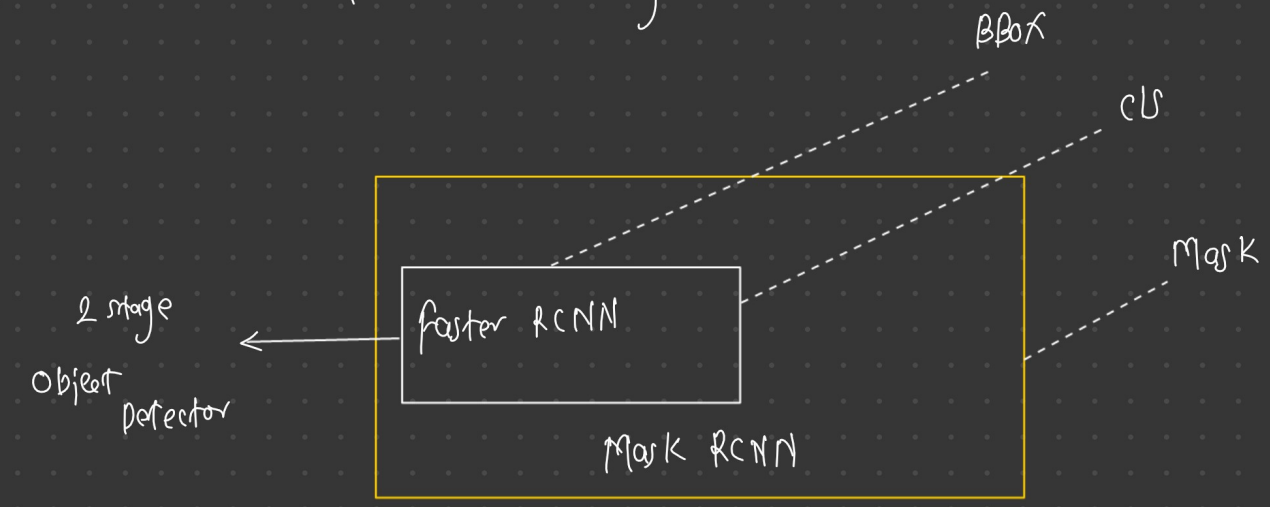
KRISHAI TECHNOLOGIES

Mask-RCNN

www.krishnaik.in

Mask RCNN

- It's a model for instance segmentation



Key improvements:

- Segmentation Branch
- RoiAlign

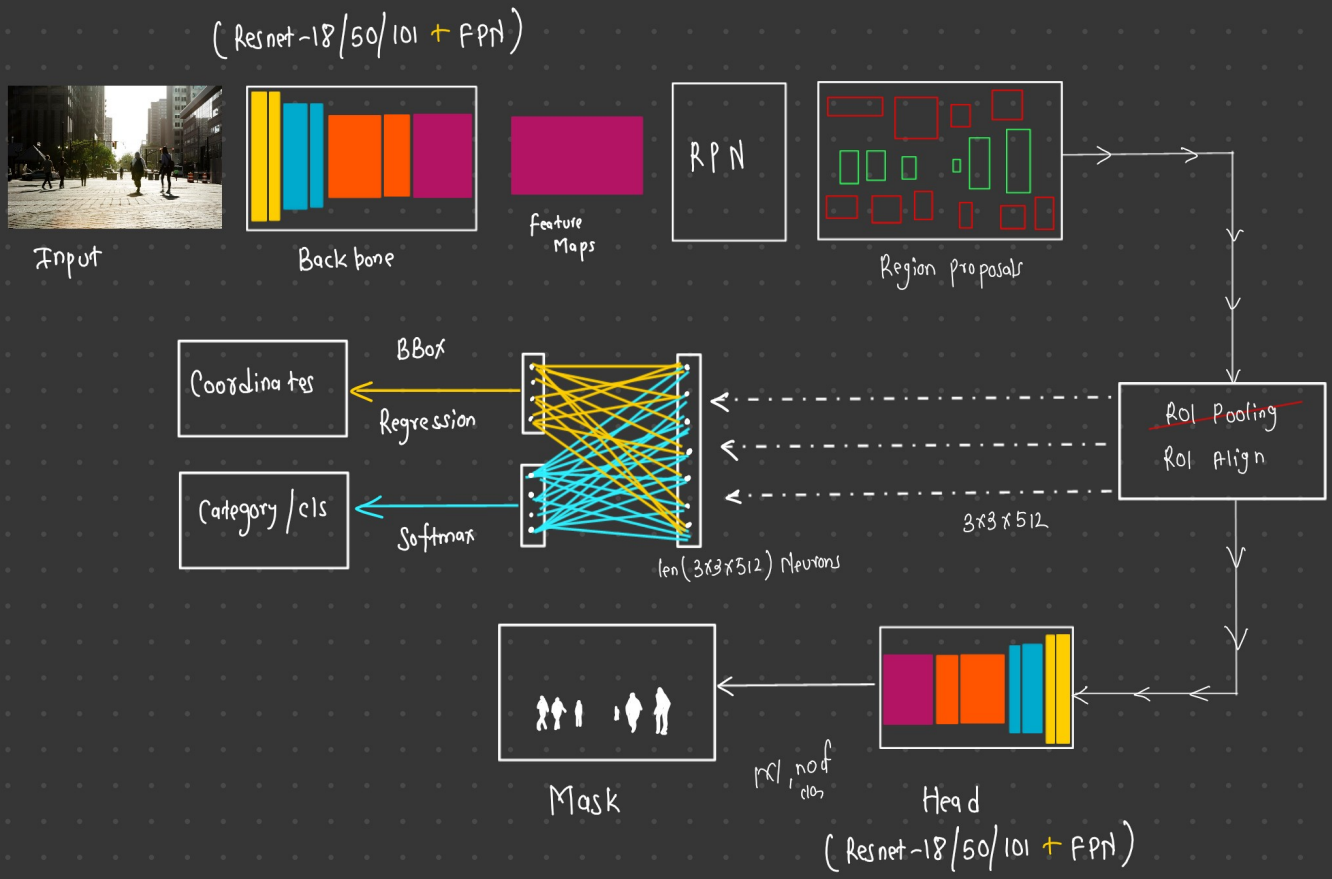
Applications:

