Region Proposal Network and Faster R-CNN

Presenter: Minoo Jafarlou
Spring 2023
MLBBQ

Outline

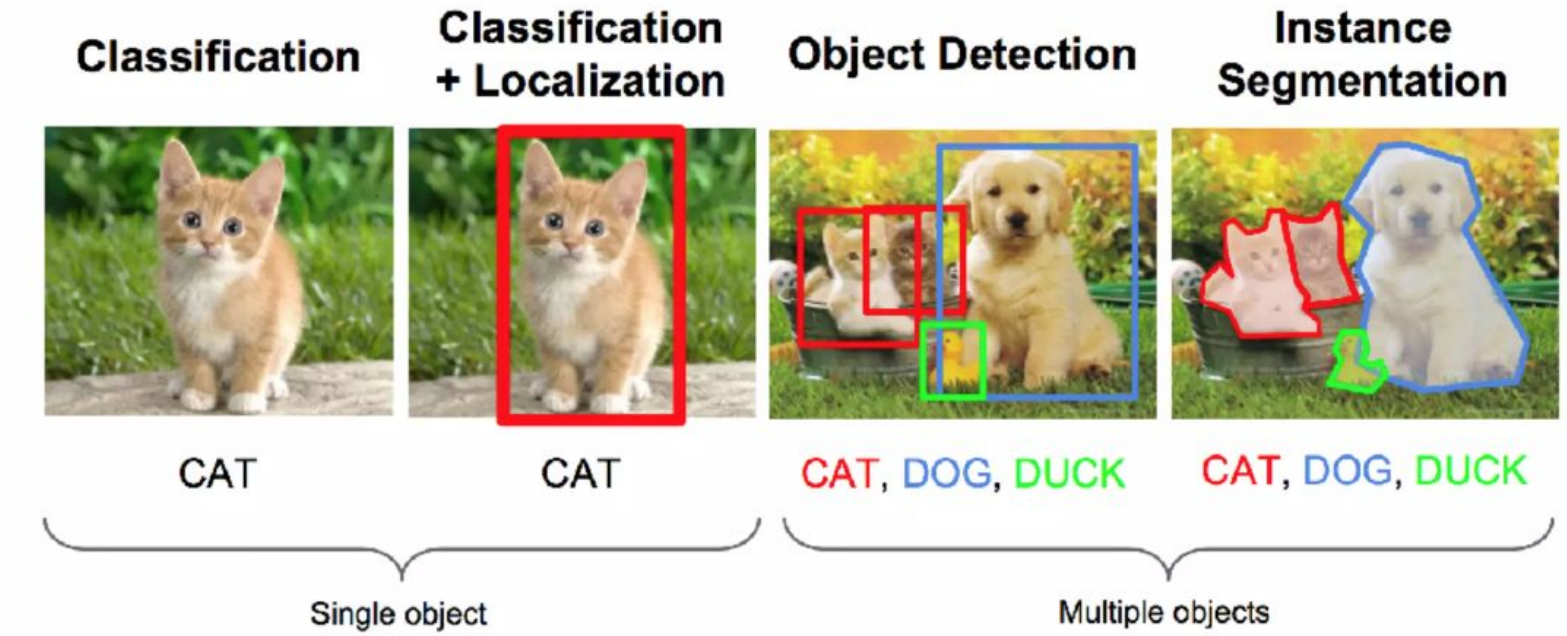
Fundamentals

- Computer vision tasks
- Localization + Regression
- Intersection Over Union

R-CNNs family’s

- R-CNN
- Fast R-CNN
- Faster R-CNN

Introduction
Computer Vision Tasks



http://cs231n.stanford.edu/slides/2016/winter1516_lecture8.pdf

Classification + Localization: Task

Classification: C classes
Input: Image
Output: Class label
Evaluation metric: Accuracy

