IET SMP 2021-Computer Vision

Welcome Guys!!!

Instructors:

M.T.Aakash Chandra-2nd Year CSE

Email :aakash.191cs132@nitk.edu.in

Contact:9110272655

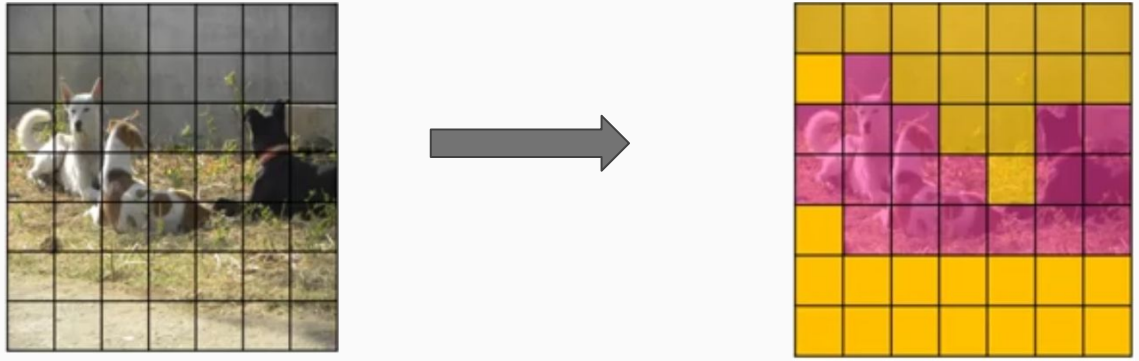
Aprameya Dash -2nd Year IT

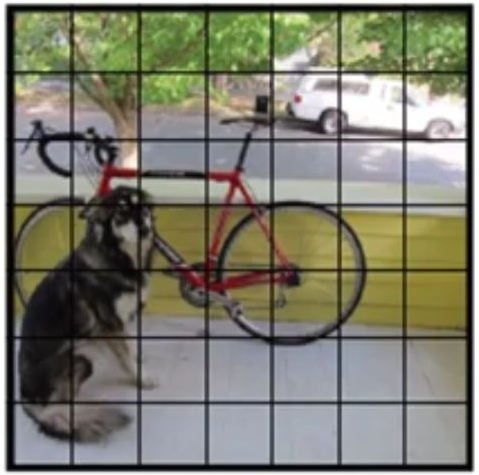
Email: aprameyadash.191it209@nitk.edu.in

Contact:7735784564

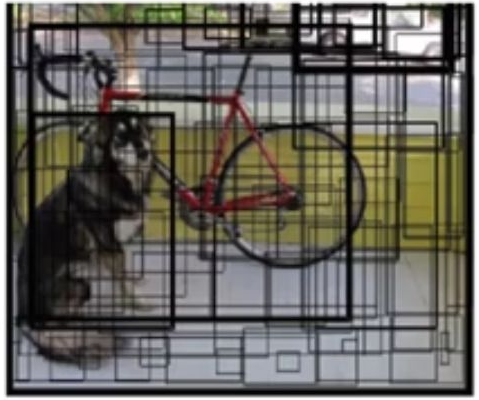
YOLO:You Only Look Once

- YOLO (“You Only Look Once”) is an effective real-time object recognition algorithm, first described in the seminal 2015 paper by Joseph Redmon .
- In object detection we basically require to find a bounding box around the object and classify on what object or what class of objects the detected object belongs to.
- Thus each bounding box can be described as (p_c,b_x,b_y,b_w,b_h,c)
- Basic Idea: Divide input into grids of equal size sxs. Predict a class and bounding box of objects each grid location either fully or partly.