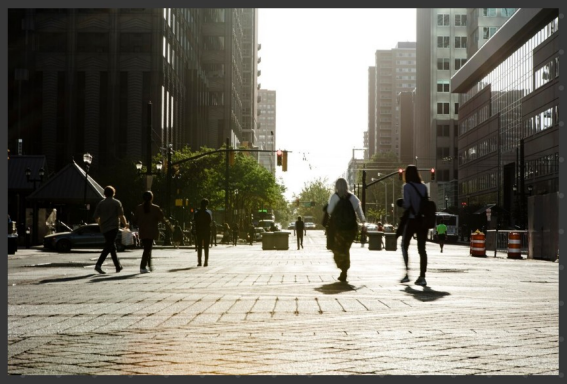
- Object Detection
- Instance Segmentation

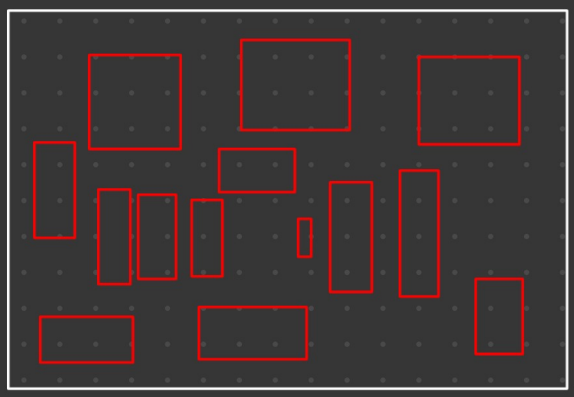
☆ Able to identify individual object

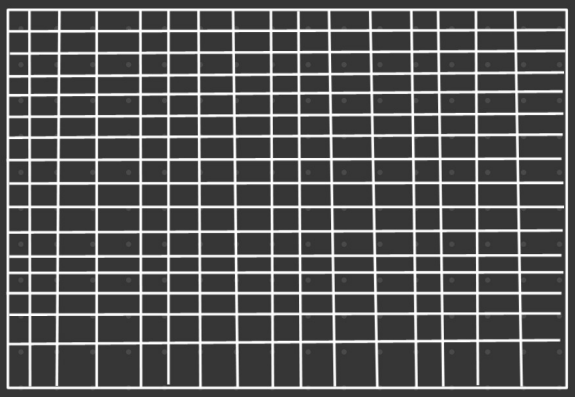
→ Car 1 , Car 2

Mask RCNN Architecture



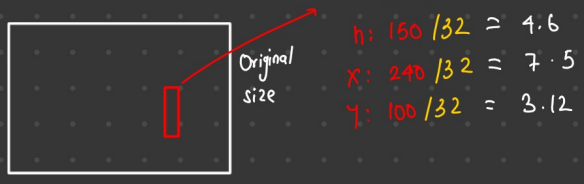


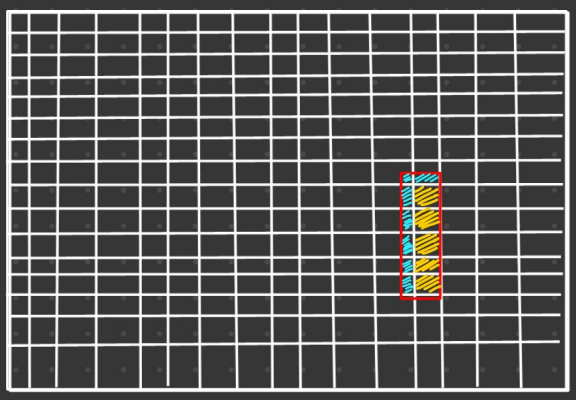




→ Downsample → 32X

Each pixel(1x1) in feature map (16x16)
sees 32x32 pixel of original image





(ROI Pooling)

