Python Case study : Masked RCNN

▼ What is difference between Semantic Segmentation and Instance Segmentation?

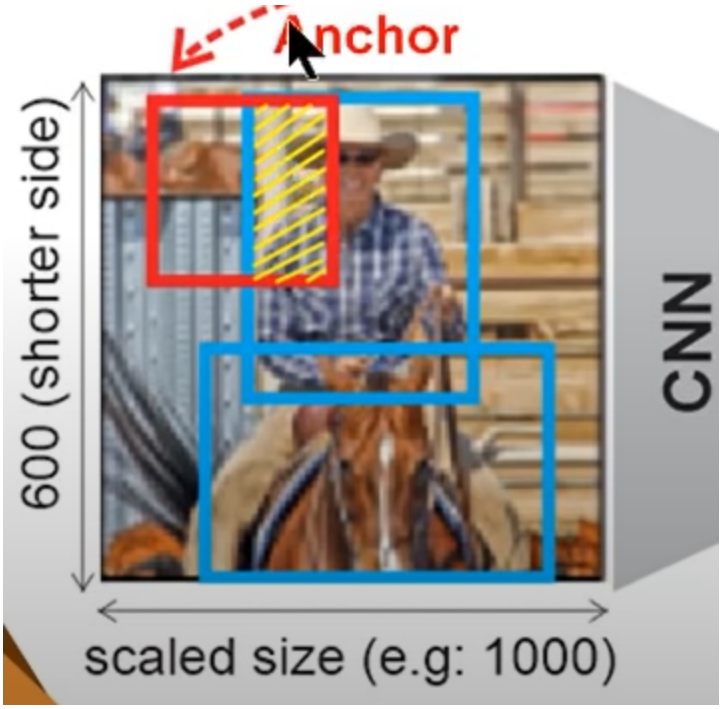
Semantic Segmentation

Instance Segmentation

Region Proposal Network

▼ What is Anchor Layer?

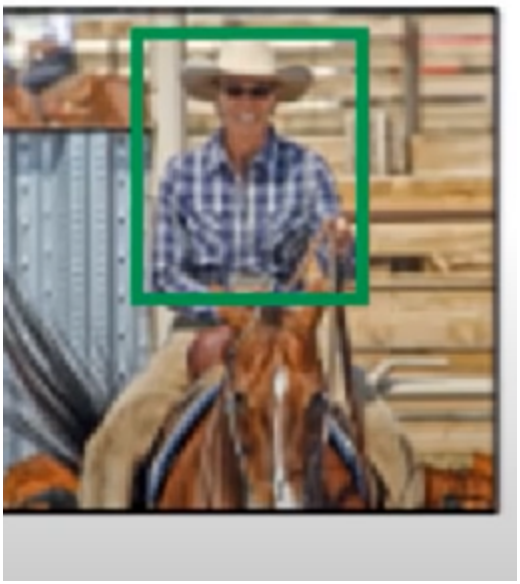
Anchor layers are sub matrices that are used to find the chance of us finding the image that we are looking for in that area .For example , lets say we are trying to find a person in our complete image



As we can see, the anchor(red rectanlge) is able to overlap over some part of the ACTUAL object that we are looking for . The bounded boxes are provided by us in an excel sheet or a csv file for each image stating this is where the actual image is present .

▼ What are GT boxes?

GT stands for ground truth boxes . These GT boxes contain information of where the actual image that we are searching for exists .For example,



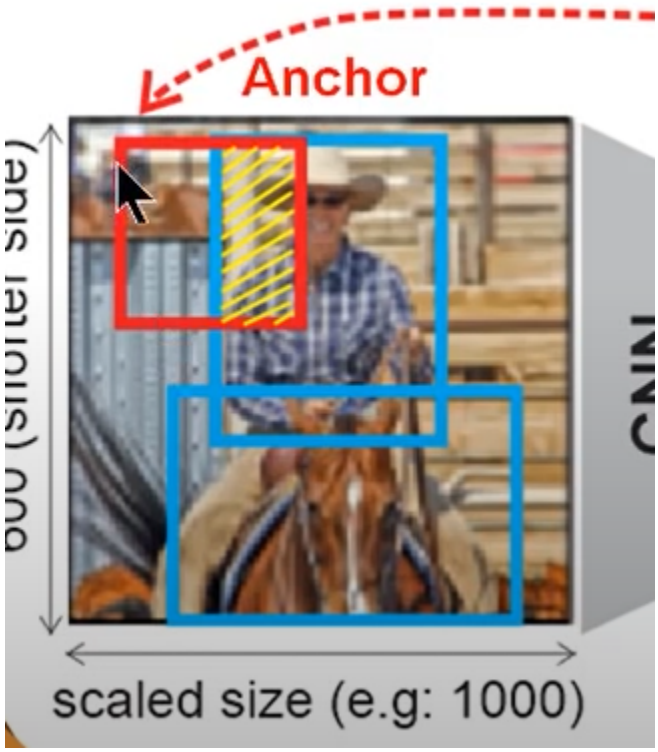
In the above image , The GT box is the green rectangle , or the person himself .

