Video Object Detection using Faster R-CNN

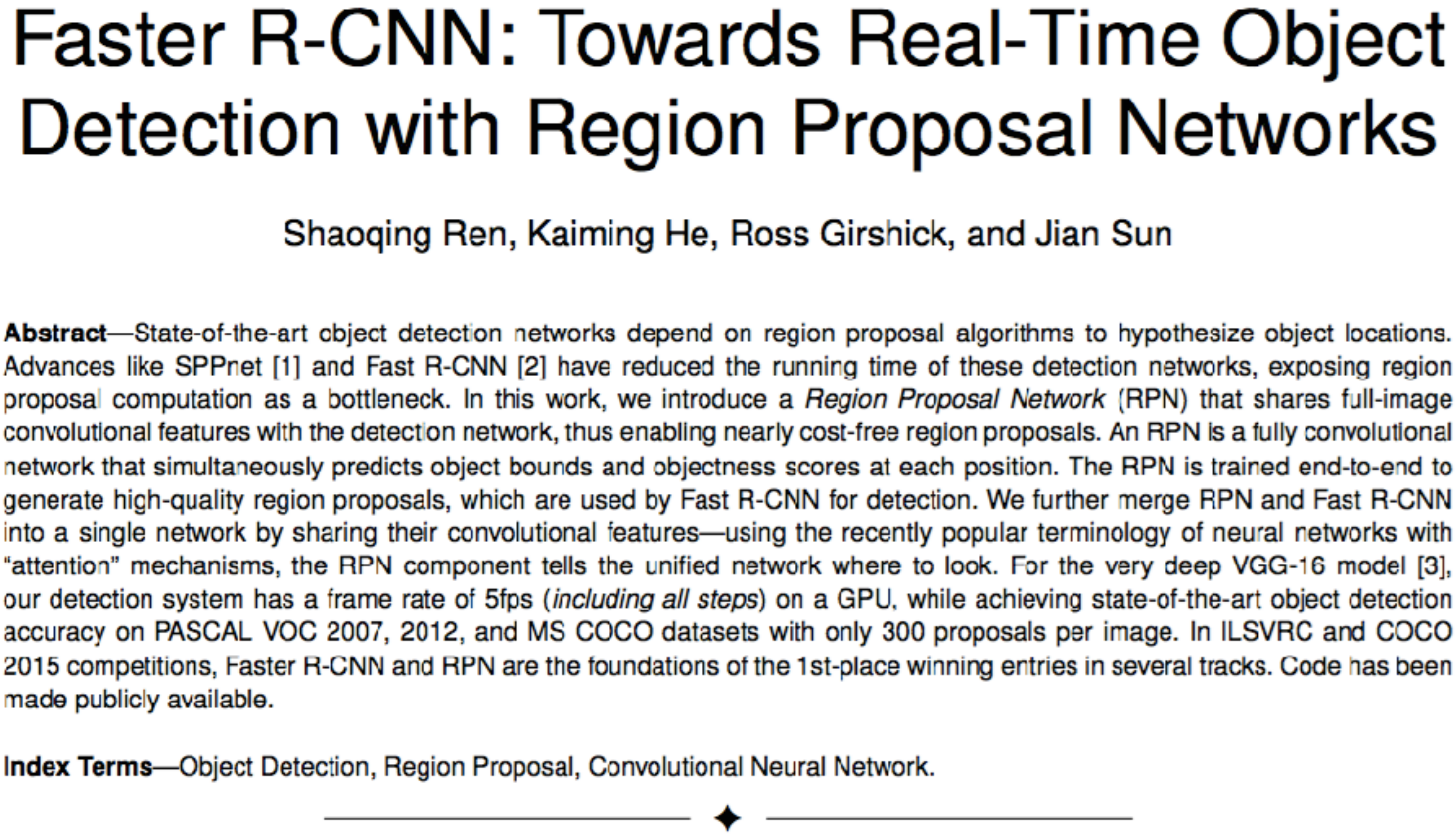
Group: 30

102061137 廖元宏

102061145 莊景堯

start from

Image Object Detection



Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun

Abstract—State-of-the-art object detection networks depend on region proposal algorithms to hypothesize object locations. Advances like SPPnet [1] and Fast R-CNN [2] have reduced the running time of these detection networks, exposing region proposal computation as a bottleneck. In this work, we introduce a *Region Proposal Network* (RPN) that shares full-image convolutional features with the detection network, thus enabling nearly cost-free region proposals. An RPN is a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position. The RPN is trained end-to-end to generate high-quality region proposals, which are used by Fast R-CNN for detection. We further merge RPN and Fast R-CNN into a single network by sharing their convolutional features—using the recently popular terminology of neural networks with "attention" mechanisms, the RPN component tells the unified network where to look. For the very deep VGG-16 model [3], our detection system has a frame rate of 5fps (*including all steps*) on a GPU, while achieving state-of-the-art object detection accuracy on PASCAL VOC 2007, 2012, and MS COCO datasets with only 300 proposals per image. In ILSVRC and COCO 2015 competitions, Faster R-CNN and RPN are the foundations of the 1st-place winning entries in several tracks. Code has been made publicly available.

