NYU-MMVC-LAB Weekly Report

CHENG QIAN

May 1, 2017

1 Motivation

To fully understand the methods and technologies behind the TextBox detection (Liao et al. 2017) and replicate its results in Python TensorFlow environment, and then proceed with other researches in computer vision areas.

2 Reading

This week I've primarily focused on Tensorflow study, the reading materials covered

- MNIST application with simple softmax classifier
- MNIST appliaction with convolutional neural network classifier
- Iris Data with deep neural network classifier
- Loop training, graph construction and feeding
- Sharing variables, threading and queues

Aside from that, I've also read the TextBox paper to familiarize myself with the topic, outlining the key components of the model structure for next week's study.

- Based on a text recognition algorithm CRNN (Shi, Bai, and Yao, 2015) which estimates sequence probability conditioned on input image
- Word-based text detection with VGG-16 architecture (Simonyan and Zisserman 2014)
- Multiple output layers (called *text-box layers*) are inserted after the last and some intermediate convolutional layers, simultaneously predicting text presence and bounding boxes
- Non-maximum suppression (NMS) process