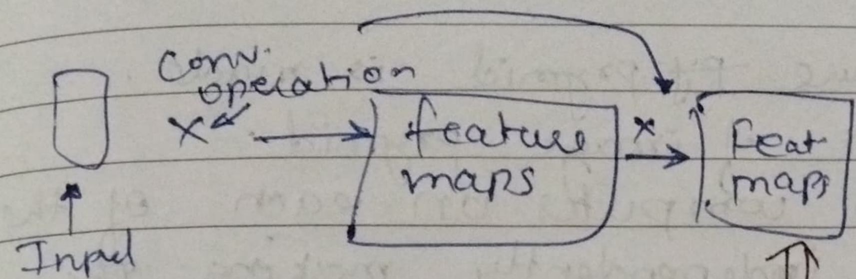
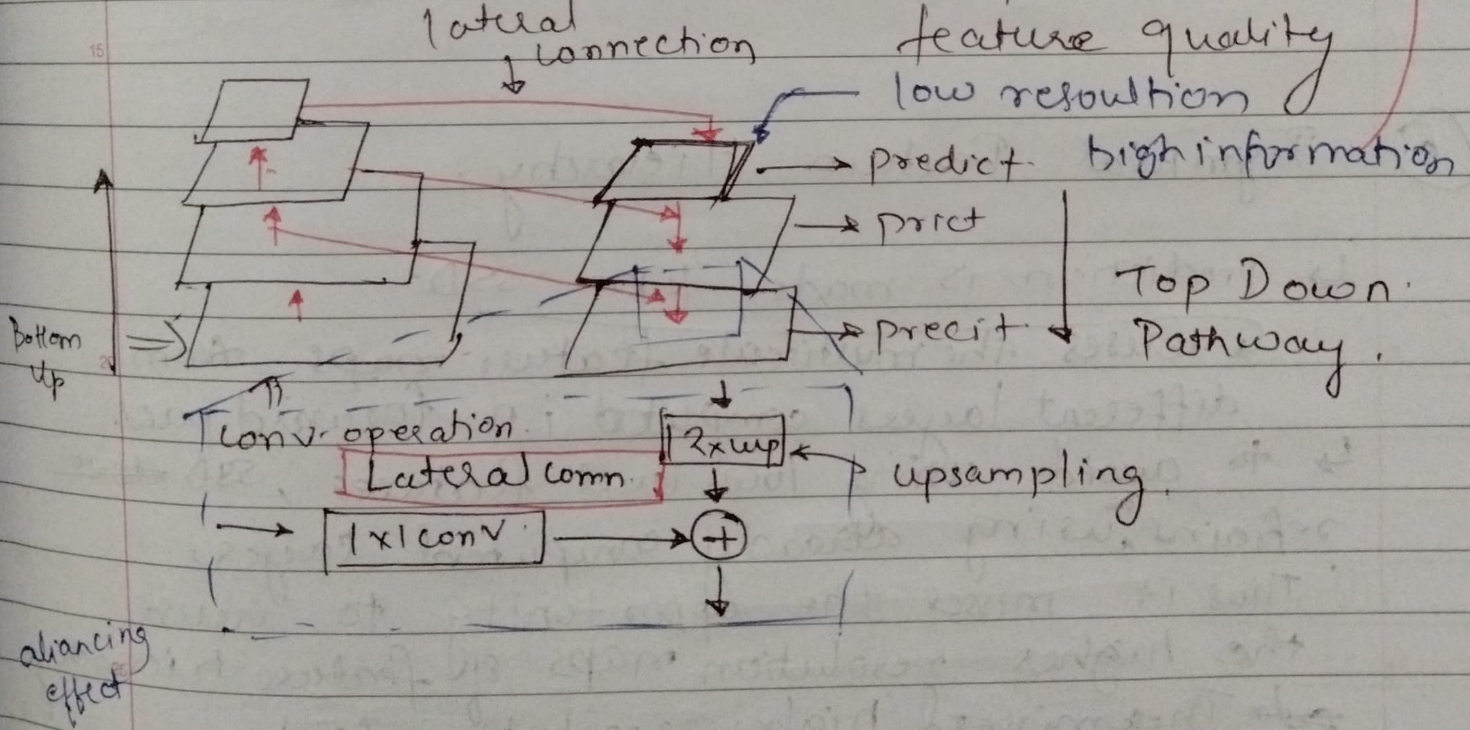
feature Py Prymid Netroork (FPN)



we merge feature maps from this layer with its previous feature maps to improve the feature quality

the o/p feature maps are of low resolution



Date / /

(a) featurized image prymid.

A feature ~~py~~ prymid is made with the use of image prymid. features are computes on each of the image scale independently making the process slow.

(b) Single feature map.

→ simple Convolutional operation on a sing image. and the predictions are at the end.

(c) Prymidal Feature Hierarchy.

↳ prediction is made like SSD.

↳ reuses the multiscale feature maps from different layers computed in forward pass.

↳ to avoid using low-level features, SSD ~~does~~ refrains re using already computed layers. Thus it misses the opportunity to reuse the higher resolution maps of feature hierarchy

↳ These missed higher resolution feature maps are necessary for detecting small objects.

(d) Feature Prymid Network.

↳ Combines low resolution, semantically strong features with high resolution, semantically weak features via top-down pathway

Neurons/units in a hidden layer are connected to each other

and lateral connections.

Feature Pyramid Network:

Bottom-up Pathway:

↳ feed fwd ~~net~~ computation of the backbone ConvNet; which computes a feature hierarchy consisting of feature maps at several scales with a scaling factor of 2. i.e.

↳ feature maps undergo a 1x1 conv. operation that reduce the dimensions.

