

YOLOv3 Analysis

Introduction

Yolov3 is an algorithm for object detection which is much faster and more accurate than other detection methods.

Bounding box prediction

Each object will contain a bounding box and the updated network produces 4 parameters for a box t_x, t_y, t_w, t_h .

If offset coordinates of top left corner is (c_x, c_y) and the width and height of prior box are p_w, p_h then the prediction coordinates are-

$$b_x = c_x + (\sigma) * t_x$$

$$b_y = c_y + (\sigma) * t_y$$

$$b_w = p_w * (\exp(t_w))$$

$$b_h = p_h * (\exp(t_h))$$

YOLOv3 detects an object using logistic regression(It is a statistical method to predict output when input is categorical. It uses logistic function $(1/(1+e^{-x}))$ to predict the object of a bounding box)

Class Prediction

To identify from which class the object belongs.

From the bounding box we get some classes for an object using multi label classification.

Now we predict one class using **logistic prediction**.

Why don't we use softmax prediction:- because it will fail when one object belongs to

multiple classes like . An object belongs to both women and person.

Prediction scale and layers formation

YOLOv3 follows a method for final bounding box prediction.

Initially it predicts 3 boxes , then upsample them with 2 layers previous and estimate one box using k-means clustering(dividing n particles into k parts with similar properties. Suppose 100 particles. 30 are red, 20 are black and rest blue, so $k=3$)

Layers formation:- The new network is a combination of previous 3 methods YOLOv2,darknet-19 and one new method.

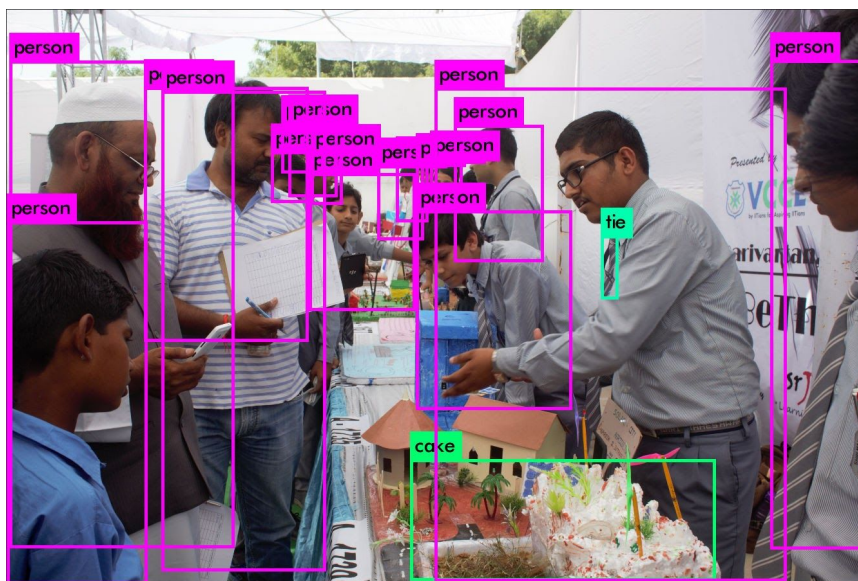
This method produces 53 convolutional layers, so the method is called darknet-53.

Training

Train using some variables like batch normalisation, data analysis, etc.

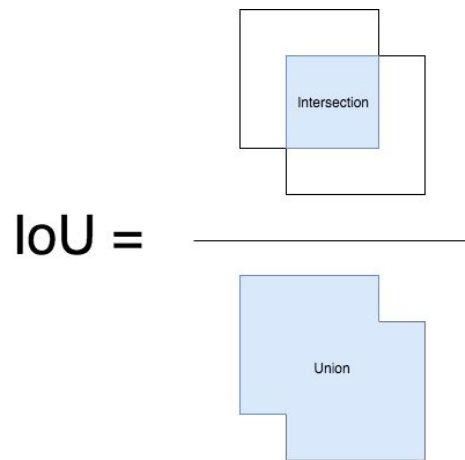
Advantage of YOLOv3

1. faster and accurate than other methods.
2. Working on small objects also. This image is predicting Tie.



Disadvantage of YOLOv3

1. Bad performance in medium and large objects.
2. When we increase the threshold it detects only larger objects.



IOU threshold

Intersection over Union

Training is done using a fixed threshold value.(0.5)

Implementation

1. Counting the number of tigers in a park.
2. Counting the number of students in a classroom.
3. Counting number of people without masks in a crowd during COVID-19 outbreak.

© Jay Prakash and Rushil