

Why anchor?

点击上方“[CVer](#)”，选择加“星标”或“置顶”
重磅干货，第一时间送达

本文转载自：SIGAI

作者知乎：

<https://www.zhihu.com/people/milestone-94/activities>

SIGAI特约作者

milestone

研究方向：计算机视觉

1、为什么需要anchor机制

anchor机制由Faster-RCNN提出，而后成为主流目标检测器的标配，例如SSD、Yolov2、RetinaNet等等。

Faster-RCNN中使用anchor机制的motivation如下：

In contrast to prevalent methods [8], [9], [1], [2] that use pyramids of images (Figure 1, a) or pyramids of filters (Figure 1, b), we introduce novel “anchor” boxes that serve as references at multiple scales and aspect ratios. Our scheme can be thought of as a pyramid of regression references (Figure 1, c), which avoids enumerating images or filters of multiple scales or aspect ratios.

即anchor机制要解决的问题是scale和aspect ratio变化范围大，之前的解决方法是pyramids of images或者pyramids of filters。pyramids of images耗时，而pyramids of filters是传统CV的做法，当时CNN-based的CV还没有一种对应pyramids of filters的方案（对应的方案得等到2016年底的FPN，即pyramids of features）。

所以作者提出了一种新的解决方案——anchor机制。

时间	模型
2015-06	Faster RCNN
2015-06	Yolov1
2015-09	DenseBox
2015-12	SSD
2016-04	MTCNN
2016-12	Yolov2
2019-03	FSAF
2019-04	FCOS
2019-04	FoveaBox

图1.1 目标检测简史



图1.2 两个gt box落到同一个cell

虽然作者解释anchor机制是为了解决scale和aspect ratio变化范围大的问题，但anchor机制还顺便解决了另外一个重要的问题——gt (ground truth) box与gt box之间overlap过大会导致多个gt box映射到一个cell（即检测head上的一个特征点），从而导致gt box丢失，如图1.2所示，人体框和球拍框中心点相同，会落到同一个cell里。有了anchor机制之后，一个cell里会有多个anchor，不同anchor负责不同scale和aspect ratio的gt box，即使有多个框会映射到同一个cell，也不会导致gt box丢失。具体地，在训练阶段，gt box丢失会导致模型不能见到全量gt；在测试阶段，gt box丢失，会导致recall降低。

简而言之，anchor机制有两个作用：分而治之（将scale和aspect ratio空间划分为几个子空间，降低问题难度，降低模型学习难度）和解决gt box与gt box之间overlap过大导致gt box丢

失问题。同理，pyramids of images和pyramids of features也有上述两个作用。

2、八仙过海，各显神通

上面已经说了，anchor机制是pyramids of images和pyramids of features的替代方案。换一句话说，pyramids of images、pyramids of features和anchor机制是解决同一个问题的三种思路。

比如DenseBox，MTCNN使用的是pyramids of images机制；而FCOS、FoveaBox使用的是pyramids of features机制；Faster-RCNN、Yolov2使用的是anchor机制；SSD、RetinaNet、Yolov3糅合了anchor机制和pyramids of features机制。

3、anchor free一定更好吗？

最近几个月anchor free的相关文章喷涌而出，大有革掉anchor based检测器命的势头，那么问题来了，anchor free就一定比anchor based的方法更好吗？

anchor free和anchor based方法各有优劣。例如，anchor free的前提是基于图像金字塔或者特征金字塔这个前提，但是无论哪种金字塔都会增加计算，降低检测速度，而anchor机制可以减少金字塔层数，进而提高检测速度。

就如同pyramids of images和pyramids of features两种方法各有优劣一样。基于pyramids of features的FPN大火的同时，基于pyramids of images的MTCNN方法依然有自己的一片天地。无论是pyramids of images、pyramids of features还是anchor机制，没有孰优孰劣，它们有各自的优缺点，应该按需使用，甚至糅合到一起使用，取每种方法的长处，规避每种方法的短处。

例如ACN（Anchor Cascade for Efficient Face Detection）结合了pyramids of images和anchor机制，从而在速度和效果上取得了一个不错的折中。

技术是螺旋前进的，一个技术领域往往有多个演进路线，而每个时代，由于硬件或者理论基础等等的限制，总会有一个当前最优的演进路线，但是没有一个技术演进路线在任何时代都是最优的。就跟机器学习里的no free lunch理论一样，没有一个模型能在所有任务上都表现最优，同理没有一个技术路线能在任何时代都代表最先进方向。

扫码添加CVer助手，可申请加入**CVer-目标检测交流群**、**图像分割**、**目标跟踪**、**人脸检测&识别**、**OCR**、**超分辨率**、**SLAM**、**医疗影像**、**Re-ID**、**GAN**、**NAS**、**深度估计**、**自动驾驶**和**剪枝&压缩**等群。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！