Proposal

Investigation into Detecting Street Furniture using Computer Vision

Jana Wessels
WSSJAN006

Supervised by:
Prof. Julian Smit

March 20, 2018

Contents

1 Introduction

Currently the world is experiencing data capture and collection on a scale never seen before, but this information has little to no real value until it has been processed into valuable information. This process can be costly and time consuming, thus this project focuses on one example of data gathering and how to streamline and automate the system that turns it into useful information.

Cities all around the world are growing a rapid rate and the management of their infrastructure and services are becoming more and more daunting. Thus this project focuses on designing a system that can create a spatial database by detecting objects in georeferenced oblique images and categorising the objects.

2 Topic

An investigation into multi-object recognition and detection of street furniture within a series of images and the identification of the real-world coordinates of the objects.

3 Statement of the Problem

Cities have been relying on paper maps and aggregated data to manage its street furniture. However management of the thousands of street furniture which is present in all cities, big or small, can pose a big challenge and a number of questions remain on how to solve this problem. The creation of a spatial database containing all the traffic lights, street signs, electric boxes, rubbish bins and more would highly beneficial to any city management team. Such a database would allow for route planning by the service providers, accurate counts on the amount of traffic lights in a city and more. All of which would ease many issues involved with maintenance and service projects.

However what still needs to be investigated is how the different aspects; multi-object detection, spatial location and a spatial database can be combined to form a coherent and effective work flow. The accuracy of the object detection also needs to be investigated to see if it would be appropriate for this type of work.

4 Aim

The aim of this project are to create a program that will identify and detect street furniture and its location in space. The program should also populate a spatial database, which should be compatible with GIS software, containing all the street furniture detected.

5 Objectives

The objectives of this project are to create a neural network capable of uniquely identifying a set list of street furniture within a certain level of certainty. The positions of the objects identified by the neural network should also be found to a certain degree of precision. This needs to be accomplished by using images which have been georeferenced to control points positioned along the route.

6 Hypotheses

By using a retrained TensorFlow Object Detection API street furniture in images can be located and identified to a high degree of accuracy.

7 Research Questions

1. Can multiple items of street furniture and their corresponding image location be identified from a single image using the re-trained Tensorflow Object Detection API?
2. What accuracy of object detection can be achieved by using the proposed method?

8 Scope

This project focuses on the image recognition aspect of the work flow thus the control of the images will be done manually and not automated. The test case site will also be limited to one portion of road in the city due to time and equipment constraints present when collecting the data.

9 Research Steps and Methods

- Investigate neural network packages and other reports, projects and papers pertaining to the underlying theory of object recognition and location.
- Select the appropriate neural network and object detection methods that will be used to create the object detection program.
- Train neural network using a sample data set to recognise the chosen set of street furniture and test its accuracy on a set of test images.
- Collect city images and set-up control for them using a traverse through town survey marks.
- Design format of the spatial database on Microsoft Access and code program to automatically add information to the spatial database.
- Do pre-processing on the collected images to prepare them for the object recognition.
- Run collected images through recognition protocol and evaluate accuracy.
- Calculate location of objects in the images using the tie-points that exist in both the image and the world system.

10 Literature Review and Theory

10.1 Object Detection Methodologies

A well established methodology within computer vision is object detection. This is a catch all phrase used to describe the techniques used to classify and locate objects within an image. There are many different ways to implement object detection and in recent years a with the help of convolutional neural networks (CNN) a lot of progress has been made in the field. A number of modern object detectors are based on CNN's such as YOLO (You Only Look Once), R-FCN (Region-based Fully Convolutional Networks), Faster R-CNN (Region-based Convolutional Neural Network) and SSD (Single Shot Multibox Detector). The deciding factor on which method would be deemed suitable for use is if speed or accuracy is the main goal of the object detection.

The object detection method of Darknet YOLO is a very recently developed method that is state-of-the art and capable of detecting over 9000 classes in real time (Redmon and Farhadi, 2016) . A trade-off for YOLO's speed is its accuracy which under preforms compared to some of the other slower detection models (Mustamo, 2018).

In Faster R-CNN the image detection happens in two stages. The stages are known as the region proposal network(RPN) and box classifier (Huang et al., 2017b). During the RPN stage the images are processed by a feature extractor which is used to predict class agnostic box proposals. During the box classifier stage these predicted boxes are used to crop features.

R-FCN uses similar concepts to Faster R-CNN differs as crops are taken from the final layer of features prior to prediction (Huang et al., 2017b). It has also been shown that this model can achieve comparable accuracy to R-CNN with similar or faster speeds.

