

Region Proposal by Guided Anchoring

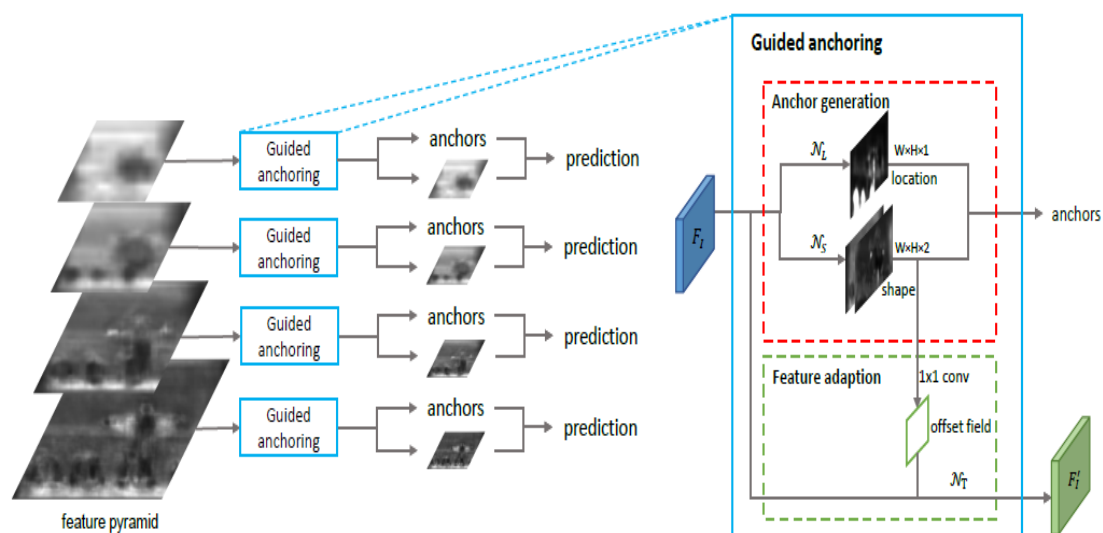
Jiaqi Wang^{1*} Kai Chen^{1*} Shuo Yang² Chen Change Loy³ Dahua Lin¹

¹CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong

²Amazon Rekognition ³Nanyang Technological University

{ck015,wj017,dhlin}@ie.cuhk.edu.hk shuoy@amazon.com ccloy@ntu.edu.sg

<https://zhuanlan.zhihu.com/p/55854246>



根据图像特征生成 anchor

Anchor 概率分布:

$$p(x, y, w, h|I) = p(x, y|I)p(w, h|x, y, I).$$

Anchor Location Prediction

特征图 FI 经过 1×1 卷积和 sigmoid 函数转换为概率 $p(i, j|FI)$ ，设定阈值过滤 90%

Anchor Shape Prediction

anchor 的 w 和 h 是不确定的，是一个需要预测的变量。我们将这个 anchor 和某个 gt 的 IoU 表示为

$$vIoU(a_{wh}, gt) = \max_{w>0, h>0} IoU_{normal}(a_{wh}, gt)$$

特征图 FI 经过 1×1 卷积生成两通道的图，一个是 dw 和 dh，另一个是通过公式：

$$w = \sigma \cdot s \cdot e^{dw}, \quad h = \sigma \cdot s \cdot e^{dh}.$$

(where s is the stride and # is an empirical scale factor (# = 8 in our experiments).) 得出 w,h。

Feature Adaptation

在 RPN 中，anchor 形状相同，但此处 anchor 形状大小不同。对原本的特征图来说，它并不知道形状预测分支预测的 anchor 形状，但是接下来的分类和回归却是基于预测出的 anchor 来做的。

思路：把 anchor 的形状信息直接融入到特征图中，这样新得到的特征图就可以去适应每个位置 anchor 的形状。

利用一个 3×3 的 deformable convolution 来修正原始的特征图，而 deformable convolution 的 offset 是通过 anchor 的 w 和 h 经过一个 1×1 conv 得到的。

可变形卷积：

是基于通过学习一个额外的偏移 (offset)，使卷积核对输入 feature map 的采样的产生偏移，集中于感兴趣的目标区域。

Table 4: The effects of each module in our design. L., S., and F.A. denote location, shape, and feature adaptation, respectively.

L.	S.	F.A.	AR ₁₀₀	AR ₃₀₀	AR ₁₀₀₀	AR _S	AR _M	AR _L
			47.5	54.7	59.4	31.7	55.1	64.6
✓			48.0	54.8	59.5	32.3	55.6	64.8
	✓		53.8	59.9	63.6	36.4	62.9	71.7
✓	✓		54.0	60.1	63.8	36.7	63.1	71.5
✓	✓	✓	59.2	65.2	68.5	40.9	67.8	79.0