Index Terms—Object Detection, Region Proposal, Convolutional Neural Network.

Faster R-CNN

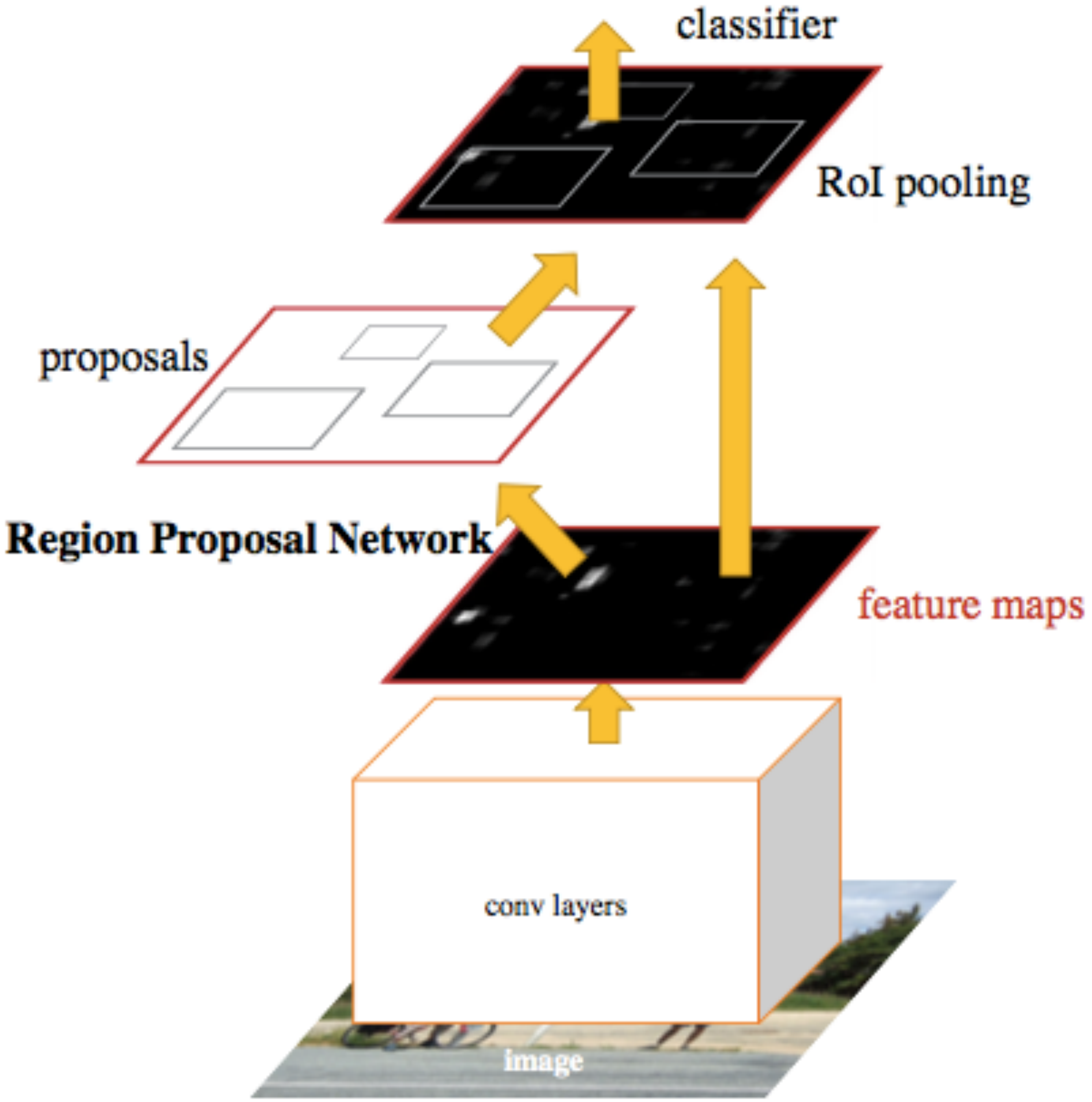


Figure 2: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.