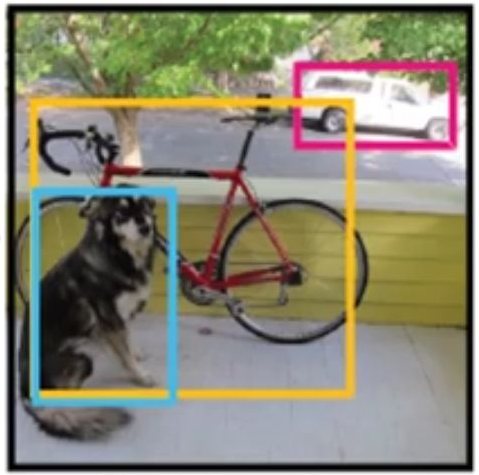


S × S grid on input



Bounding boxes + confidence

Class probability map



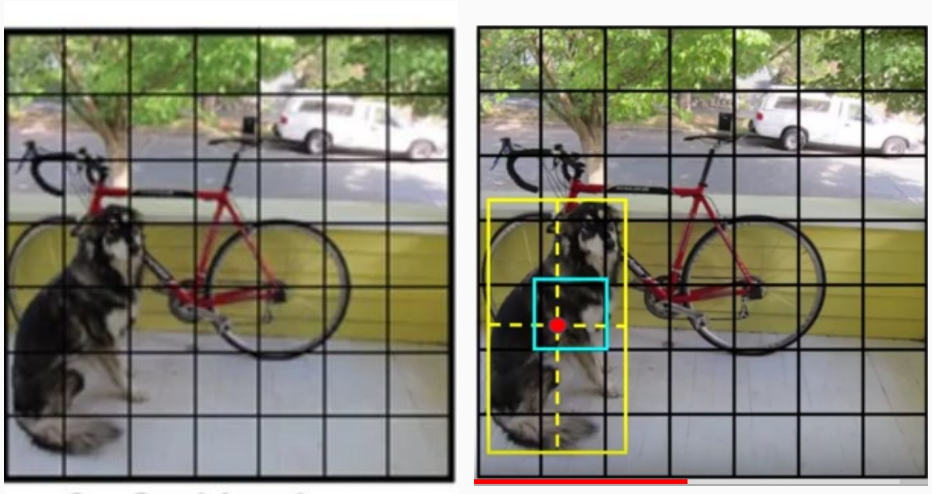
Final detections

Redmon et al. CVPR 2016.

https://arxiv.org/abs/1506.02640

YOLO Algorithm

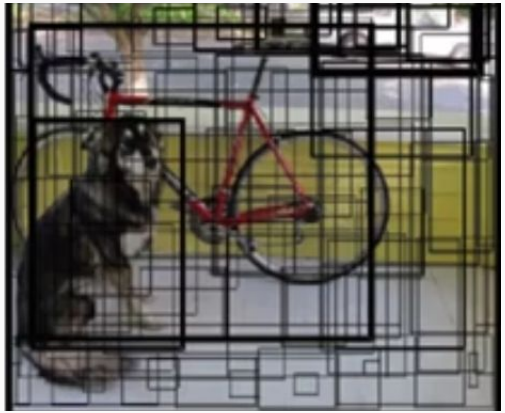
Split the image into 7x7 grids.If a midpoint of an object(ground truth bounding box) falls into a particular grid cell, then this grid cell is responsible for detecting the object.



YOLO Algorithm contd

Each grid cell predicts a **B** number of bounding boxes in a particular cell with a confidence score for each of these boxes.

Each cell predicts the class probability.

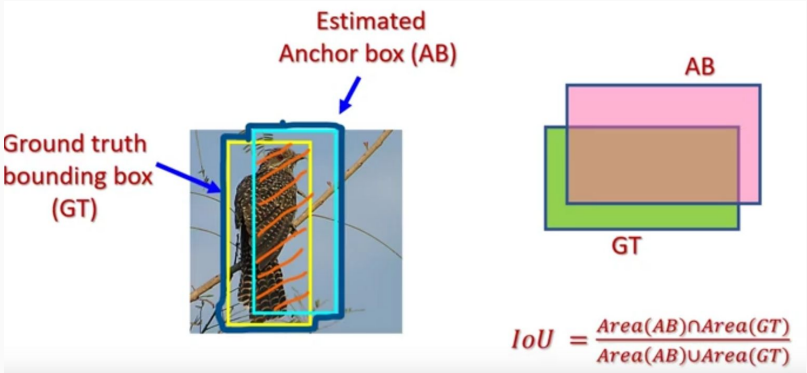


Dealing with redundant anchor boxes

This is dealt used two stage filtering. First set a threshold on confidence of a box detecting a class.Select only one box when several boxes overlap with each other and detect the same object

Box confidence can be defined as **Probablilty x IoU**

IoU-Measure of overlap between ground truth bounding box and predicted bounding box.