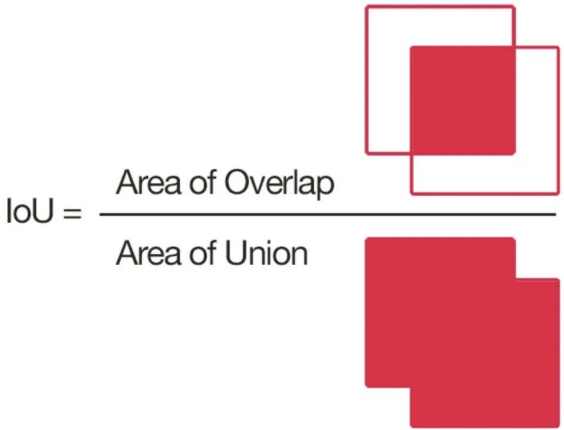
→ CAT

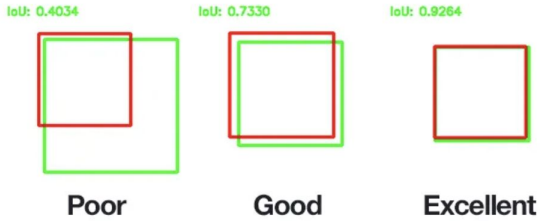
Localization:
Input: Image
Output: Box in the image (x, y, w, h)
Evaluation metric: Intersection over Union

→ (x, y, w, h)