Region Proposal Network (RPN)

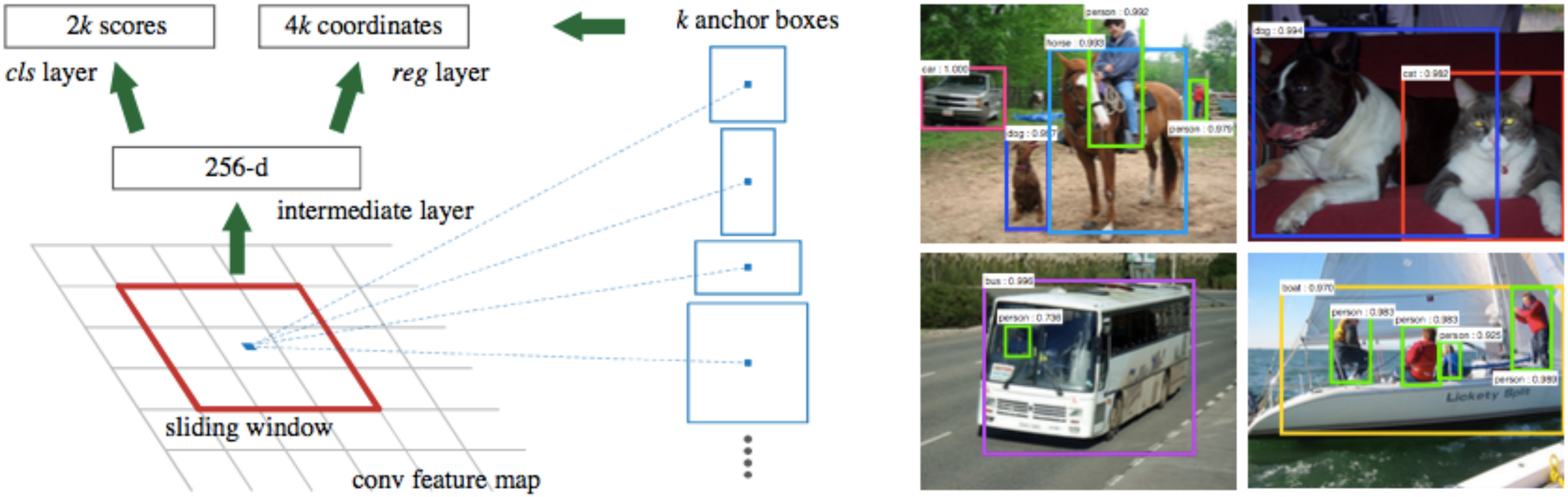
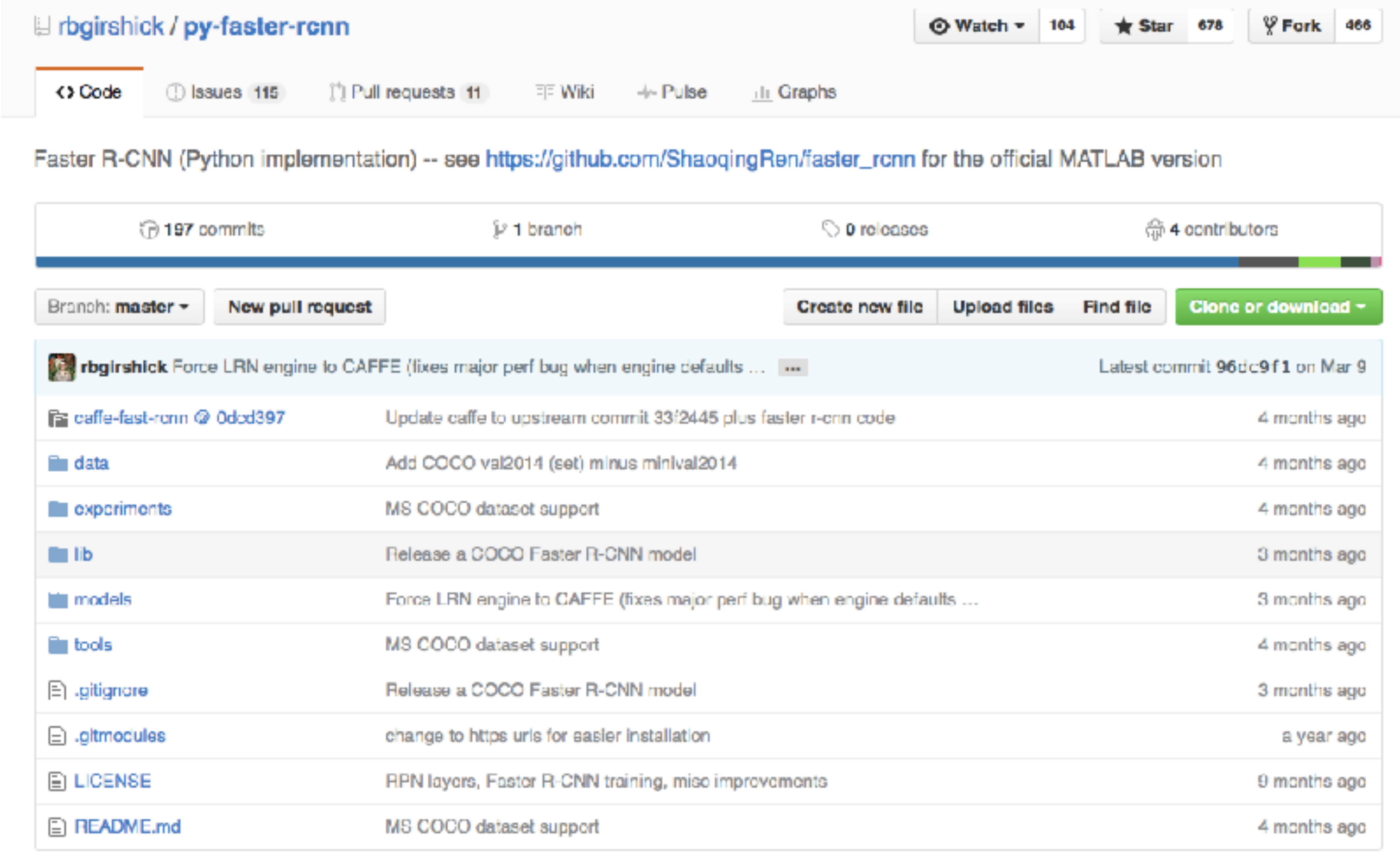


