Detecting, Mapping, and Grading Sidewalks using Street View Images and Secondary Sources for the city of Dallas

Andrew Abbott1, Alex Deshowitz1, Dennis Murray1, Dr. Eric Larson1

1 Southern Methodist University, Dallas, Texas

**Abstract.** In this paper, we explore the feasibility of using machine learning methods to find and map sidewalks and to then assess their quality and accessibility. We seek to create a training set of images to be used to train a convolutional neural network to correctly identify and grade Street View images.

1 Introduction

Documentation and mapping of a neighborhood feature, such as sidewalks, has previously been carried out through in-person (subjective) “Neighborhood Audits” or through information cataloged in Geographic Information Systems. The completeness, timeliness, and quality of the information gathered in these methods may be below what is needed for a comprehensive view of the availability of a feature like sidewalks across a city. There is a myriad of reasons for the city of Dallas to focus a portion of spending on sidewalk creation and repair. For the mobility impaired, availability of sidewalks is an essential requirement for moving about the metropolitan area. Additionally, sidewalks allow citizens to move in a safe manner around the city without the risks associated with walking in the streets amongst cars driven by ever-increasingly distracted drivers. Sidewalks also provide health benefits since they provide a means of pedestrian travel to near-home destinations for those who may otherwise choose to not travel. Lastly, sidewalks in disrepair tend to contribute to the “broken-windows” theory surrounding many under-privileged areas of the metroplex. Overall, sidewalks represent a worthwhile investment for the city and provide a means of safe and healthy travel for those who choose pedestrian transportation.

This project seeks to provide a means of grading and prioritization of sidewalk projects to the city of Dallas via an unbiased, algorithmic approach to sidewalk repair recommendation and scoring.

2 Problem Statement

Many municipalities struggle with how to correctly allocate funding for necessities such as pedestrian sidewalks. Oftentimes, it is hard to prioritize those items that are in the most need for repair. The city of Dallas is one such municipality that needs this sort of help. In a metropolitan area, as large and diverse as Dallas, many times projects are prioritized by those items garnering the highest number of complaints or those areas with the most political influence. This project seeks to “level the playing field” in order ultimately allow the city to create an optimal budget for sidewalk repairs.

This project first seeks to create a grading system for sidewalks based on Street View imagery. This grading system will consider features such as “obstacles”, condition of the sidewalk, and presence of sidewalk transitions in order to provide a grade for a set of sidewalk training data. Once an appropriate training set has been created, the goal is to train a convolutional neural network to correctly classify ungraded sidewalks in order to recommend and prioritize sidewalk repairs. After the initial training phase is complete, the model may be modified to take into account other open-source data structures such as satellite imagery and personal geo-location data in order to help the model better generalize to the test data.

3 Project Plan

Domain Knowledge Research

Understand Social and Public Health considerations of sidewalks

Pedestrian Safety concerns

Persons with Physical Difficulties

Other Approaches to the Problem – Neighborhood Audits

Python Image Processing Knowledge

Google/Bing Street View API Knowledge

Machine Learning/Neural Networks with Images

Best Practices for improving performance in image classification problems

GIS (Geographic Information System) file formats and interchanges

Development of Machine Learning/Neural Network, Cross Validation, Testing

Documentation

Final Poster Presentation

4 Ethical Considerations

With data being sourced primarily through “Street View” photographs collected by major search engines, the coverage of images must be evaluated to ensure that there is equitable coverage of neighborhoods of all income and demographic varieties.

5 Data Quality Considerations

Use of images sourced via the Street View collection method represents several potential shortfalls. Aside from the previously addressed potential for under-represented neighborhoods, paths away from the street (through parks or alleys) may also not be represented. Age of the images also represents a challenge, as sidewalks may have been built after the image, or sidewalk quality may have changed since the image was created.

6 Github

The address for this project’s Github page is: [github.com/dpmurraygt/CapstoneProject](file:///C:\SMU%20Data%20science\capstone\AppData\Local\Temp\Temp1_CapstoneProject-master.zip\CapstoneProject-master\github.com\dpmurraygt\CapstoneProject)

7 Project Timeline