Classification + Localization: Do both

Accuracy of the predicted bounding boxes with **Intersection Over Union (IoU)**





Outline

Introduction

- Computer vision tasks
- Classification + Localization

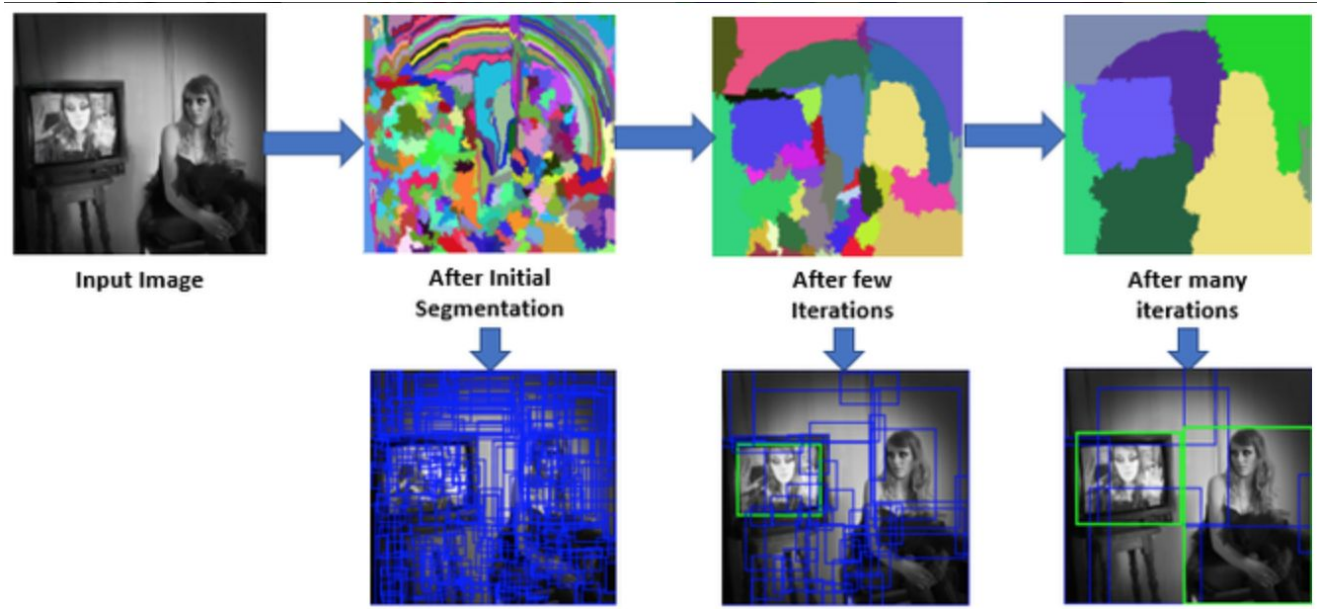
R-CNNs

- R-CNN
- Fast R-CNN
- Faster R-CNN

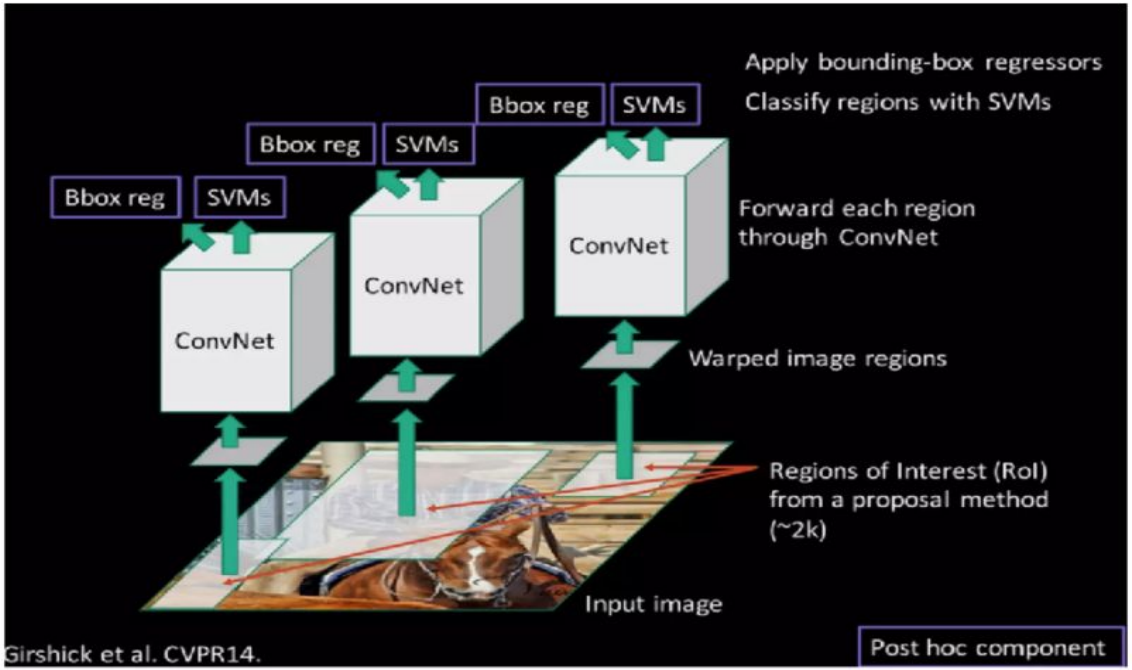
Conclusion

R-CNN

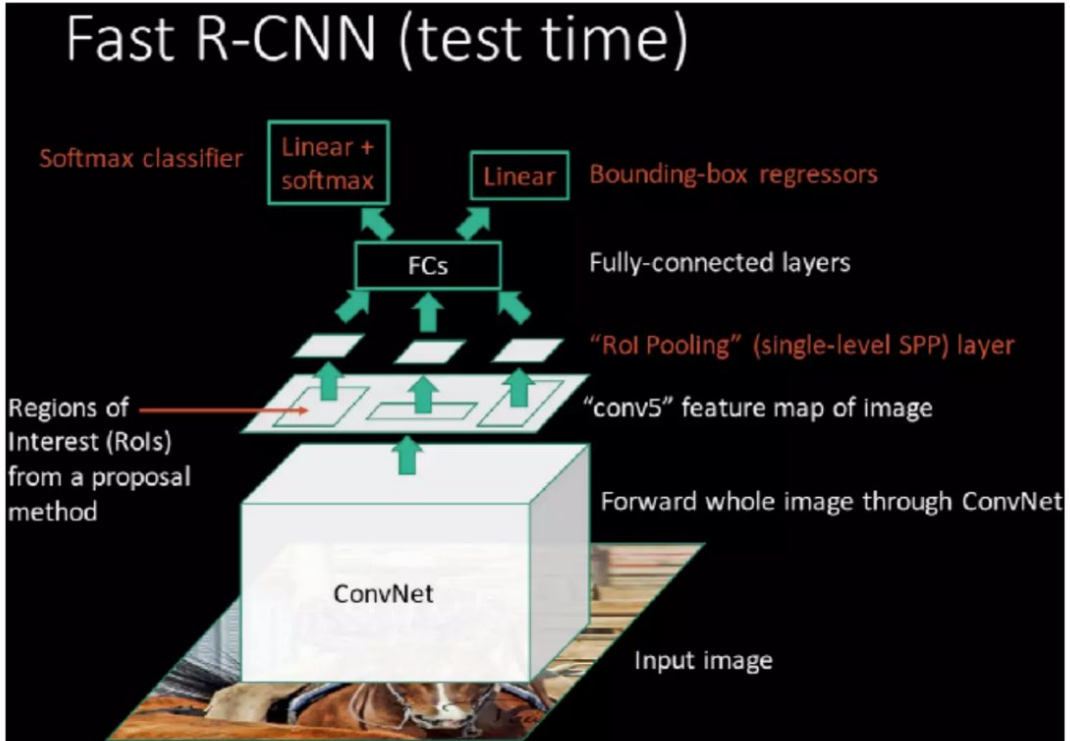
Selective search:



