Master Thesis

Czech Technical University in Prague

F3

Faculty of Electrical Engineering
Department of Cybernetics

Deep neural network for city mapping using Google Street View data

Varun Burde

Supervisor: Ing. Michal Reinštein, Ph.D.
May 2019

Acknowledgements   Declaration

Abstract

With the advancement computation power and large datasets finally made a huge improvement of Deep neural network leading many widespread applications. One of such application is solving computer vision problems like classification and segmentation. Also Competition like ImageNet[1] Large Scale Visual Recognition Challenge, took the solution to next level, in some cases classification is better than Human.

This report describes the evaluation of pre-trained deep neural network on Google StreetView Images[2]. Pretrained model used is Mask RCNN[3] .Implementation is done in Python[4] using Keras and TesnsorFlow-GPU framework. User interface for the application execution, processing of the input images and visualization of the results is realized using Google Colab[5] with repository in GIT[6].

A pipeline was created for the task, user provide the parameters like coordinates for the Google StreetView API [2]. Python[4] script downloads the available images to that location then images were pre-processed to fit into classifier. Classification is done on the Images depending on the architecture of the Neural Network.A vectorized map is generated as a result of classified Images and put into map using Folium.

Map is represented in various form as heat-map for better understanding of scenario. Object markers in map are the objects classified with the Mask RCNN .

Keywords:

Abstrakt

Rozvíjíme . . .

Klíčová slova:

Contents

Figures   Tables

Chapter 1

Introduction

1.1 Aim and objective of project

The aim is to design, implement and experimentally evaluate a deep neural network based solution for city mapping using Google Street View images. The proposed software solution should allow the user to request Google Street View imagery for any given location specified as geojson, perform analysis and feature extraction using deep neural network(s) and output vectorized description projected and visualized over an underlying map. User interface for the application execution, processing of the input images and visualization of the results should be realized using Google Colab to utilize Google TPUs. Existing pre-trained models should be explored first, thorough experimental evaluation on publicly available datasets should follow. Comparison with related state-of-the-art work is integral part of the work and should be presented in the final thesis. Recommendation: implementation should be done in Python, using Keras and TensorFlow frameworks.

1.2 Overview of Project

The project is software solution with the aim of achieving the task described above. There are few important part of the project which is implemented as a individual task in order to make it easy to debug.

The core implementation of project is depends on the Google Street View [2] imaginary. The API consist of helper function and parameter to get the images at certain location available in Google street view.Task was to Use API with Python [4] script to download the required images under the area by the coordinates provided .

Then this Images are processed with Deep Neural Network for classification and segmentation of objects, Mask R-CNN [3] with pre-trained model consist of 82 classes from matterport [?]s used . The resulted Images are classified and segmented with the bounding box with there confidence. For purpose of project ,the confidence threshold is kept 0.7 for classification.

Experimentation is done to localized the object in map.Map is created with Folium [?]nd represented in different ways ,the detected objects are put with

different marker in the map and view is represented as heatmap to get better understanding of density of object and its type in the scenario

Chapter 2

State of the art

2.1 Layered Interpretation of Street View Images

Proposed a layered street view model to encode both depth and semantic information on street view images for autonomous driving. They propose a 4-layer street view model, layers encode semantic classes like ground, pedestrians, vehicles, buildings, and sky in addition to the depths. The only input to our algorithm is a pair of stereo images. Deep neural network was used to extract the appearance features for semantic classes.

2.2 Building Instance classification using street view images [7]

Land-use classification based on spaceborne or aerial remote sensing images has been extensively studied over the past decades. Proposed a general framework for classifying the functionality of individual buildings. The proposed method is based on Convolutional Neural Networks (CNNs) which classify façade structures from street view images, such as Google StreetView[2], in addition to remote sensing images which usually only show roof structures. Geographic information was utilized to mask out individual buildings, and to associate the corresponding street view images. In addition, the method was applied to generate building classification maps on both region and city scales of several cities in Canada and the US

2.3 Automatic Discovery and Geotagging of Objects from Street View Imagery

This paper describe about solution to localize the object from multiple view using geometry.To geolocate the object in image they developed Markov Field model to perform object triangulation.They use 2 SOTA FCNN for semantic segmentation and monocular depth estimation.The geolocalization is done with Google street view images with Triangular based MRF model described in paper. End Result of the projects looks similar to the output of

the resulting output but lacks the precision which is done with the help of depth image and Triangulation-based MRF model

Chapter 3
Methodology

here the method and the

Chapter 4
Conclusion

ghghjkghkgkh [8]

Chapter 5

Bibliography

Bibliography

[1] Stanford Vision Lab, "Large Scale Visual Recognition Challenge (ILSVRC)," 2015.

[2] Developers.google.com, "Developer Guide," 2018.

[3] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2017-Octob, pp. 2980–2988, 2017.

[4] Python Software Foundation, "Python," 2019.

[5] Colab.research.google.com, "Welcome to Colaboratory!."

[6] S. F. Conservancy, "git –fast-version-control."

[7] J. Kang, M. Körner, Y. Wang, H. Taubenböck, and X. X. Zhu, "Building instance classification using street view images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 44–59, 2018.

[8] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 1800–1807, 2017.