

Object Detection Using Convolutional Neural Networks

Progress

Guided by:
Dr. Sushil Kumar

Amlaan Bhoi
A2305213303
7CSE4-Y

Dataset

Large Scale Visual Recognition Challenge 2016

<http://image-net.org/challenges/LSVRC/2016/browse-det-synsets>

Search

Massive plantigrade carnivorous or omnivorous mammals with long shaggy coats and strong claws

1688 pictures 91.45% Popularity Percentile Words IDs

Treemap Visualization Images of the Synset Downloads

Numbers in brackets: (the number of synsets in the subtree)

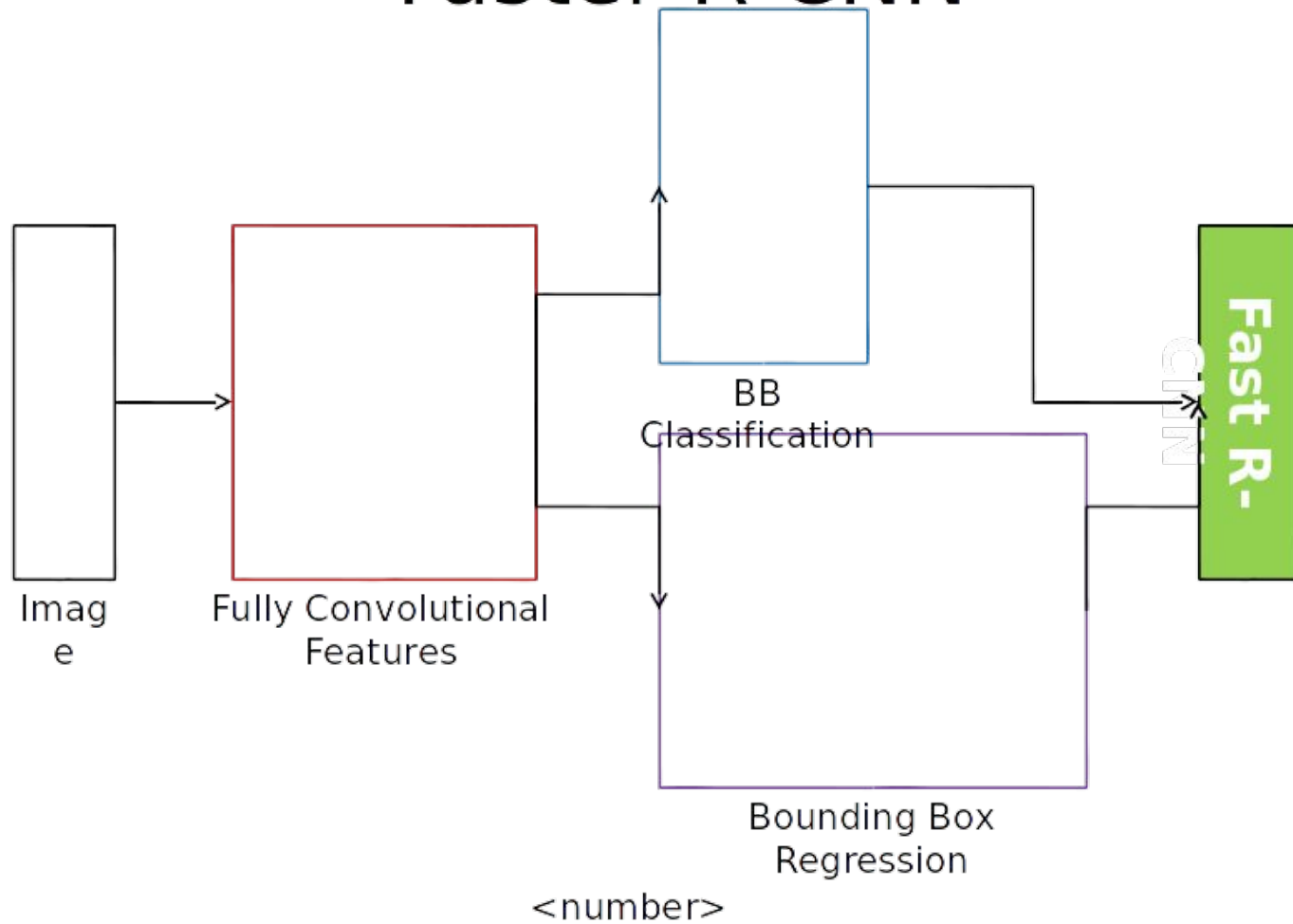
- ImageNet 2011 Fall Release (32326)
 - plant, flora, plant life (4486)
 - geological formation, formation (17)
 - natural object (1112)
 - sport, athletics (176)
 - artifact, artefact (10504)
 - fungus (308)
 - person, individual, someone, some
 - animal, animate being, beast, brute
 - invertebrate (766)
 - homeotherm, homoiotherm, home
 - work animal (4)
 - darter (0)
 - survivor (0)
 - range animal (0)
 - creepy-crawly (0)
 - domestic animal, domesticated
 - molt, moulder (0)
 - varmint, varment (0)
 - mutant (0)
 - critter (0)
 - game (47)
 - young, offspring (45)
 - poikilotherm, ectotherm (0)
 - herbivore (0)
 - peeper (0)
 - pest (1)
 - female (4)
 - insectivore (0)
 - net (0)