Figure 3: **Left**: Region Proposal Network (RPN). **Right**: Example detections using RPN proposals on PASCAL VOC 2007 test. Our method detects objects in a wide range of scales and aspect ratios.

Source Code on Github (framework : Caffe)



Faster R-CNN (Python implementation) -- see https://github.com/ShaoqingRen/faster_rcnn for the official MATLAB version

rbgirshick Force LRN engine to CAFFE (fixes major perf bug when engine defaults … … Latest commit 96dc9f1 on Mar 9

caffe-fast-rcnn @ 0dcd397	Update caffe to upstream commit 33f2445 plus faster r-cnn code	4 months ago
data	Add COCO val2014 (set) minus minival2014	4 months ago
experiments	MS COCO dataset support	4 months ago
lib	Release a COCO Faster R-CNN model	3 months ago
models	Force LRN engine to CAFFE (fixes major perf bug when engine defaults …	3 months ago
tools	MS COCO dataset support	4 months ago
.gitignore	Release a COCO Faster R-CNN model	3 months ago
.gitmodules	change to https urls for easier installation	a year ago
LICENSE	RPN layers, Faster R-CNN training, misc improvements	9 months ago
README.md	MS COCO dataset support	4 months ago

Extend PASCAL VOC to ImageNet

PASCAL2

Pattern Analysis, Statistical Modelling and Computational Learning

PASCAL VOC 2007
20 Categories

↓

IMAGENET

ImageNet
200 Categories

Training Details

Duration : 16 hours

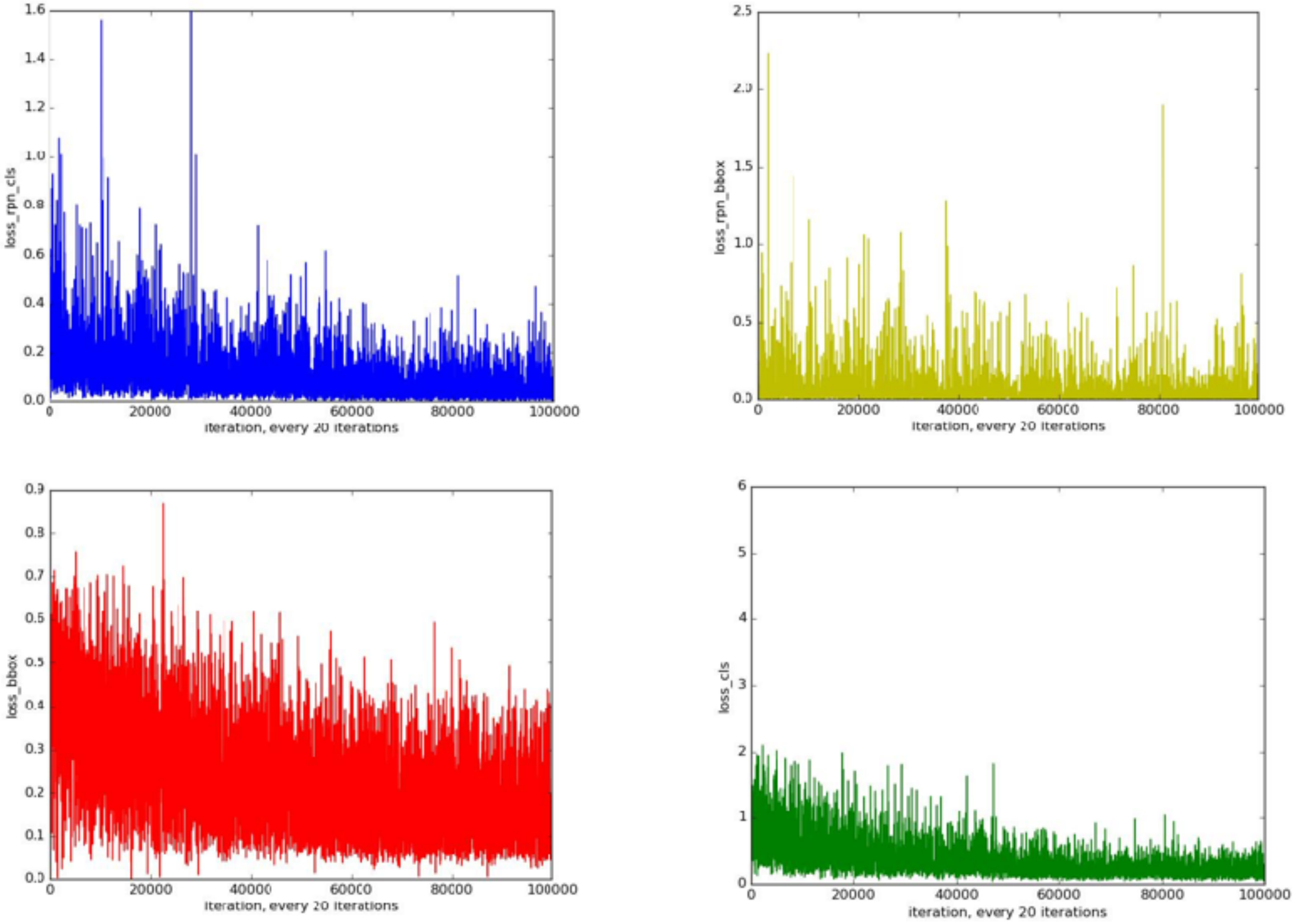
Tool : Geforce GTX Titan X

Train Data : ImageNet Validation Set 1

Final Model: 100000 iterations

Training Details

Loss :



Training Details

Accuracy:

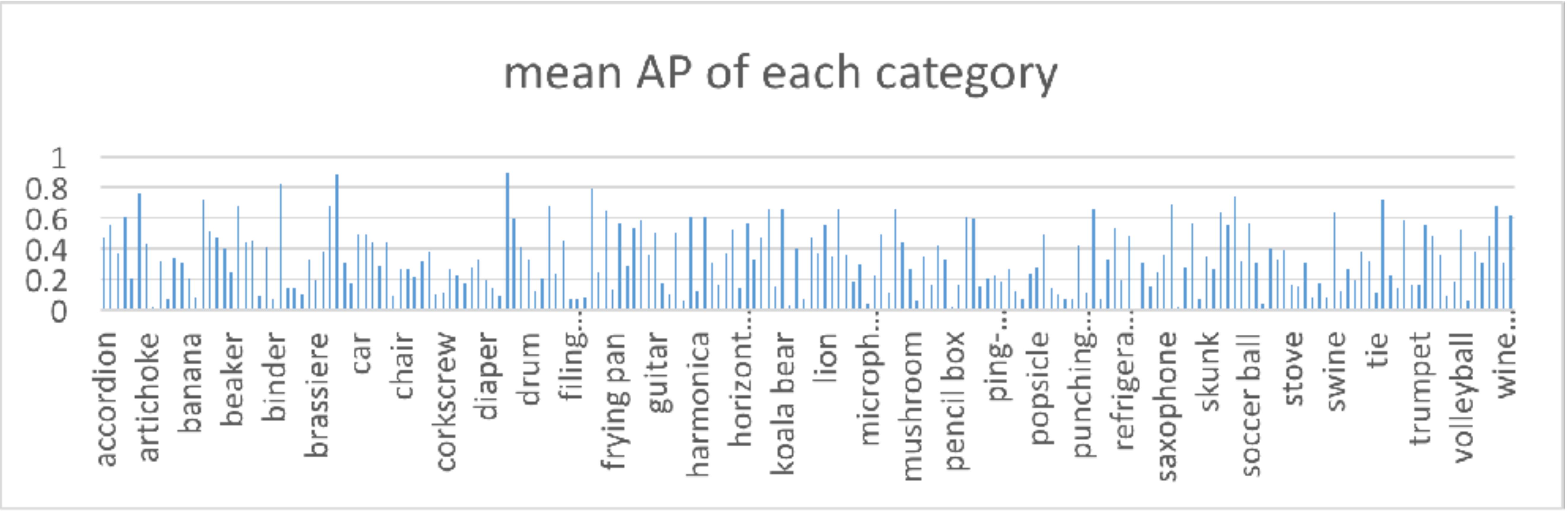
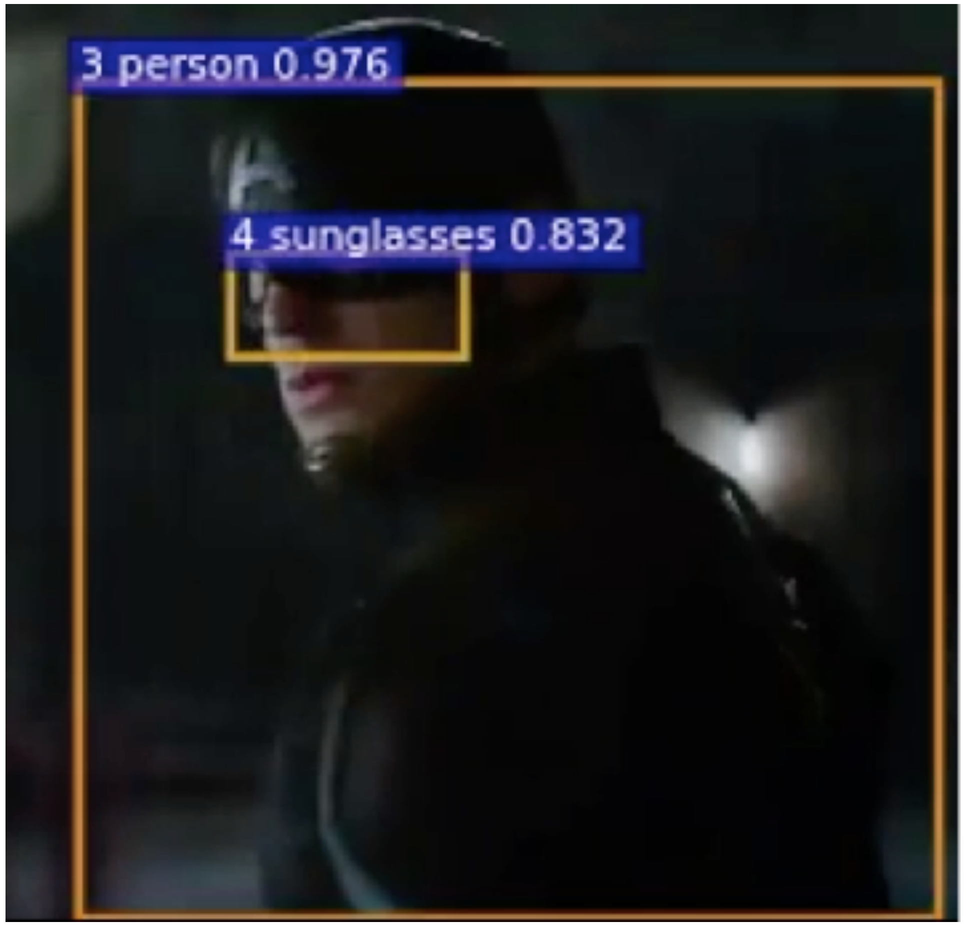