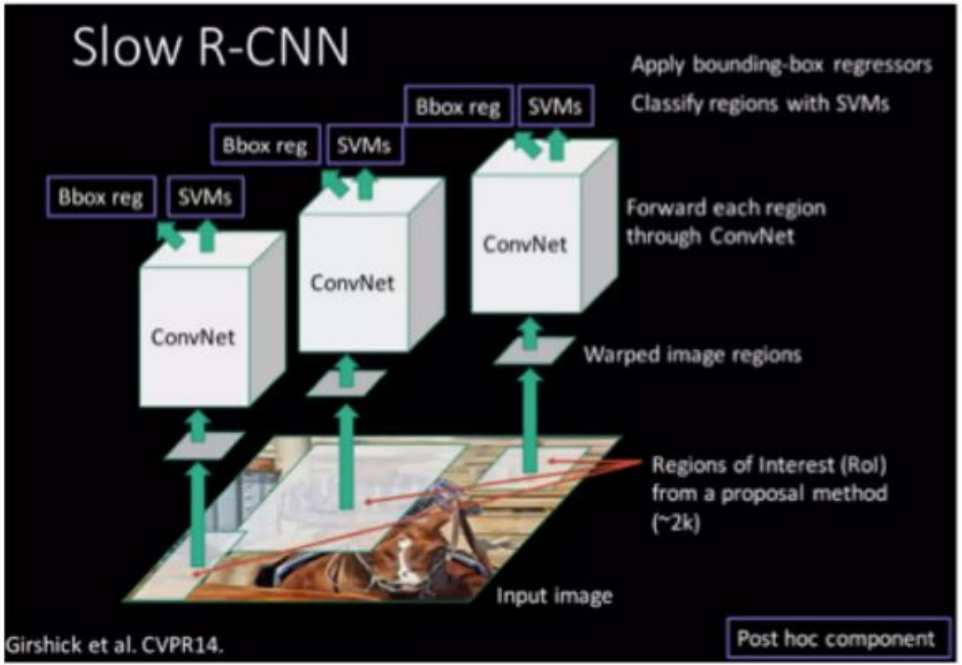
R-CNN architecture

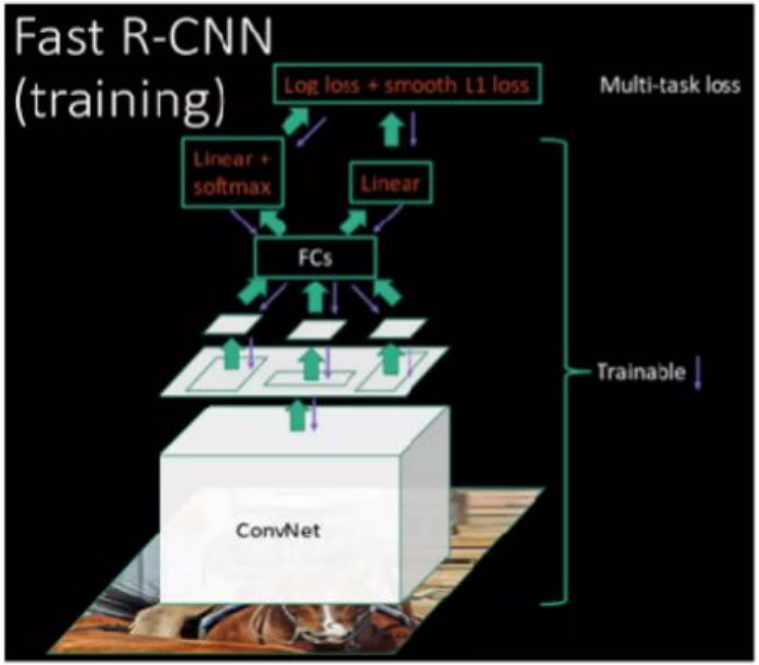


Fast R-CNN architecture



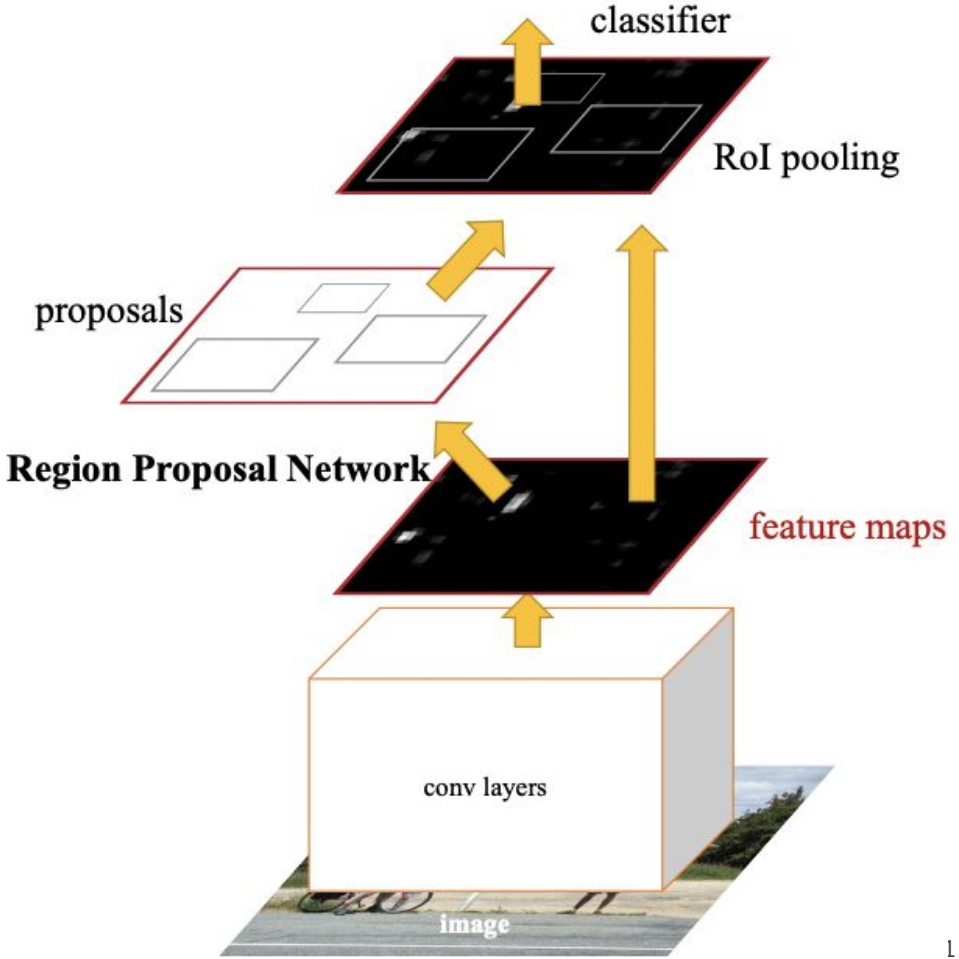
Architecture comparison between R-CNN and Fast R-CNN