Images of the Synset

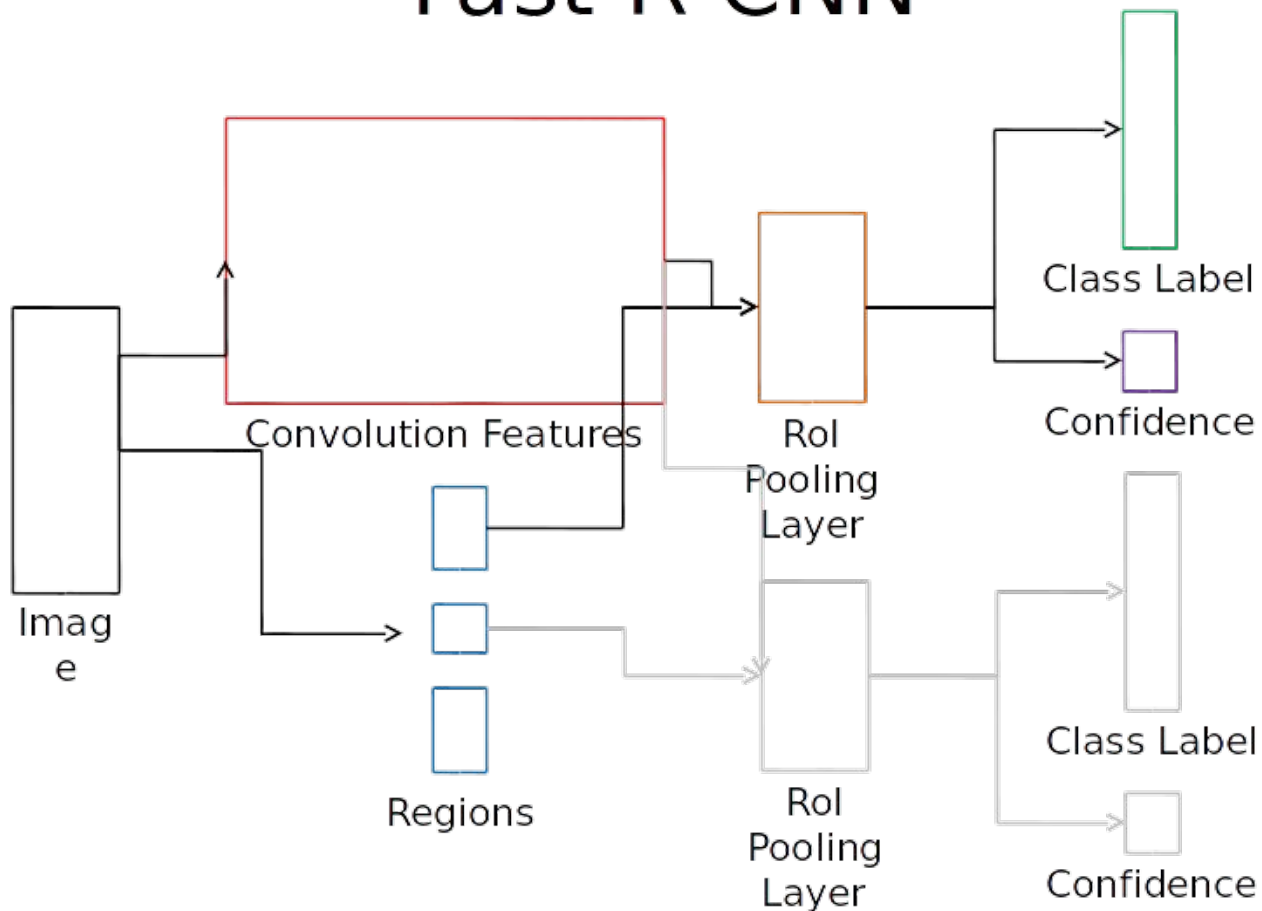
*Images of children synsets are not included. All images shown are thumbnails. Images may be subject to copyright.

