Faster RCNN

The basic process is as follows a variant of CNN (VGG / ResNet / InceptionNet) is used as a feature extractor. The extracted features are then used as inputs for the region proposal network as well as the classifier. The region proposal network is responsible for generating candidate anchor boxes which are then converted to bounding boxes. While the classification side of it is the standard run of the mill image classification network.

Class Config: This class contains the base information related to the network. The most important aspects here are the image size, anchor sizes and ratios. These are varied depending on the application.

Get Data: This function is the input parser. It takes as input the annotation.txt file and stores the filename , bounding box – x , y , w , h and the class of the image in a dictionary called all\_imgs.

RoiPooling Layer: This function contains definition for Spatial Pyramid Pooling . This is used because the convolutional layer that is employed for classification cannot take fractured inputs. ROI Pooling takes the corresponding feature map for the given region of interest, converts it to a fixed size input. Here spatial pyramid pooling is used. In spatial pyramid pooling different spatial bins of varying sizes are maintained. According to each filter in each bin it is pooled and it is flattened to before being used as input to the fully connected classification layers.

<https://medium.com/analytics-vidhya/review-spatial-pyramid-pooling-1406-4729-bfc142988dd2#:~:text=Spatial%20pyramid%20pooling%20is%20applied,proposal%20to%20generate%20final%20predictions>.

<https://arxiv.org/abs/1406.4729>

nn\_base : defines the base vgg-16 network used as a feature extractor

rpn\_layer : This function defines the region proposal network. The region proposal network takes as input the feature maps and has two outputs. A binary classification between object and background and predictions for anchor box coordinates.

Classifier\_layer: This function defines the image classification network. It also takes as input the feature maps .ROIPooling is applied before classification. It performs multiclass classification and also finetunes the bounding box predicitons.

Union,intersection,iou – Helper functions to calculate intersection over union.

calc\_rpn : We have selected 3 anchor box size and aspect ratios. Giving us a total of 9 anchor boxes that can be generated. Ground Truth Anchor or GTA – refers to the bounding box coordinates specified in the RSNA dataset. This function takes as input the specified bounding box coordinates and checks which anchor box has the highest iou with the specified bounding box. An iou > 0.7 indicates a positive anchor box, iou < 0.3 indicates a negative anchor box . 0.3 < iou < 0.7 indicates a neutral anchor box. We ensure for every bounding box there is at least one positive anchor box . Negative and Neutral anchor boxes are not used in training of the RPN

augment : This function augments the images.

get\_anchor\_gt : Function that calls calc\_rpn after preprocessing images (augment). Img\_length\_calc\_function is just a parameter that passes the name of the function to invoke in order to obtain output feature dimensions.

rpn\_loss\_regr , rpn\_loss\_class , class\_loss\_regr , class\_loss\_class : These functions define the loss function for the rpn bounding box regression , rpn classification for foreground or background (logistic loss) , the loss function the classifier bounding box regression and multiclass classification loss respectively.

Non\_max\_supression : If multiple bounding boxes are drawn around the same object . Non-max suppression is the process that is utilized filter and keep the best bounding boxes. Of all the bounding boxes the one with the highest confidence score is kept , boxes which have a high IoU with this box are removed.

apply\_regr\_np : This function generates bounding box coordinates for all the predicted anchor boxes for which non-max suppression is applied.

rpn\_to\_roi : This function converts the predicted anchor box coordinates to x1,y1,x2,y2 bounding box coordinates.

calc\_iou : converts x1,y1,x2,y2 to x,y,w,h bounding box coordinates for boxes which satisfy the overlap criterion with the ground truth box.