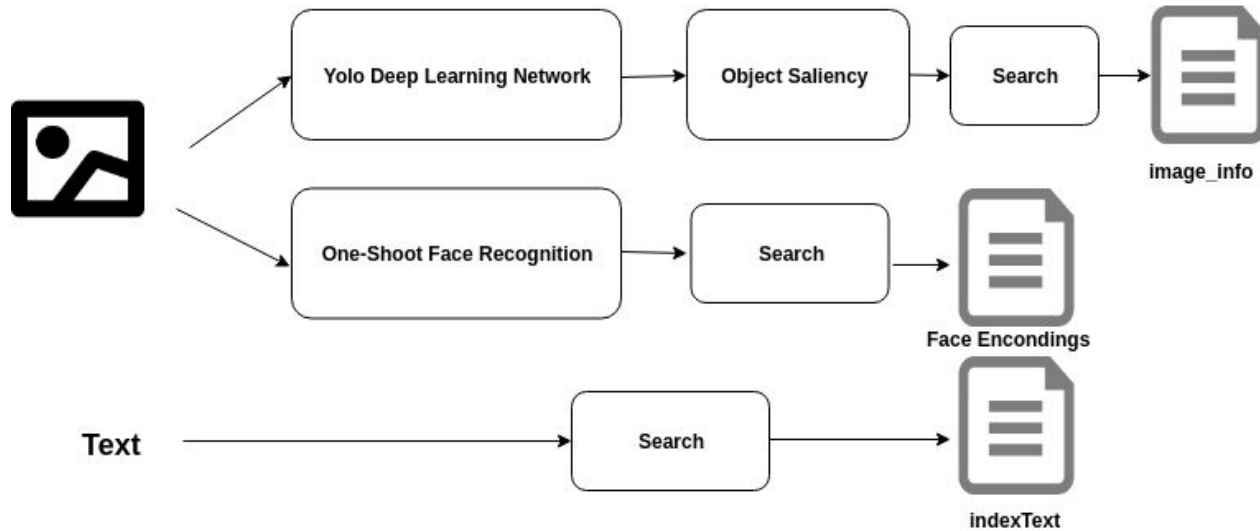
Transfer Learning

Transfer Learning

Indexing Data flow

Search Data flow

Ranking

TABLE I
OBJECT WEIGHT CALCULATION

y	$\hat{y}$	cost
0	0	0
0	> 1	$log(1 + \hat{y}) * e^{rel(\hat{y})}$
> 1	0	$1 + log(y) * e^{rel(y)}$
> 1	> 1	$|log(y) - log(\hat{y})| * e^{|rel(y) - rel(\hat{y})|}$

TABLE II
FACE DISCOUNT CALCULATION

$d <= 6$	-10
$0.6 < d < 1$	$min(log((d - 0.6) * 2.5)), -10)$
$d > 1$	0

Results

Query
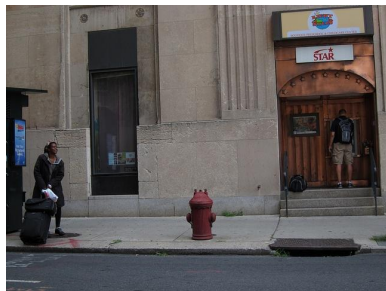
Results

Query: "nokia" →

Query →