▼ What does IoU stand for and how do we calculate it?

IoU stands for Intersection over union . The formula to calculate it is

$$IoU = \frac{A \cap Gt}{A \cup Gt} \begin{cases} > 0.7 = \text{object} \\ < 0.3 = \text{not object} \end{cases}$$

Its the total are of intersection (between the anchor and the Ground truth box)divided by total are of union.



In this particular case there is not much overlap .Hence the anchor is not considered to have object inside .

▼ In masked RCNN we have two aspects .We need to not only identify whether there is an object in the image , but we also need to find **WHERE** the image is. How do we achieve this?

RPN loss function

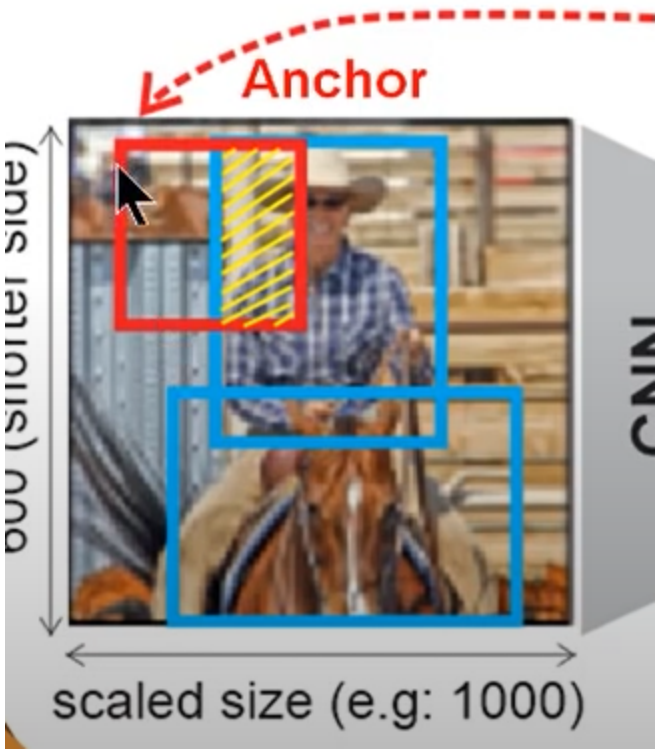
$$L(p_i, t_i) = \underbrace{\frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*)}_{\text{object/not object classifier}} + \underbrace{\lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)}_{\text{box regressor}}$$

So two aspects to the loss function(which should ideally be less than 2) .It should have object classfier(logistic regression model via conventional softmax) AND we are also throwing in the box regression model , which will give us where the object is present .

Is the object there

Where is the object present

▼ How do we calculate the box regression ?



From the above image, we can see extract , different parameters

λ = constant value

N_{reg} = number of total anchors (~2000)

p_i^* { 1 for pos anchor

0 for neg anchor

$$
\begin{aligned}
t_x &= (x - x_a)/w_a, \quad t_y = (y - y_a)/h_a, \\
t_w &= \log(w/w_a), \quad t_h = \log(h/h_a), \\
t_x^* &= (x^* - x_a)/w_a, \quad t_y^* = (y^* - y_a)/h_a, \\
t_w^* &= \log(w^*/w_a), \quad t_h^* = \log(h^*/h_a),
\end{aligned}
$$

Note that after creating anchors for all the anchor values , x* represents the anchors for negative anchors and x represents the anchor for positive anchors .We do this across ALL possible combinations for a specific class of object.

And then finally we calculate the following parameters. From above .

$$
+\lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)
$$

$$
L_{reg} = \text{smooth}_{L_1}(t_i - t_i^*) \quad \text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise,} \end{cases}
$$

if the overlap is less than 1 then we use 0.5*x*x . But if its significant then we check mod(x)-0.5 as mentioned in the smooth function

Priority Nu 1:

step 1 : Identify the car (python)

if not car : exit

if car :

step 2: Upload images from the android. (Write a service in python (Webservice) .)

step 3 : Identify the damage front side , back or headlights , scratches (show box)

step 4: Identify Severity Percentage

step 5: Give the result back to the android.