Introduction to Object Detection with Convolution Neural Networks using the TensorFlow API

Abel Brown

August 10, 2017

Outline

What this lab is

What this lab is not

What we hope you take away

Object Detection API in TensorFlow

The MSCOCO dataset

Launching the Lab

What this lab is

- Get going fast!
- Leverage the Google Object Detection API in TensorFlow
- Gain experience using state-of-the-art object detection network architectures and understand engineering trade-offs
- Gain familiarity with the Microsoft COCO dataset
- Build intuition and codes for evaluating detector performance

What this lab is not

- Not an intro to machine learning from first principles
- Not a rigorous formalism of convolutional neural networks
- Not a survey of all the features and options of TensorFlow

Key Takeaways

- Baseline for state-of-the-art object detection w/ deep learning
- Understanding of the complexities in evaluating performance
- Familiarity with industry standard COCO dataset
- Everything you need to start leveraging top performing deep learning based object detection models in your own work

Object Detection API in TensorFlow

- The TensorFlow Object Detection API is an open source framework built on top of TensorFlow
- This API makes it easy to construct, train and deploy object detection models
- The API includes a selection of trainable detection models:
  - Single Shot Multibox Detector (SSD) with MobileNets
  - SSD with Inception V2
  - Region-Based Fully Convolutional Networks (R-FCN) with Resnet 101
  - Faster RCNN with Resnet 101
  - Faster RCNN with Inception Resnet v2
- All models have been pre-trained on the COCO dataset and are ready for use out-of-the-box
- Convenient local training scripts included as well

Object Detection API in TensorFlow

- The SSD models that use MobileNet are lightweight, so that they can be comfortably run in real time on mobile devices
- The Faster RCNN models are much more computationally intensive but are significantly more accurate

The Microsoft COCO Dataset

- In this lab we're going to use the Common Objects in Context (COCO) dataset from Microsoft .
- COCO is a new image recognition, segmentation, and captioning dataset. COCO has several features:
  - Object segmentation
  - Recognition in Context
  - Multiple objects per image
  - More than 300,000 images
  - More than 2 Million instances
  - 80 object categories
  - 5 captions per image
  - Keypoints on 100,000 people

Lets Get Started . . .

- Navigate to nvlabs.qwiklab.com[1]
- Login or create a new account
- Select the "DLI Instructor-Led Labs? Class"
- Find the lab called "Introduction to Object Detection with TensorFlow", select it, click Select, and finally click Start
- When prompted, enter your 16-digit token code
- After a short wait, lab Connection information will be shown
- Please ask Lab Assistants for help!

[1] Chrome browser usually works best