Image Detection Output



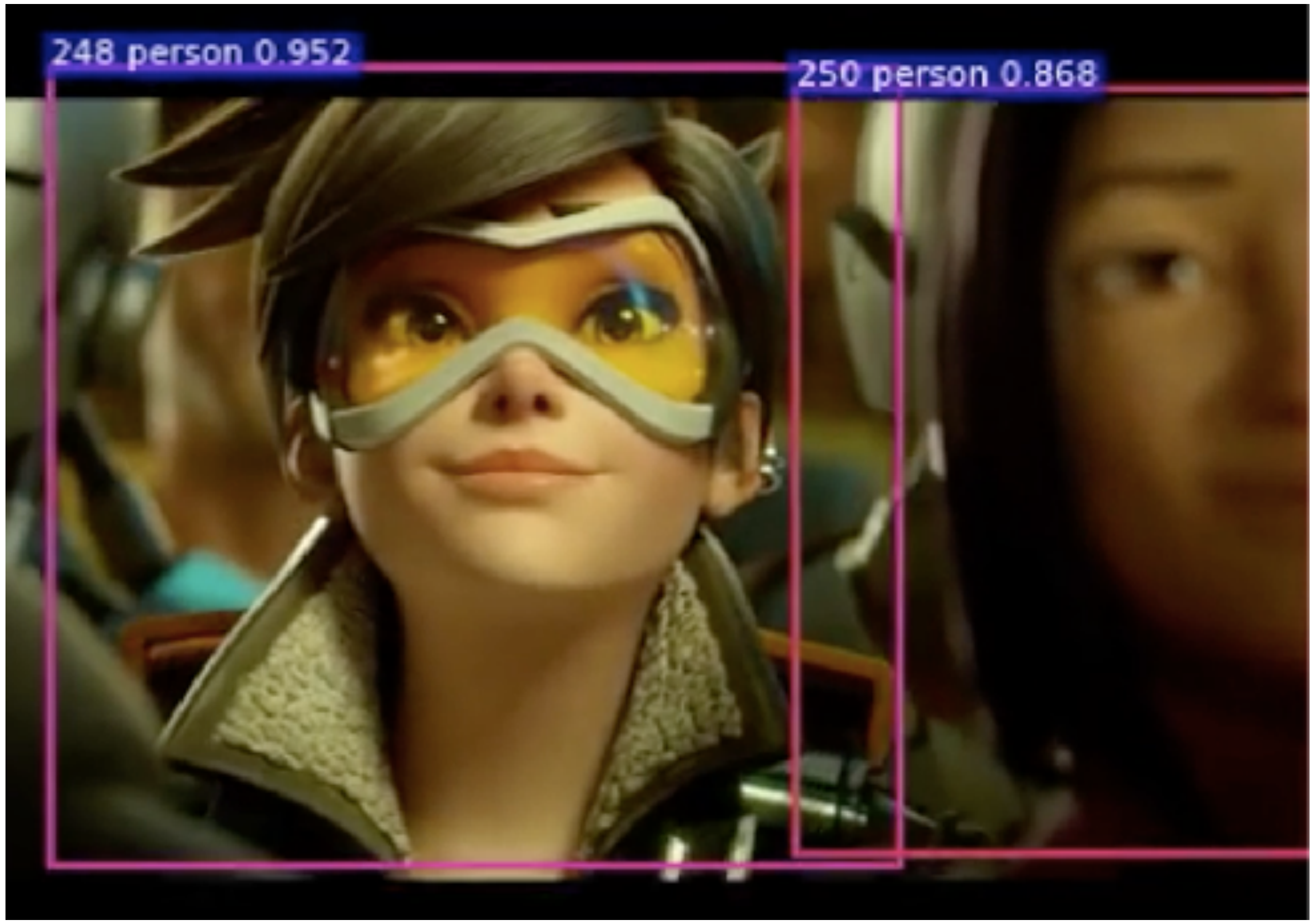
