

Anchor-free方法随笔

点击上方“[CVer](#)”，选择加“星标”或“置顶”

重磅干货，第一时间送达

作者: yyfyan

<https://zhuanlan.zhihu.com/p/72672611>

本文已由作者授权，未经允许，不得二次转载

Anchor-----Anchor-----Anchor or Free-----Free-----Free

周末公司训练平台崩了，只能复习基础、看看论文、写写东西了：

）最近也在做关于anchor free的方法，把一些东西记录下。

很多文章很早就看了，没留下笔记，这里也不重新写了。

估计没什么干货~~~~后面还会继续更新的。

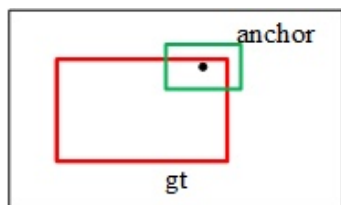
基于anchor free的目标检测方法最近一年比较流行，出现了很多"有趣"的工作。而原本基于anchor的目标检测方法，借助anchor更容易的回归不同尺度、不同长宽比的目标，是RPN中本首先采用的，之后是目标检测的标配。

这么简单有效的anchor，仍然存在以下缺点(quan shi chao de)：

- 一般需要针对不同数据集不同任务设计不同尺度(sc)与长宽比(ar)的anchor，如对于人脸检测，一般设置长宽比为1的anchor，而对于行人检测则设置为0.41等。anchor具体的设置可以视为hyper-parameters，RetinaNet中说明了该超参数可以提升~4%的AP。
- 泛化性能差，基于anchor的目标检测器对于不同的任务需要设置不同的anchor(同上)。
- 为了提高检测的recall，一般需要密集地平铺大量的anchor，典型800大小的输入，FPN大约有180K anchor boxes。这一方面导致匹配计算IOU时的计算量增加，另一方面导致了正负样本的极度不平衡。

为什么anchor-free?

- 下图中绿色框为anchor，红色为gt，基于anchor的目标检测方法通常认为anchor与gt之间的IOU大于0.7/0.5为正样本，下图由于IOU小于0.5/0.7而被认为是负样本或ignore。而anchor-free的方法一般会认为中心点落在gt内部即为正样本。因此，anchor-free一般会得到更多的正样本。(这里只是举一个极端的例子~~~)



- anchor-free方法存在很明显的歧义问题：使用点描述目标可能会出现一个点落到不同目标，而大多数anchor-free方法处理该问题使用FPN，FPN的不同层处理不同尺度的目标，极大的缓解了该歧义问题。

- anchor方法存在正负样本极度不平衡的问题，在anchor-free中也存在，甚至更明显。而RetinaNet中提出的Focal loss是处理该问题最简单有效的方法。RetinaNet是真的强~~。绝大多数anchor free方法分类分支均使用Focal loss。

anchor-free方法的分类

anchor-free大致可以分为两大类：

- 一类是以bottom-up方法为主：该类方法首先检测keypoints，然后对keypoints进行后处理以确定目标；这类方法有：CornerNet/ExtremeNet/CenterNet；
- 另一类是以FPN为主：该类方法类似于RetinaNet，只是改变了bbox的编码方式，同时也针对性的改变了FPN不同层的分配问题。FCOS/CenterNet/FoveaBox/FASA/RepPoints；

方法总结

这里就列4篇我自己比较喜欢的~~~

- CenterNet-----Objects as Points



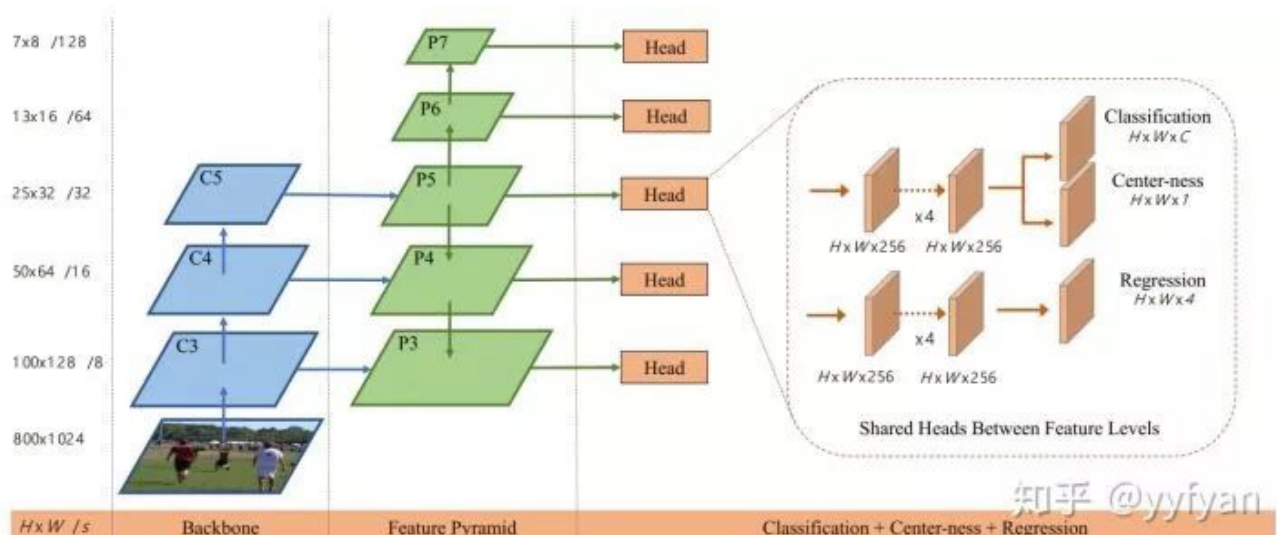
keypoint heatmap [C]

local offset [2]

object size [2]

