

# FCOS: Fully Convolutional One-Stage Object Detection

笔记本: Detection

创建时间: 2019/4/29 1:51

更新时间: 2019/11/3 9:57

作者: ming71

## 1. Motivation

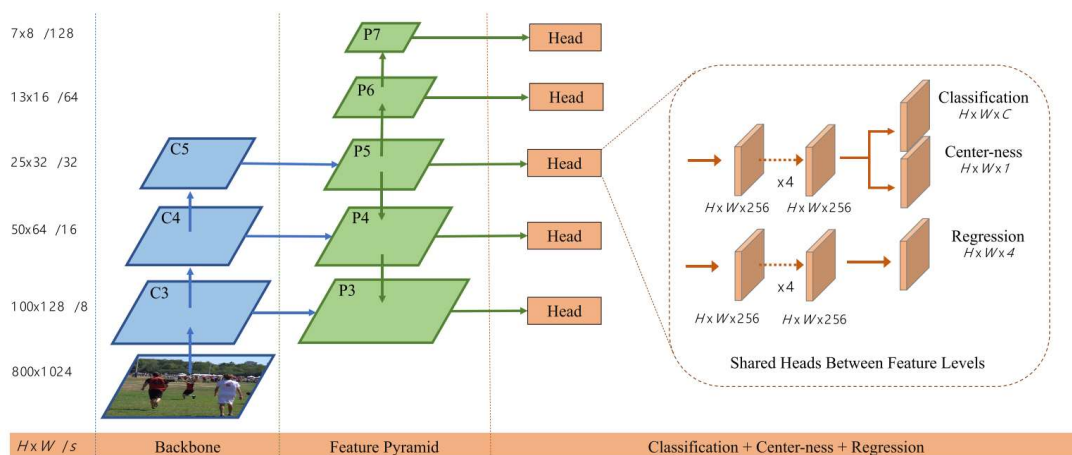
anchor-based缺点:

- 检测性能和anchor的尺寸数目形状等超参数有关, 需要精确设置
- 实际物体具有较大的尺度变化性, 手工设置的固定anchor仍性能有限; 而且有失一般性, 对不同的任务泛化能力差
- 为了保正anchor的召回率避免漏掉, 往往会用大量密集的anchor覆盖物体, 其中负样本过多, 类别不平衡
- 过多的anchor导致计算量大

Insight:

- 能否像FCN做语义分割一样在单个像素上进行预测?
- 采用FPN融合多尺度信息
- 像素级预测肯定会产生过多的低质量box, 采用到gt中心的距离来筛选掉低质量box降低计算

## 2. Design



### 2.1 Fully Convolutional One-Stage Object Detector

对于某张需要预测的特征图, 将其映射到原图可以得到每个cell的中心位置, anchor-based方法就是在这些位置放上anchor的中心进行anchor的回归, 而本方法中对其直接回归bbox.

将这个位置(x,y)划分正负样本：如果它落入某个gt内，那么其类别标签c为该gt的类 (coco: 1-80) ， 否则是负样本， label c=0(背景).同时还为每个位置预测一个4维向量 $t^* = (l^*, t^*, r^*, b^*)$ ， 表征该位置到左上右下四条边的距离。如果一个位置落入了多个gt中，定义其为模糊样本，简单地选取面积最小的gt为回归的目标（后面会用多级预测解决这个问题）。

#### ◦ Network Outputs

分别通过不同的分支进行分类和回归。相较于其他detector训练多分类器，这里分类是训练C个二分类器；对于预测的四条边距离的向量，由于真实距离一定是正的（点在gt内）将预测值拉伸到  $(0, +\infty)$  。

#### ◦ Loss Function

$$L(\{p_{x,y}\}, \{t_{x,y}\}) = \frac{1}{N_{\text{pos}}} \sum_{x,y} L_{\text{cls}}(p_{x,y}, c_{x,y}^*) + \frac{\lambda}{N_{\text{pos}}} \sum_{x,y} \mathbb{1}_{\{c_{x,y}^* > 0\}} L_{\text{reg}}(t_{x,y}, t_{x,y}^*)$$

上式中， $N_{\text{pos}}$ 是正样本个数，求和是对所有特征图而言的， $L_{\text{cls}}$ 是分类误差focal loss，lambda为1用于权衡分类回归的比重，花体1是示性函数，只计算前景的回归损失，回归损失采用IoU loss（UnitBox中的）。

后面对上式还要加上center-ness损失。

#### ◦ Inference

前向传播得到分类得分和回归结果，取分类分数大于0.5的认为是正样本，并计算对应回归结果的bbox。

## 2.2 Multi-level Prediction with FPN for FCOS

- 需要解决的问题：（1）anchor-based检测器中，降采样步长过大（如32）容易导致anchor覆盖的最高召回率低，他们可以通过降低筛选的IoU进行这种弥补，但是anchor-free面临步长大而忽视部分目标时，需要采取其他办法解决这个问题（2）上面提到的模糊样本，一个location落入多个gt，那么应该向哪个回归（这也是FCN-based检测器性能劣化的重要原因）
- 解决方法：多尺度预测。对每个尺度特征图（P3-P7）设置location到四条边的距离 $\max(l^*, t^*, r^*, b^*)$ 阈值，如果 $\max(l^*, t^*, r^*, b^*)$ 落入合理的区间才会视为正样本，过大或过小视作负样本不进行回归；实际上阈值是连续设置的，和FPN不同的level进行对应，设置为0-64-128-256-512- $\infty$ 。将物体分配到不同的尺寸进行预测。
- 输出激活映射：之前2.1提到，对于输出采用exp函数将输入映射到 $0 \sim +\infty$ 可以拉伸特征空间范围，但是对于不同的level预测结果范围是不同的，采用相同映射不合理，这里采用可学习的 $\exp(\text{si}^*x)$ ，为不同level学习合理的si基数。

## 2.3 Center-ness for FCOS

- 问题：对每个location生成的box中有很多中心远离gt的低质量box，导致回归效率低下。（理解：例如一个gt内落入了两个location，其中距离中心位置近的比较容易回归，属于高质量的box，距离中心很远甚至到边缘的box较难回归，属于低质量box）
- 解决办法：预测中心

## 3. 结构流程图

