

3.10 候选区域（选修）（Region proposals (Optional)）

如果你们阅读一下对象检测的文献，可能会看到一组概念，所谓的候选区域，这在计算机视觉领域是非常有影响力的概念。我把这个视频定为可选视频是因为我用到候选区域这一系列算法的频率没有那么高，但当然了，这些工作是很有影响力的，你们在工作中也可能会碰到，我们来看看。



你们还记得滑动窗法吧，你使用训练过的分类器，在这些窗口中全部运行一遍，然后运行一个检测器，看看里面是否有车辆，行人和摩托车。现在你也可以运行一下卷积算法，这个算法的其中一个缺点是，它在显然没有任何对象的区域浪费时间，对吧。

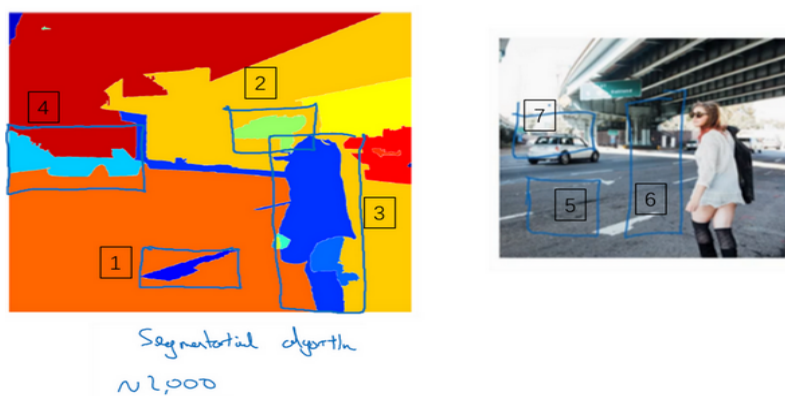


所以这里这个矩形区域（编号 1）基本是空的，显然没有什么需要分类的东西。也许算法会在这个矩形上（编号 2）运行，而你知道上面没有什么有趣的东西。

所以 **Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik**，在本幻灯片底部引用到的论文中提出一种叫做 **R-CNN** 的算法，意思是**带区域的卷积网络，或者说带区域的 CNN**。这个算法尝试选出一些区域，在这些区域上运行卷积网络分类器是有意义的，所以这里不再针对每个滑动窗运行检测算法，而是只选择一些窗口，在少数窗口上运行卷积网络分类器。

选出候选区域的方法是运行图像分割算法，分割的结果是下边的图像，为了找出可能存在对象的区域。比如说，分割算法在这里得到一个色块，所以你可能会选择这样的边界框（编

号 1)，然后在这个色块上运行分类器，就像这个绿色的东西（编号 2），在这里找到一个色块，接下来我们还会在那个矩形上（编号 2）运行一次分类器，看看有没有东西。在这种情况下，如果在蓝色色块上（编号 3）运行分类器，希望你能检测出一个行人，如果你在青色色块(编号 4)上运行算法，也许你可以发现一辆车，我也不确定。



所以这个细节就是所谓的分割算法，你先找出可能 2000 多个色块，然后在这 2000 个色块上放置边界框，然后在这 2000 个色块上运行分类器，这样需要处理的位置可能要少的多，可以减少卷积网络分类器运行时间，比在图像所有位置运行一遍分类器要快。特别是这种情况，现在不仅是在方形区域（编号 5）中运行卷积网络，我们还会在高高瘦瘦（编号 6）的区域运行，尝试检测出行人，然后我们在很宽很胖的区域（编号 7）运行，尝试检测出车辆，同时也在各种尺度运行分类器。

这就是 **R-CNN** 或者**区域 CNN** 的特色概念，现在看来 **R-CNN** 算法还是很慢的。所以有一系列的研究工作去改进这个算法，所以基本的 **R-CNN** 算法是使用某种算法求出候选区域，然后对每个候选区域运行一下分类器，每个区域会输出一个标签，有没有车子？有没有行人？有没有摩托车？并输出一个边界框，这样你就能在确实存在对象的区域得到一个精确的边界框。

澄清一下，**R-CNN** 算法不会直接信任输入的边界框，它也会输出一个边界框 b_x ， b_y ， b_h 和 b_w ，这样得到的边界框比较精确，比单纯使用图像分割算法给出的色块边界要好，所以它可以得到相当精确的边界框。

现在 **R-CNN** 算法的一个缺点是太慢了，所以这些年来有一些对 **R-CNN** 算法的改进工作，**Ross Girshik** 提出了 **Fast R-CNN** 算法，它基本上是 **R-CNN** 算法，不过用卷积实现了滑动窗法。最初的算法是逐一对区域分类的，所以 **Fast R-CNN** 用的是滑动窗法的一个卷积实现，这和你在本周第四个视频（3.4 卷积的滑动窗口实现）中看到的大致相似，这显著提升了 **R-CNN** 的速度。

Faster algorithms

→ R-CNN: Propose regions. Classify proposed regions one at a time. Output label + bounding box. ←

Fast R-CNN: Propose regions. Use convolution implementation of sliding windows to classify all the proposed regions. ←

Faster R-CNN: Use convolutional network to propose regions.

[Girshik et. al, 2013. Rich feature hierarchies for accurate object detection and semantic segmentation]

[Girshik, 2015. Fast R-CNN]

[Ren et. al. 2016. Faster R-CNN: Towards real-time object detection with region proposal networks] Andrew Ng

事实证明，**Fast R-CNN** 算法的其中一个问题是得到候选区域的聚类步骤仍然非常缓慢，所以另一个研究组，任少卿（**Shaoqing Ren**）、何凯明（**Kaiming He**）、**Ross Girshick** 和孙剑（**Jiangxi Sun**）提出了更快的 **R-CNN** 算法(**Faster R-CNN**)，使用的是卷积神经网络，而不是更传统的分割算法来获得候选区域色块，结果比 **Fast R-CNN** 算法快得多。不过我认为大多数 **Faster R-CNN** 的算法实现还是比 **YOLO** 算法慢很多。

候选区域的概念在计算机视觉领域的影响力相当大，所以我希望你们能了解一下这些算法，因为你可以看到还有人在用这些概念。对我个人来说，这是我的个人看法而不是整个计算机视觉研究界的看法，我觉得候选区域是一个有趣的想法，但这个方法需要两步，首先得到候选区域，然后再分类，相比之下，能够一步做完，类似于 **YOLO** 这个算法，在我看来，是长远而言更有希望的方向。但这是我的个人看法，而不是整个计算机视觉研究界的看法，所以你们最好批判接受。但我想这个 **R-CNN** 概念，你可能会想到，或者碰到其他人在用，所以这也是值得了解的，这样你可以更好地理解别人的算法。