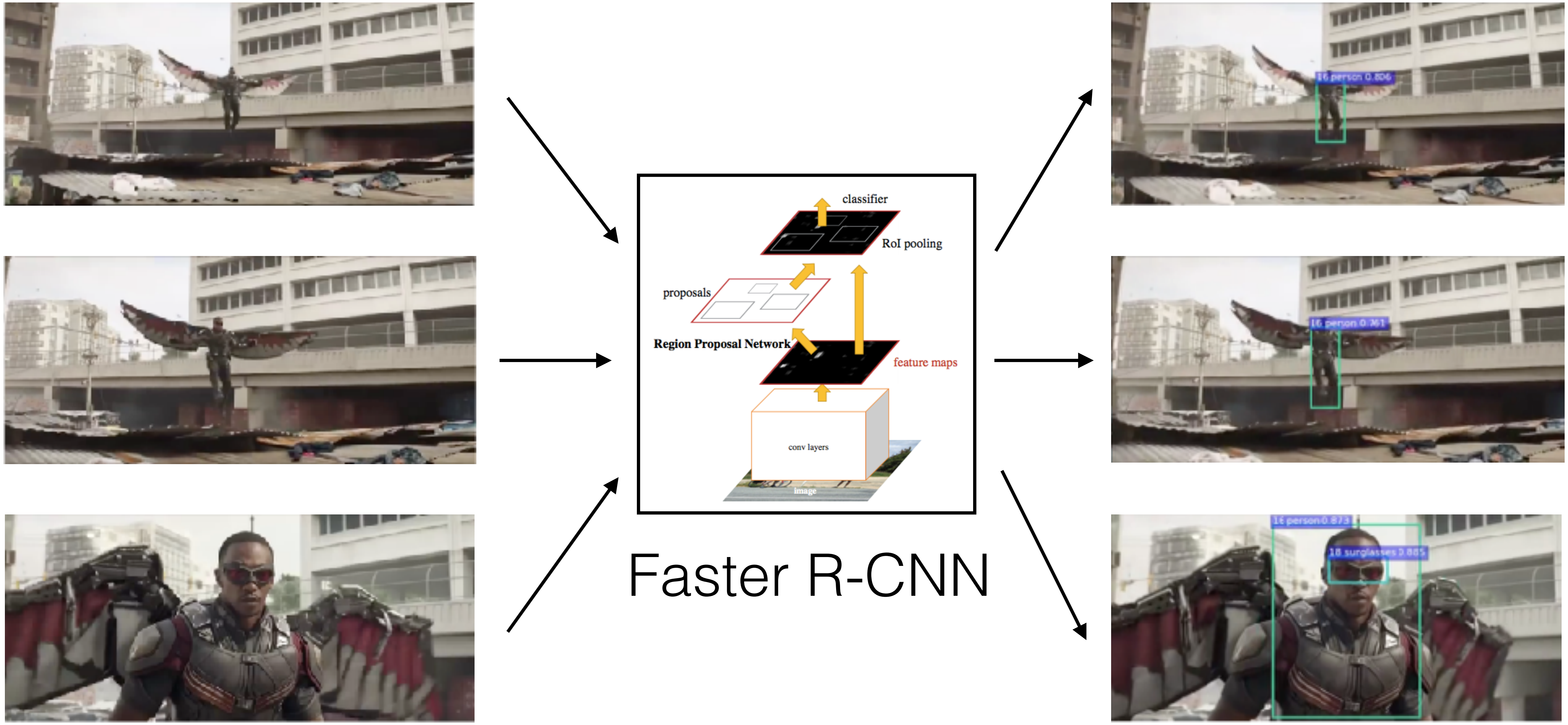


