PAGE
DATE

→ State of Computer Vision

→ (week3) Object Detection
1- identify object
2- localize

→ object localization
. inorder to design object recognition we must learn about object localization.
. if we want to localize the object in image we modify the neural network to have output units the boundey box $b_x, b_y, b_h, b_w$
($b_x, b_y$ → midpoint; $b_h$ → height; $b_w$ → width)
. localize focuses determing the positions of objects in image

→ land Mark detection
. in other cases NN have xy cordinates, which are important point in image these points called landmarks we want the model to recognize
. used in Snapchat filters & Posture detection

→ object detection
. we will take about using convnet for obj detection using something called sliding window.
. for exi: Car detection only. first: we need to create a label training set with closely croped x & y examples of cars

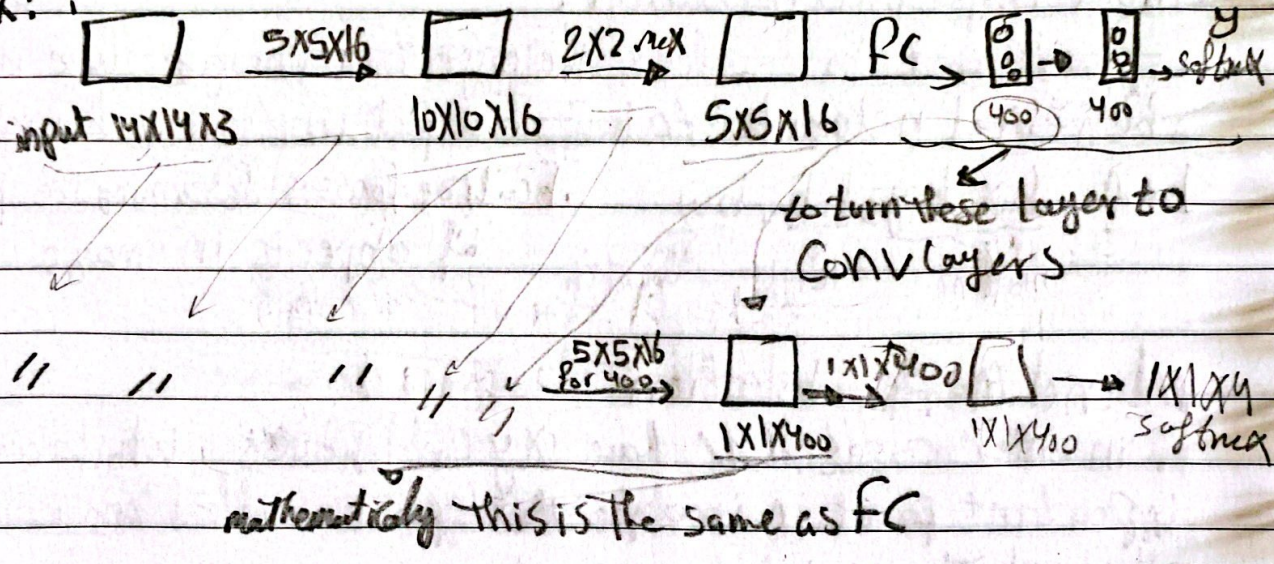

second train Convnets using this set to output 0 or 1 if there is car or not

Third use's slide window by picking certin window size and apply it to the image then input it to conv net that make prediction

PAGE
DATE

→ Convolutional implementation of sliding windows.
• we saw that sliding window was to slow & has a high computional cost, how can we implement it convolutionaly
• let see first how can we transform a fully connected layer of neural network into conv layer

ex:




mathematicaly this is the same as FC

→ intersection over union (IOU)
• we use it to evaluate obj. detection alg.

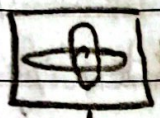
$$= \frac{\text{size of intersection}}{\text{size of union}} \quad \text{if iou} \geq 0.5 \text{ :)}$$

→ non-max suppression
• its a way to make sure that the alg. only detects obj once
• each output $\begin{bmatrix} P_c \\ b_x \\ b_y \\ b_h \\ b_w \end{bmatrix}$ → Discard $P_c < 0.6$
↳ picks the box with the highest $P_c$ & output that as prediction

PAGE
DATE

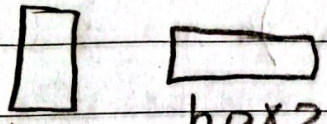
## anchor boxes

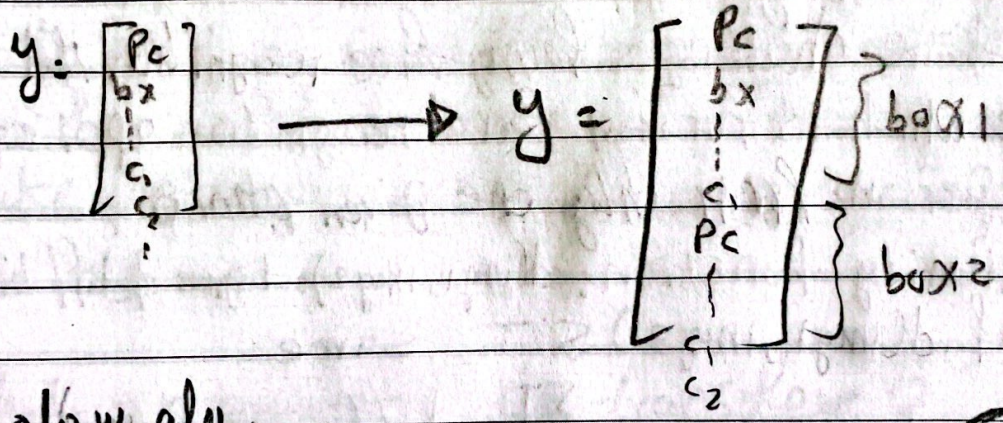
- if each of grid cells detect one object what if grid cell wants to detect multiple objects heres anchor boxes are used



overlapping objects

anchor boxes predfined two box shape




→ yolow obj.

if $y = 3 \times 3 \times 2 \times 8$

anchors, num of classes