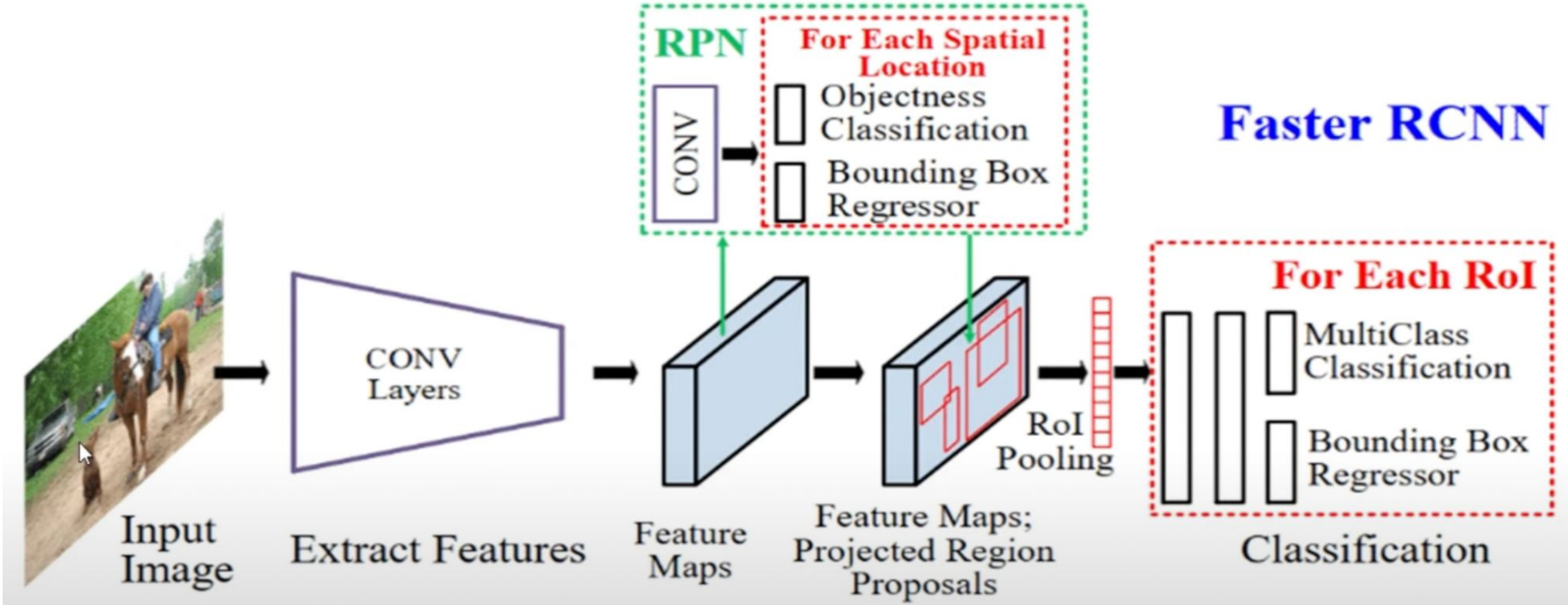




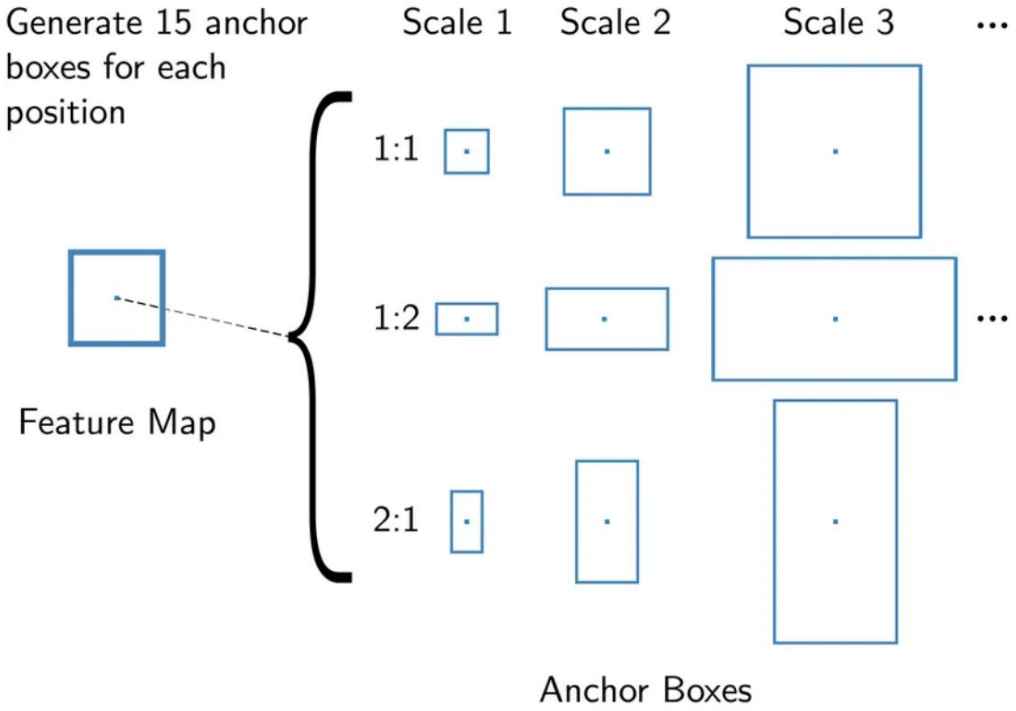
Faster R-CNN

Faster R-CNN=PRN+Fast R-CNN




RPN
For Each Spatial Location
CONV
Objectness Classification
Bounding Box Regressor
Faster RCNN
For Each RoI
MultiClass Classification
Bounding Box Regressor
CONV Layers
RoI Pooling
Input Image
Extract Features
Feature Maps
