2019-08-06 2:45:21 PM

IOU net is proposed to take an image and a bounding box as input and predict the IOU between that box and the ground truth bounding box of the target object in that image

It seems to be mostly targeted as an add-on module or layer for existing detectors where it can be applied either in place of NMS or for bounding box regression

In case of NMS, output of the IOU net is used as the target objective for selecting the best bounding box instead of classification score

classification score of the final predicted box corresponding to each cluster is taken to be the maximum among all of the boxes in that cluster while the box location remains that given by the box with the maximum IOU net score

In the case of regression, it fits into the faster RCNN / FPN pipeline by simply replacing the ROI pooling layer with the proposed precise ROI pooling layer which is differentiable with respect to its input bounding box coordinates

Precise ROI pooling achieves this differentiability by using bilinear interpolation to convert the feature map into a continuous function

It also seems to be trainable by itself without using the RPN proposed the bounding boxes by simply applying some caution jittering to the ground truth box to generate a set of random boxes whose IOU with the ground truth remains greater than a threshold like 0.5 and then training on all of these boxes

gain in performance is very small or nonexistent except when high a localization accuracy metrics such as AP90 are used