We are working with 16x16 grid and we
can only consider integers

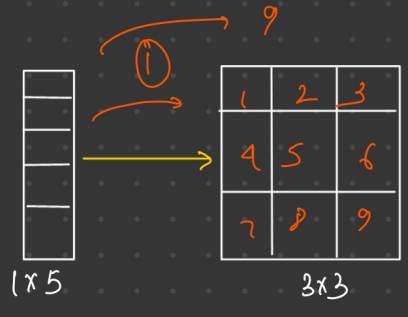
Last Data

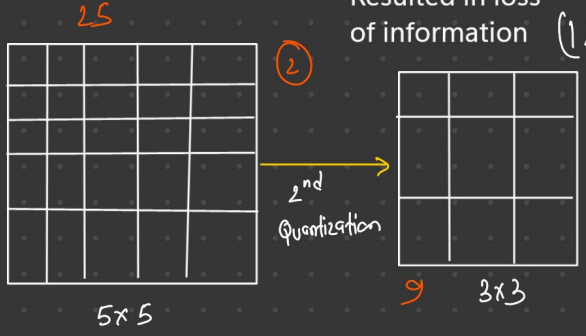
Next step : Quantization for roundoff

w: 80 /32 ≃ ~~2.5~~ 2
h: 150 /32 ≃ ~~4.6~~ 4
x: 240 /32 ≃ ~~7.5~~ 7
y: 100 /32 ≃ ~~3.12~~ 3

→ Resulted in loss of information (I)

ROI Pooling

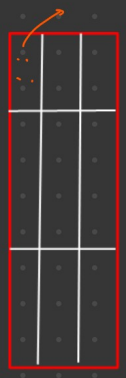




Resulted in loss
of information (II)

~~Step : Quantization for roundoff~~ No info loss

Step : Applies bilinear interpolation to calculate values using surrounding cells



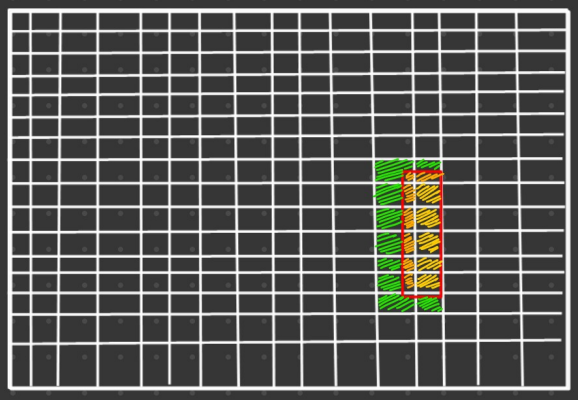
1. Divide ROI into 3x3 boxes

2. Create 4 sampling points for within each cell

3. Use bilinear interpolation for 1 of 4 points and use cell weights for calculation.

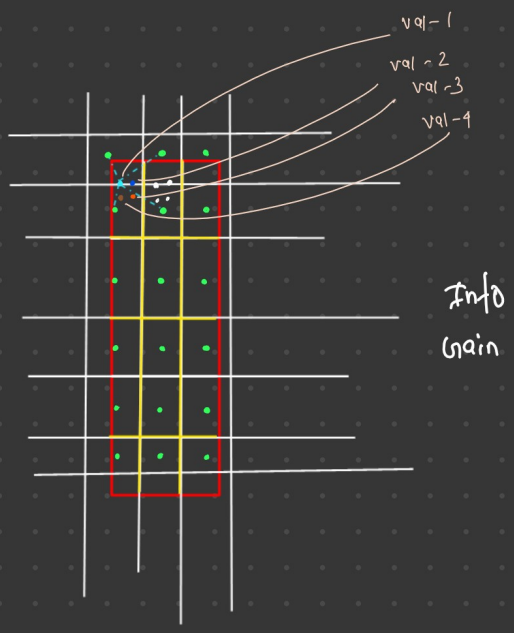
4. For 4 points we will get 4 values, apply max pool on 4 values

5. Do for all to get 3x3 matrix



(RoI Align)

Gain



Info Gain

max(val-1, val-2, val-3, val-4) → val(1,1)

(1,1)	(1,2)	(1,3)
(2,1)	—	—
—	—	(3,3)

3x3