Image v.s Video

Image : Single Frame

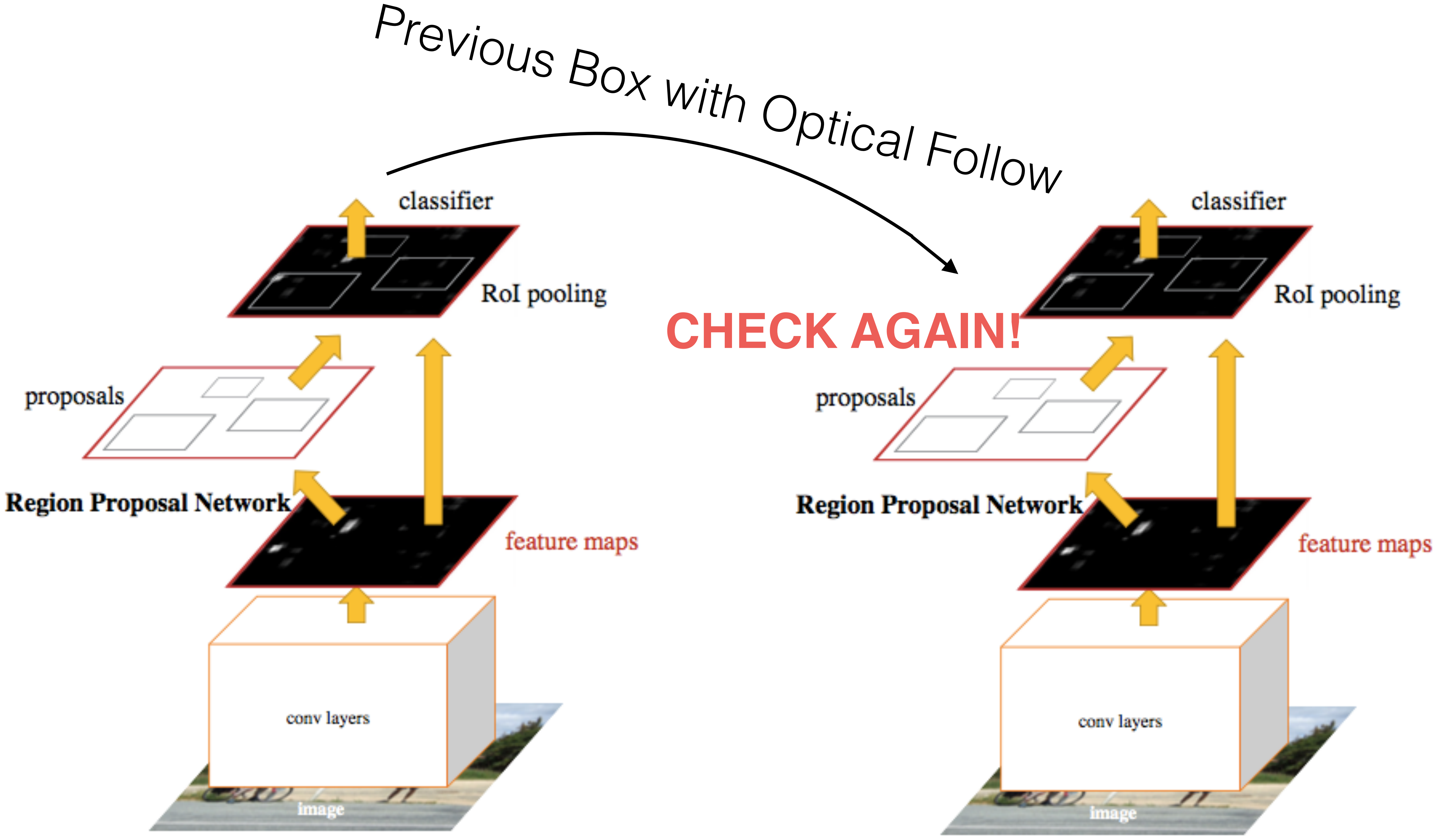
Video : Multiple Frame

Extend Image to Video



Faster R-CNN

Tracker



Previous Box with Optical Follow

CHECK AGAIN!

frame t

frame t+1

Contribution

1.Set up the environment

2. Modify Faster R-CNN to adapt **IMAGENET**

3. Extend Image Object Detection to **Video**

4. Construct **Tracker** to make the bounding box stable

5. Result **Visualisation**

Demo: Captain America

https://www.youtube.com/watch?v=zebSqDt6oMM

Q&A