# 阅读纲要

## 1 自己的总结、评价以及应用

该论文主要讲了Faster R-CNN在Fast R-CNN的基础上做的两项改进，使得算法在accurancy和running time两方面得到了极大地改善。

## 2 文章的主要问题（abstract、疑问句中）

Fast R-CNN存在的两个问题:accurancy和running time.

Faster R-CNN在两个方面做出了改进：

①使用RPN（Region Proposals Network）替代SS（Selective Search），其作用为the RPN module tells the Fast R-CNN module where to look.

本文的核心就在于研究object/region proposal methods的问题

②we propose a training scheme that alternates between fine-tuning for the region proposal task and then fine-tuning for object detection.使得region proposal和object detection可以协调进行。

## 3 结论（abstract以及conclusion中）

两个创新点：  
①We have presented RPNs for efficient and accurate region proposal generation.

②By sharing convolutional features with the down-stream detection network, the region proposal step is nearly cost-free.

## 4 思路脉络（小标题中的关键句）

1 INTRODUCTION

该部分主要介绍了Fast R-CNN算法存在的突出问题，并提出了RPN：

We introduce novel Region Proposal Networks (RPNs) that share convolutional layers with state-of-the-art object detection networks.

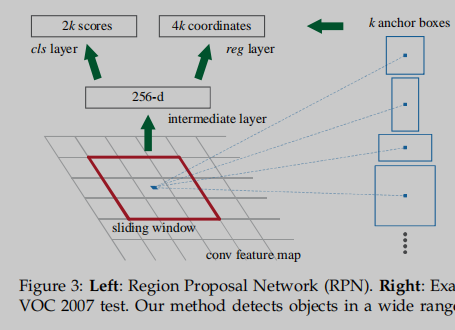
主要优点是有效降低了computation cost.

还提到了RPN的实现原理:anchor box（不理解）

2 RELATED WORK

3 FASTER R-CNN

3.1 Region Proposal Networks



3.1.1 Anchors

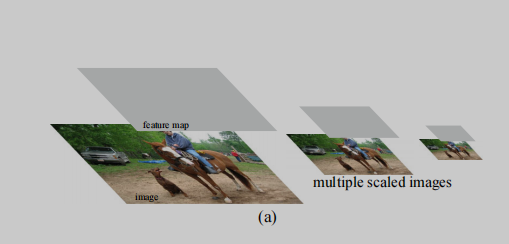
Translation-Invariant Anchors

关于“平移不变”特性的解释：. If one translates an object in an image,the proposal should translate and the same function should be able to predict the proposal in either location.

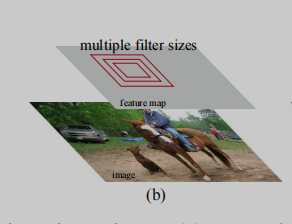
Multi-Scale Anchors as Regression References

大体解释：

通常有两种方式there have been two popular ways for multi-scale predictions处理多尺寸预测问题：  
① The first way （对图片进行切割、变形）is based on image/feature pyramids, e.g., in DPM [8] and CNN based methods [9], [1], [2]. The images are resized at multiple scales, and feature maps (HOG [8] or deep convolutional features [9], [1], [2]) are computed for each scale (Figure 1(a)).



② The second way （使用不同的filters）is to use sliding windows of multiple scales (and/or aspect ratios) on the feature maps.



而anchor-based method：

is built on a pyramid of anchors, which is more cost-efficient. Our method classifies and regresses bounding boxes with reference to anchor boxes of multiple scales and aspect ratios. It only relies on images and feature

maps of a single scale, and uses filters (sliding windows on the feature map) of a single size（我们的anchors尺寸不变）.

3.1.2 Loss Function

3.1.3 Training RPNs

3.2 Sharing Features for RPN and Fast R-CNN

有三种机制解决这个问题：

①Alternating training.（交替训练）这也是这篇论文中使用的方法

In this solution, we first train RPN, and use the proposals to train Fast R-CNN.

The network tuned by Fast R-CNN is then used to initialize RPN, and this process is iterated.

②近似联合训练Approximate joint training.（没看懂）

③非近似联合训练Non-approximate joint training.（没看懂）

1. Step Alternating Training

通过四个步骤，最终使得RPN和R-CNN共享一个同样的卷基层，并且形成了一个统一网络。

share the same convolutional layers and form a unified network.

3.3 Implementation Details

4 EXPERIMENTS

5 CONCLUSION

## 5 文献中难理解的点

主要有两点：

①RPN的工作原理，其核心在于anchors机制

②如何实现RPN与Fast R-CNN协作，主要有三种机制，本文使用的是交替训练这一方法，分为四个步骤。