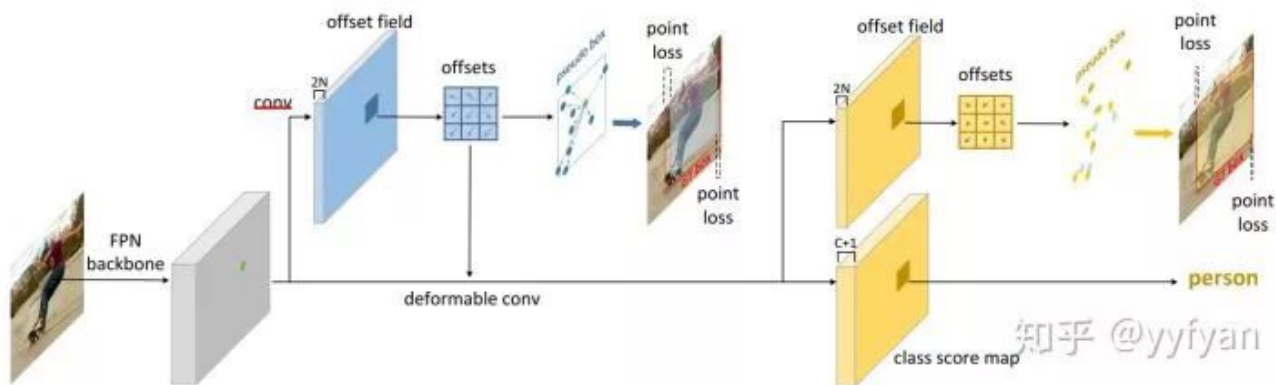
该文章首先以keypoints算法找到目标中心点，然后使用类似于Guassian的趋势对中心点周围的点计算loss。同时计算中心点的offset，以减少中心点的量化误差。为了能确定bbox，同样预测目标的size：w和h。其中，offset与size是所有类别共享的。整体类似于CornerNet。

- FCOS-----Fully Convolutional One-Stage Object Detection



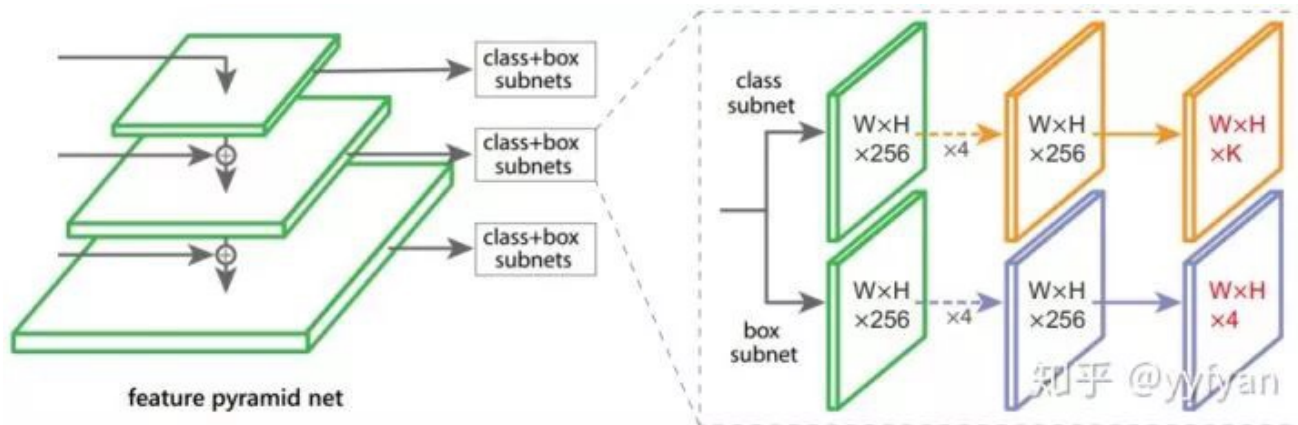
整体上类似于RetinaNet, 1) 把特征图上的点映射到原图相比于基于anchor方法, 会增加更多正样本, 每一个点预测一个目标, 预测C类+4个offset(预测点到四个边上的距离作为回归目标)。2) 由于一个点可能对应不同的gt bbox, 利用FPN可以有效解决; 3) 由于远离目标中心点的点预测的结果含有大量误检测, 因此, 利用单独一个分支学习点到中心的加权系数center-ness, 类似于一个attention map。

- RepPoints: Points Set Representation for Object Detection



基于anchor的方法, 让anchor多次回归到gt, 进而检测目标。但如果gt与anchor之间的距离很远的时候, 性能降低。RepPoints使用9个点去预测目标, 9个点经过一次refine。refine之间其实是deformable conv。每一次refine时使用的是左上角与右下角之间的l1 loss。网络整体还是基于RetinaNet。

- FoveaBox: Beyond Anchor-based Object Detector



和FCOS比较类似, 区别在于分配到FPN不同层级的方式变了, 目标bbox的编码方式变了, 正负样本的确定方式也有所区别。

思考

改改目标的编码方式, 然后anchor free, 不用anchor做参考, 也能很好的检测出目标。难道是数据集变大的原因(这些方法基本都跑的是COCO)。anchor free目前仍没有大幅度优于anchor方法, 谁优谁劣尚不明朗~~不过技术总是在进步, 如果anchor free真的有效, 大幅超越anchor方法也是必然的。或许后面很多任务可以借助anchor free去做, 比如基于bottom-up的实例分割, 全景分割等。没了:)

CVer-目标检测交流群

扫码添加CVer助手，可申请加入CVer-目标检测交流群。一定要备注：**研究方向+地点+学校/公司+昵称**（如目标检测+上海+上交+卡卡）



▲长按加群



▲长按关注我们

麻烦给我一个在看！