

# IoUNet: Acquisition of Localization Confidence for Accurate Object Detection

笔记本: Detection

创建时间: 2019/6/3 19:55

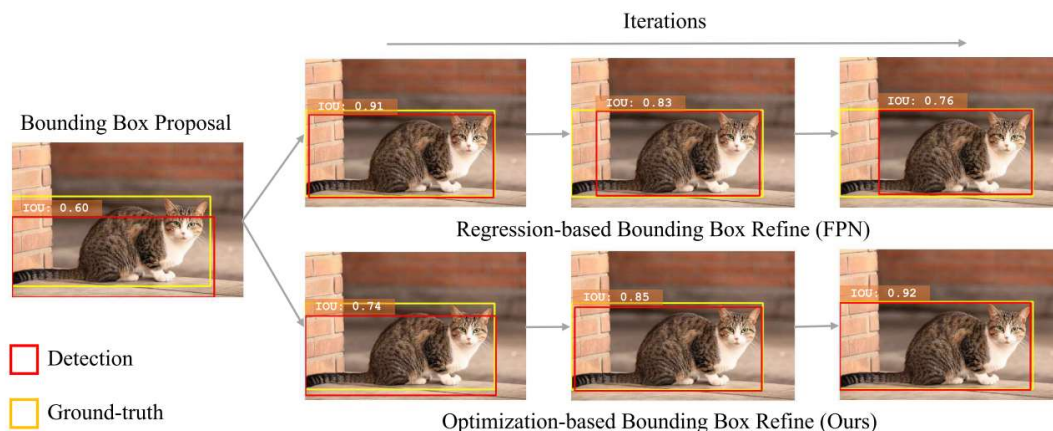
更新时间: 2019/11/7 16:50

作者: ming71

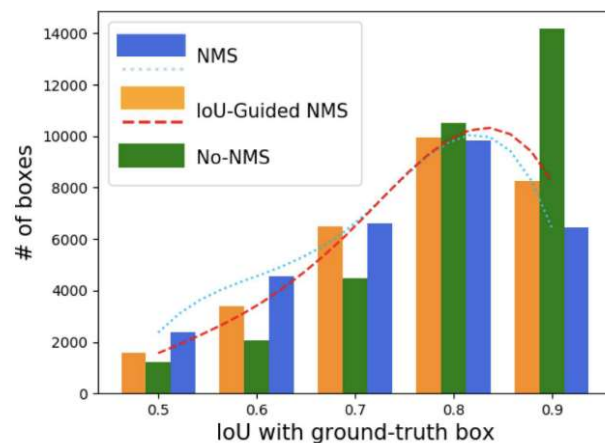
标签: Insight

## 1. Motivation

two-stage模型的proposal生成阶段, 分类和回归是分开进行的, 互相并行独立; 尽管共享部分卷积核共同训练, 但是评判机制上是各自独立的分类任务中, 通过分类置信度得分来给proposal打上类标签; 回归任务却直接得到期望的box坐标, 缺少对于box结果置信度的评价, 缺乏置信度的监测, 会导致边框的回归可解释性低。此外, 传统的bbox回归都是通过监督xywh损失来进行回归的, 理论上应该是迭代越多效果越好, 但是Cascade RCNN指出并非如此, 该过程并非单调的, bbox回归迭代过多反而会导致其可能下降。而采用iou作为回归置信度进行监督, 就不会受到这种限制。如下图:

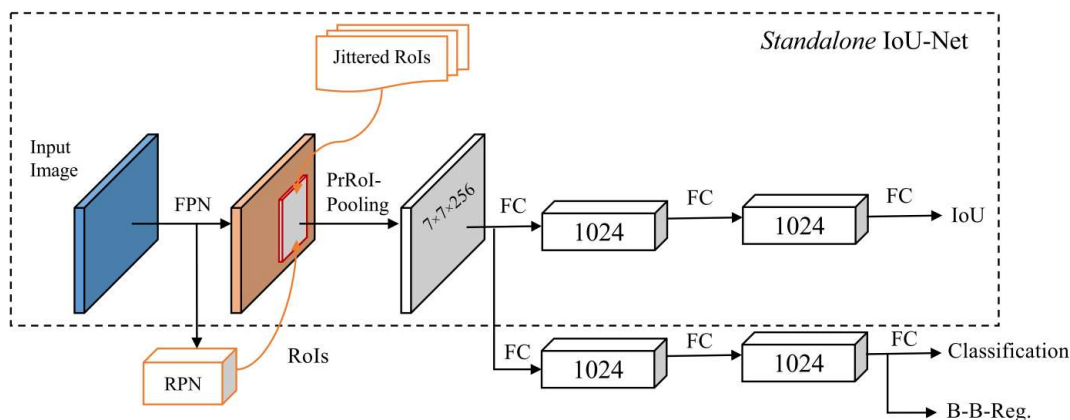


下图, 与未经NMS的原始box相比, 经过传统NMS,  $\text{IoU} > 0.9$ 的优质box约有1/3被剔除掉了, 这显然是不合理的



## 2. IoUNet

对于独立的IoU-Net网络，独立训练，训练时输入样本为gt经过增强后筛选出一批box，标签为这些box与对应gt的IoU值；测试时输入为proposal，输出为IoU值。



- **Learning to predict IoU**

IoU预测网络不是从RPN取proposal训练，而是用随机的参数gt进行增强，用于训练。而且IoU分支是独立训练的(实验中尝试了联合训练效果比单独训练好稍微好一点点)。

- **IoU-guided NMS**

NMS只用在检测最后阶段，RPN是直接根据score进行NMS的。（最终的目的）

- **Bounding box refinement as an optimization procedure**

bbox多次回归更新算法：

由下面分支预测出很多bbox后，同时上路分支计算其IoU score，记为PrevScore，并记下梯度grad。进行一步box参数更新，公式为：

$$b_j \leftarrow b_j + \lambda * scale(grad, b_j)$$
，然后用这个参数从上支路再计算一个IoU,为NewScore，如果满足PrevScore与NewScore差值小于早停阈值或

NewScore-PrevScore小于定位容忍界限，则停止参数更新，这个最新的box就是目标结果，取出来。然后对其他的待评估box继续如此操作，直到全部box都更新完毕。

该方法由于加入了iou监督，可以很好地解决多次回归的退化问题，使用多次bbox回归取得更好效果。

- **Precise RoI Pooling**

RoIAlign存在的问题是，每个bin都只取四个点，不具有可调整性，对于很大的目标，bin很大，更多的点也许更合适。

PrRoIPool采取的是插值积分形式，将roi的点看成是连续的，求积分除以面积，相当于精确的平均值池化。

- **Joint training**

IoU predictor训练的损失函数是smooth-L1 loss训练，训练样本是单独生成的，如上面的Learning to predict IoU描述。inference阶段，对最终产生的box进行IoU-guided NMS，选取100个box，再进一步采用后面的优化更新算法进行迭代求解最优bbox。

## 4. Experiment

- **IoU-guided NMS**

对高iou的AP指导效果明显，低iou的AP反而差，平均AP70左右和其他效果持平。

解决低iou效果差的方法是多保留proposal并下调iou阈值。

- **Optimization-based bounding box refinement**

效果同上。