SSD uses a single image feed-forward convolutional neural network to predict the classes (Huang et al., 2017b) and is one of the fastest algorithms for object detection currently available (Zheng et al., 2018). SSD predicts category scores, confidence level of classification, and offsets for the predicted bounding boxes. SSD uses a end-to-end training method that simplifies the process and improves the accuracy/speed trade off (Liu et al., 2015).

The TensorFlow Object Detection API was created to ease the process of identifying multiple objects in a single image and is an open sources framework built on top of Google's TensorFlow. TensorFlow is a software library (Huang et al., 2017a) that can be used for a range of tasks which include machine learning and neural networks. The API eases the process of constructing, training and deploying object detection models.

10.2 Previous usage of TensorFlow in Object Detection

There has been many different applications of TensorFlow for object detection even though its public release was only in 2017.

One study focused on using TensorFlow for object detection in sports and specifically focused on the detection of soccer players on a pitch during a game (Mustamo, 2018). The particular project did not use images but rather focused on object detection in a video, which largely uses the same methods but just applies them to a video dataset.

Another project which is more relevant to this project focused on the object detection and tracking within a city for traffic analysis (Wei et al.). This detection included objects such a street signs and traffic lights which is within the scope of this project. Again video was used rather than images due to the traffic tracking aspect of the project.

Thus TensorFlow has been proven to work for custom object detection project given that it is set up correctly and trained on the appropriate dataset.

References

Jonathan Huang, Vivek Rathod, Derek Chow, Chen Sun, Menglong Zhu, Alireza Fathi, and Zhichao Lu. Tensorflow Object Detection API, June 2017a. URL `https://github.com/tensorflow/models/tree/master/research/object_detection`.

Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, and Kevin Murphy. Speed/accuracy trade-offs for modern convolutional object detectors. *eprint arXiv:1611.10012*, April 2017b. URL `https://arxiv.org/abs/1611.10012`.

Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng Yang Fu, and Alexander C. Berg. SSD: Single Shot MultiBox Detector. *CoRR*, abs/1512.02325, 2015. URL `http://arxiv.org/abs/1512.02325`.

Prikko Mustamo. *Object detection in sports: TensorFlow Object Detection API case study.* Bachelor's Thesis, University of Oulu, Finland, January 2018.

Joseph Redmon and Ali Farhadi. YOLO: Better, Faster, Stronger. *eprint arXiv:1612.08242*, December 2016.

Yi Wei, Nenghui Song, Lipeng Ke, Ming-Ching Chang, and Siwei Lyu. Street Object Detection / Tracking for AI City Traffic Analysis.

Liwen Zheng, Canmiao Fu, and Yong Zhao. Extend the shallow part of Single Shot MultiBox Detector via Convolutional Neural Network. *eprint arXiv:1801.05918*, January 2018. URL `https://arxiv.org/ftp/arxiv/papers/1801/1801.05918.pdf`.

Bibliography

Jonathan Huang, Vivek Rathod, Derek Chow, Chen Sun, Menglong Zhu, Alireza Fathi, and Zhichao Lu. Tensorflow Object Detection API, June 2017a. URL `https://github.com/tensorflow/models/tree/master/research/object_detection`.

Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, and Kevin Murphy. Speed/accuracy trade-offs for modern convolutional object detectors. *eprint arXiv:1611.10012*, April 2017b. URL `https://arxiv.org/abs/1611.10012`.

Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng Yang Fu, and Alexander C. Berg. SSD: Single Shot MultiBox Detector. *CoRR*, abs/1512.02325, 2015. URL `http://arxiv.org/abs/1512.02325`.

Andres Milioto and Cyrill Stachniss. Bonnet: An Open-Source Training and Deployment Framework for Semantic Segmentation in Robotics using CNNs, February 2018. URL `https://arxiv.org/abs/1802.08960v1`.

Prikko Mustamo. *Object detection in sports: TensorFlow Object Detection API case study*. Bachelor's Thesis, University of Oulu, Finland, January 2018.

Joseph Redmon and Ali Farhadi. YOLO: Better, Faster, Stronger. *eprint arXiv:1612.08242*, December 2016.

Yi Wei, Nenghui Song, Lipeng Ke, Ming-Ching Chang, and Siwei Lyu. Street Object Detection / Tracking for AI City Traffic Analysis, a.

Yi Wei, Nenghui Song, Lipeng Ke, Ming-Ching Chang, and Siwei Lyu. Traffic Analysis Using Visual Object Detection and Tracking, b. URL `https://pdfs.semanticscholar.org/8b64/93a65285fb72252221e7ef8ad1091bd96b32.pdf`.

Liwen Zheng, Canmiao Fu, and Yong Zhao. Extend the shallow part of Single Shot MultiBox Detector via Convolutional Neural Network. *eprint arXiv:1801.05918*, January 2018. URL `https://arxiv.org/ftp/arxiv/papers/1801/1801.05918.pdf`.