↳ O/P of last layer at each stage is choose as ref set of feature maps, used to enrich the ~~prymid~~ pyramid.

Top-down pathway.

↳ The top-down path experiences perception of higher resolution features by upsambed. spatially coarser, (feature maps of inferior quality) but semantically stronger feature maps from higher pyramid levels.

↳ The spatial resolution is upsampled by a factor of 2., using nearest neighbour upsampling for simplicity.

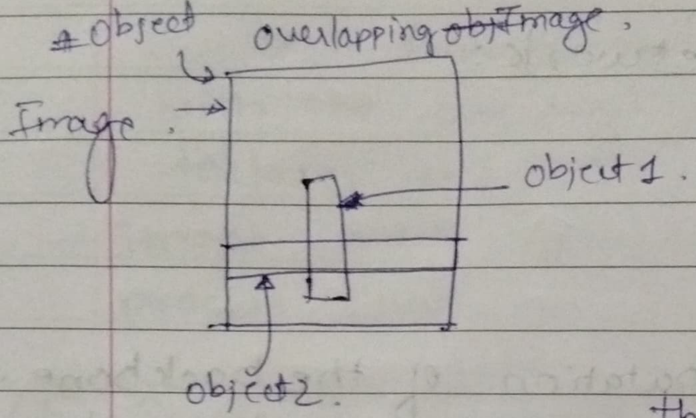
↳ lateral connections merge feature maps.

bottom up + top down Pathway

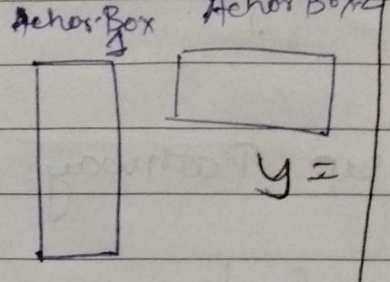
↳ by elementwise addition

Anchor boxes:

In case of overlapping objects.



we assume bboxes to be of this shape.



$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} \quad (3 \times 3 \times 8)$$

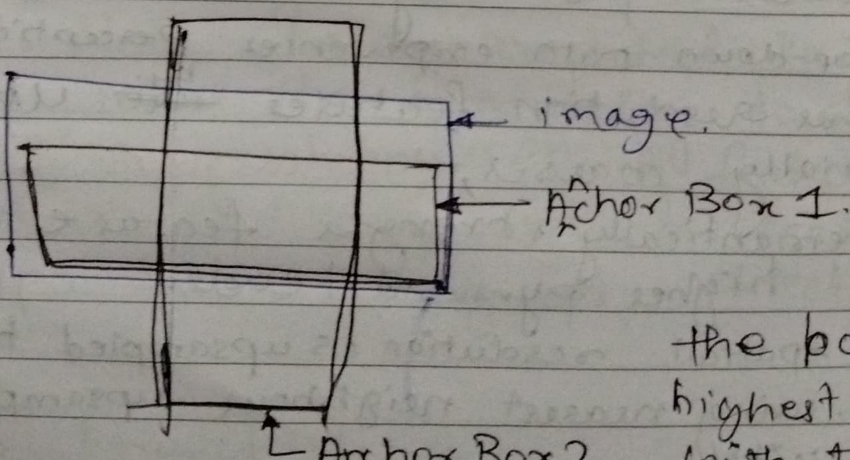
class $\{c_1, c_2, c_3\}$ — no of classes

the updated ~~an~~ 'y' vector becomes:

$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \\ p_c \\ \vdots \\ c_3 \end{bmatrix} \quad \begin{array}{l} \text{Achor box 1} \\ \\ \text{Achor box 2} \end{array} \quad (3 \times 3 \times 16)$$

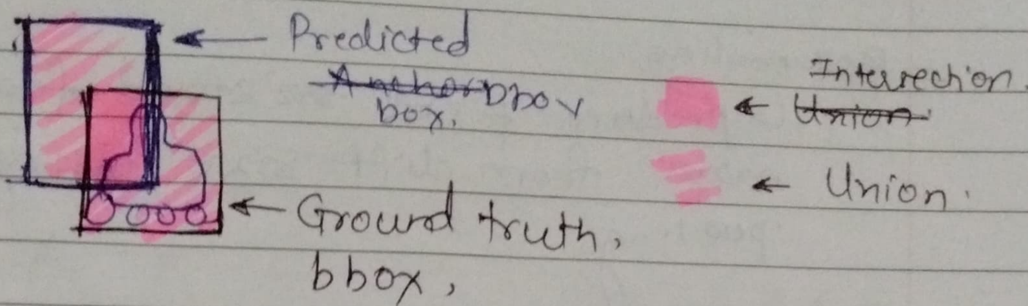
Now in this case it is diffcult to draw bbox for ~~predict~~ multiple objects in single image.

each object in training image is assigned to grid cell that contains object's midpoint and anchor box for the grid cell with highest IOU.



the box with highest intersection with the target object is the bounding box

Intersection Over Union:



Calculted the intersection & Union of both bboxes

IOU = $\dfrac{\text{size of}}{\text{size of}}$

"Correct" IOU ≥ 0.5

Perfect IOU = 1.

FPN for RPN ⟹ sliding window class agnostic object detector.

Original RPN.

⇒ a small subnet evaluated on 3x3 sliding window on a single scale conv. feature map.

⇒ performed with 3x3 conv ⇒ then two sibling 1x1 conv.

$\underbrace{\text{3x3 conv} \Rightarrow \text{then two sibling 1x1 conv}}_{\text{network head}}$

Researchers Approach

⇒ replce single-scale feature map with FPN.

⇒ a head of same 3x3 (conv & two sibling 1x1 conv) to each level on feature pyrmid.