Prev 1 2 3 4 5 6 7 8 9 10 ... 48 49 Next

Faster R-CNN



Fast R-CNN



<number>

Models

Python Implementation

<https://github.com/rbgirshick/py-faster-rcnn>

Matlab Implementation

https://github.com/ShaoqingRen/faster_rcnn

Issues?

GPU Implementation

GPU: Titan, Titan Black, Titan X, K20, K40, K80.

1. Region Proposal Network (RPN)
 - 2GB GPU memory for ZF net
 - 5GB GPU memory for VGG-16 net
2. Object Detection Network (Fast R-CNN)
 - 3GB GPU memory for ZF net
 - 8GB GPU memory for VGG-16 net

CPU (56 seconds for 10k images) vs GPU (26 seconds for 10k images)

Trial* using Caffenet (AlexNet variant ~2012)

Training file: testy.py

Output: testy_run_log.txt

Results:

Mean-subtracted Values:

[('B', 104.0069879317889), ('G',
116.66876761696767), ('R',
122.6789143406786)]

Predicted Class:

281

Output Label:

n02123045 tabby, tabby cat

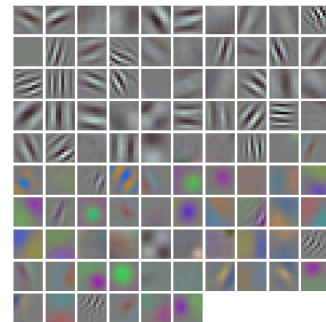
CPU Forward Pass Time:

6.66855120659

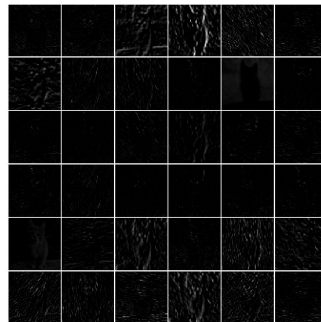
* Image classification



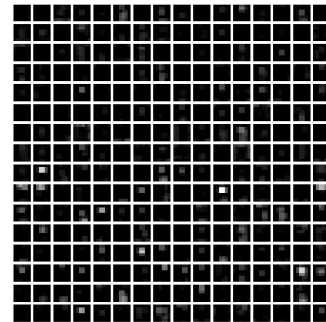
Original Image



filter1



conv1



pool5

Next Steps

- Faster-RCNN implementation
- H/W requirements (AWS possible solution)
- Pre-trained Model Zoo

(<https://github.com/BVLC/caffe/wiki/Model-Zoo>)