Feature Maps; Projected Region Proposals
Classification

Anchor boxes



Anchor Boxes Example

Loss function

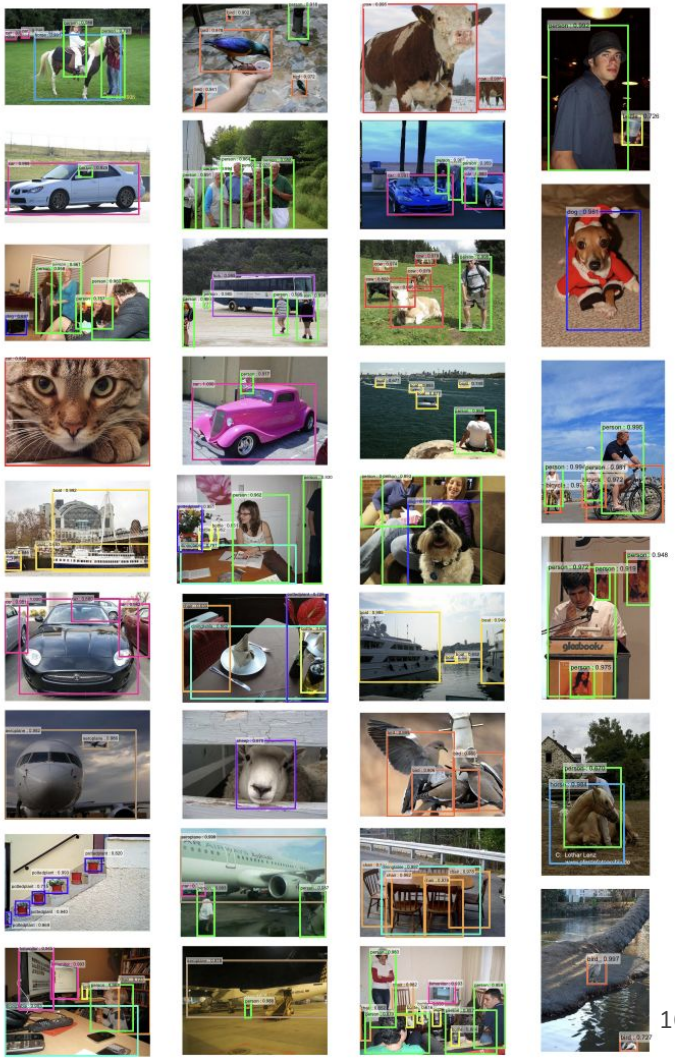
$$L(\ p_i\ ,t_i\ ) = \underbrace{\frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*)}_{\text{object/not object classifier}} + \underbrace{\lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)}_{\text{box regressor}}$$

