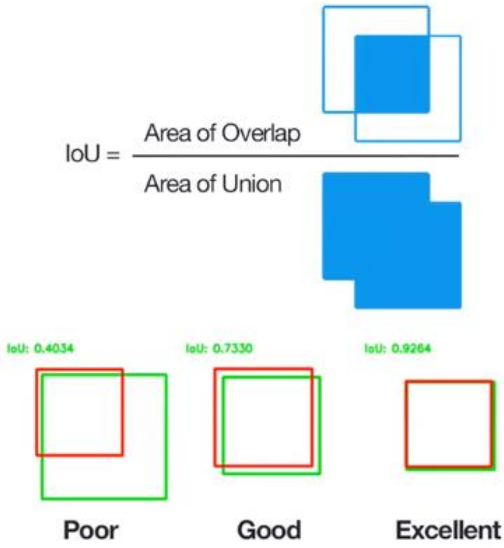
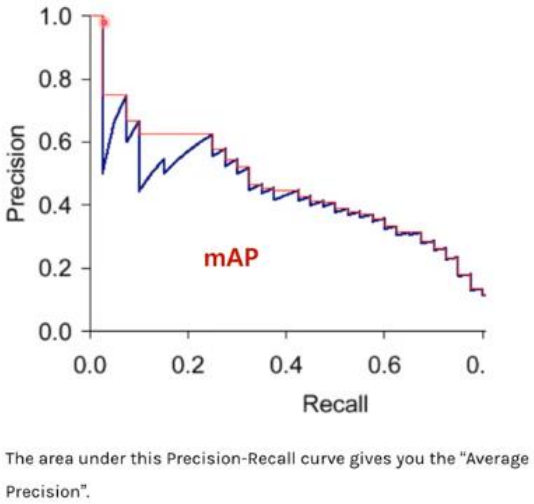
Yolov3: Incremental Improvement

- **mAP**: min Average Precision
- **IoU**: Area of Overlap / Area of Union (intersection of Union)

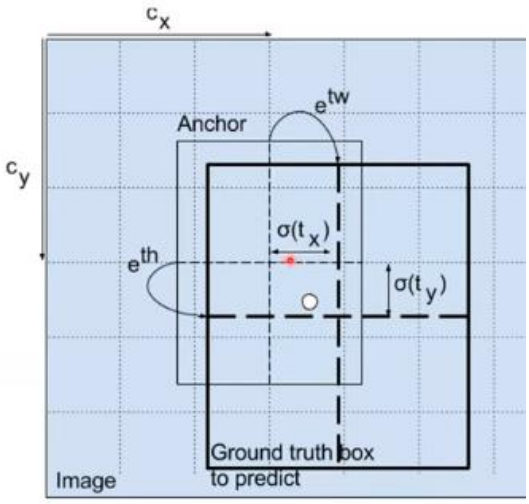




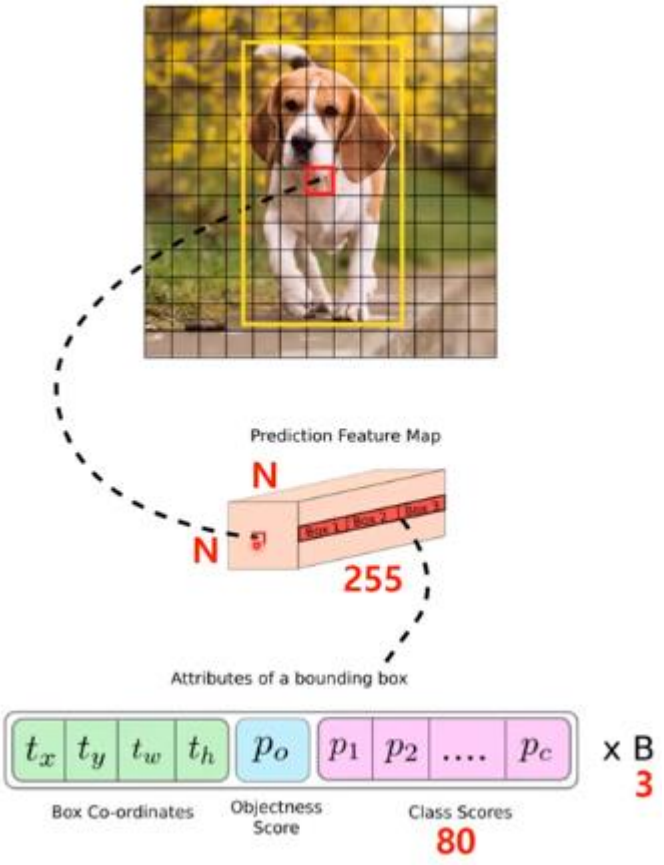
The area under this Precision-Recall curve gives you the "Average Precision".

- **Bounding Box Prediction** (기존 yolov2와 다를 바 없다)

$$b_x = \sigma(t_x) + c_x$$
$$b_y = \sigma(t_y) + c_y$$
$$b_w = p_w e^{t_w}$$
$$b_h = p_h e^{t_h}$$



- YOLOv3 only assigns one bounding box prior for each ground truth object (ground truth 하나당 예측한 box 하나) // 기존에는 gt하나에 iou>0.6이상 box matching ㅇㅇ
- YOLOv3 predicts boxes at 3 scales, YOLOv3 predicts 3 boxes at each scale



-> in total 9 boxes (so tensor is N x N x (3 x (4 + 1 + 80))

- Anchor Boxes >> use k-means clustering to determine bounding box priors

No. of Bounding Boxes

- YOLOv1 predicts 98 boxes (7x7 grid cells, 2 boxes per cell @448x448)
- YOLOv2 predicts 845 boxes (13x13 grid cells, 5 anchor boxes @416x416)
- YOLOv3 predicts 10,647 boxes (@416x416)
- YOLOv3 predicts more than 10x the number of boxes predicted by YOLOv2

-

- Darknet-53

Backbone	Top-1	Top-5	Bn Ops	BFLOP/s	FPS
Darknet-19 [15]	74.1	91.8	7.29	1246	**171**
ResNet-101[5]	77.1	93.7	19.7	1039	53
ResNet-152 [5]	**77.6**	**93.8**	29.4	1090	37
Darknet-53	77.2	**93.8**	18.7	**1457**	78

- Darknet-53 is better than ResNet-101 and 1.5x faster. Darknet-53 has similar performance to ResNet-152 and is 2x faster.
- Darknet-53 also achieves the highest measured floating point operations per second. This means the network structure better utilizes the GPU, making it more efficient to evaluate and thus faster.