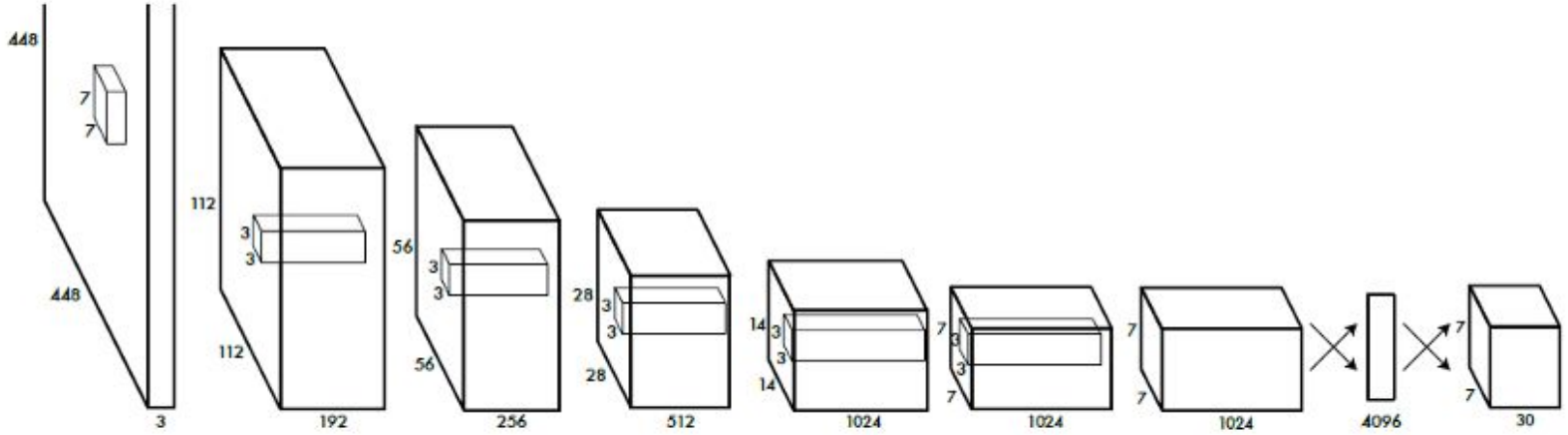
Dealing with redundant boxes contd.

Remove all boxes whose scores are less than threshold.

Non Max Supression: Select the box with highest score. Compute its overlap with other boxes and remove boxes that are really close to it until we get a single box.

In YOLO, this ground truth is taken as the grid with maximum probability.

This thus gives us the required bounding box and its coordinates .


448
7
7
448
3
112
3
3
112
192
56
3
3
56
256
28
3
3
28
512
14
3
3
14
1024
7
3
3
7
1024
7
7
1024
4096
7
7
30
Conv. Layer
7x7x64-s-2
Maxpool Layer
2x2-s-2
Conv. Layer
3x3x192
Maxpool Layer
2x2-s-2
Conv. Layers
1x1x128
3x3x256
1x1x256
3x3x512
Maxpool Layer
2x2-s-2
Conv. Layers
1x1x256
3x3x512
×4
1x1x512
3x3x1024
Maxpool Layer
2x2-s-2
Conv. Layers
1x1x512
3x3x1024
×2
3x3x1024
3x3x1024-s-2
Conv. Layers
3x3x1024
3x3x1024
Conn. Layer
Conn. Layer

Project

We shall work with a pre trained object detection YOLO model. We shall work with YOLO V3. The dataset we will be working with is the COCO Dataset.

The COCO dataset has the following classes as shown in the link below”

Check the Link:https://gist.github.com/AruniRC/7b3dadd004da04c80198557db5da4bda

Let’s head over to Google Colab

Thank You Guys!!