Faster R-CNN: Results

	R-CNN	Fast R-CNN	Faster R-CNN
Test time per image (with proposals)	50 seconds	2 seconds	**0.2 seconds**
(Speedup)	1x	25x	**250x**
mAP (VOC 2007)	66.0	**66.9**	**66.9**

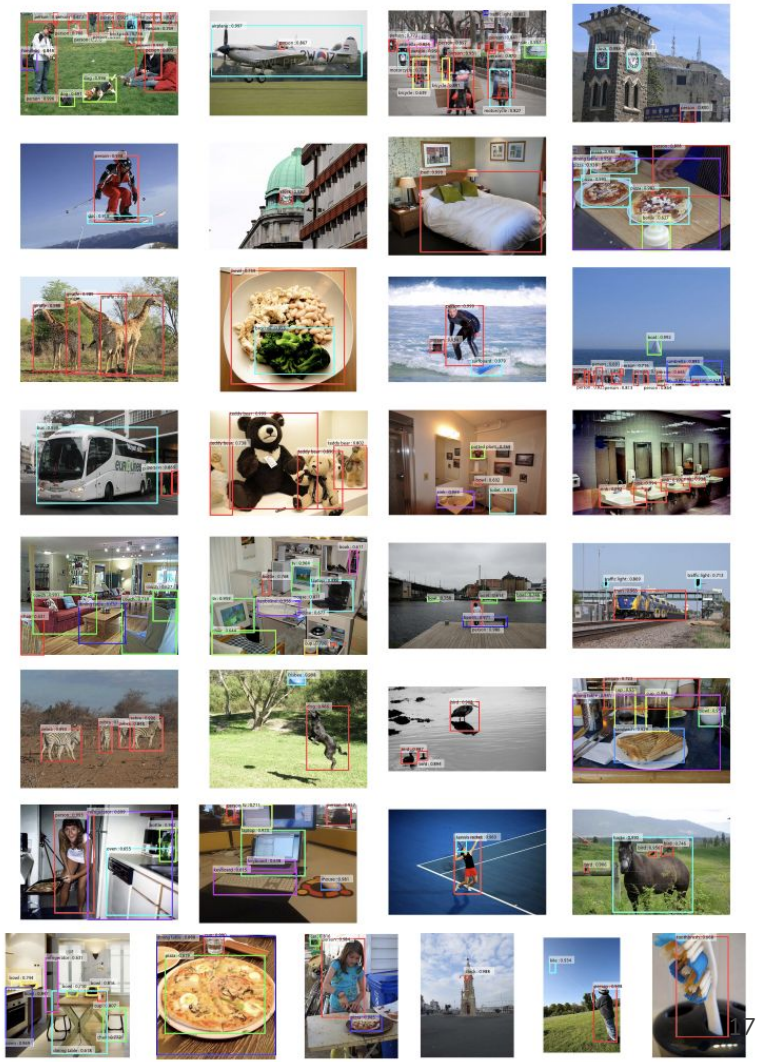
Object detection results on the PASCAL VOC 2007 test set using the Faster R-CNN system

- Selected examples of object detection results on the PASCAL VOC 2007 test set using the Faster R-CNN system. The model is VGG-16.



Object detection results on the MS COCO test-dev set using the Faster R-CNN

- Selected examples of object detection results on the MS COCO test-dev set using the Faster R-CNN system. The model is VGG-16 and the training data is COCO trainval.



References:

1. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.

2. Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

3. http://cs231n.stanford.edu/slides/2016/winter1516_lecture8.pdf