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. This project will focus on the city of Dallas; however, if successful, this approach will provide meaningful results to all municipalities.

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 lower than 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.

In the fiscal year 2015, Dallas had a proposed budget for infrastructure projects, such as sidewalks, of $7,135,208. That number has grown each of the past several years as the city has increased in size. The city of Dallas has grown in both its inhabited sprawl and in the density of the population at double digit rates over the past 5 years. This growth can be explained by the city’s pro-business mentality and the relative value proposition that such a large city provides. Dallas boasts a centrally located position in the country and businesses can send employees to any destination in the world with relative ease. The City and State also have pro-business tax and incentive policies, which have encouraged this growth. With this growth, the city must acquire a better way of allocating the budget for sidewalks and other infrastructure. Today, the city of Dallas essentially responds to complaints about sidewalks through its street services program. As inquiries come in, the city will generally put that particular inquiry on a list for assessment. There is no priority granted for severity of the situation. Additionally, maintenance in suburban areas is the responsibility of the home or property owner and this is not necessarily considered in the ranking process. The current process takes 2-3 months in order to get an assessment and cost estimate for each incident or property owner. Once the assessment has been done, either the city or the property owner will plan the project and pay for the repairs. One nice program in Dallas allows the city to reimburse home owners up to $500 or 50% of the repair cost, whichever is less. Under these circumstances, the city spends a great deal of time and resources just doing the assessments. Using the algorithmic approach that this paper lays out, the city would be able to feed images of these incidents into the model and immediately receive a grading of the sidewalk in question. The city officials could then compare this grading to the grading of previous works to know whether this sidewalk was an immediate issue, who owns the sidewalk, and how much the repair may cost. This would allow the city to prioritize project desires into bins such as: critical, severe, moderate, and low-risk. Therefore, Dallas could dispatch crews to the areas where their services will be the most impactful to the safety and health of the public.

This project provides 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 to 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 Previous Research

Image recognition is not a new field. The use of machines to recognize images has been around for decades. As early as 1963, the electrical engineering department at MIT began using computers to recognize 3D images. While these initial applications were somewhat simple compared to those that we currently use today, they paved the way for what has now become a commonplace practice across industries.

One project which has provided meaningful advancements for this project is Bahlman, Zhu, and Pelkofer’s work (3). In this paper, the authors built upon their previous research involving shape and color recognition to help classify street signs and street lights. Their work is interesting in that it has a 2-step approach where if the model fails on the first classification step, the image is thrown out. This model is important to our work, because it shows how an algorithm such as Adaboost can be used to detect both anomalous and important features for an image-based problem proposal.

Another important piece of research is Perona’s “A Bayesian Hierarchical Model for Learning Natural Scene Categories” (4). In this paper, the authors provided a very meaningful approach that allowed for very hands-off model building. This model will potentially provide a structure for our model should we encounter any issues with sparse image objects that are hard to classify correctly. In the model, the computer attempts to use human-based rules to classify image objects. Essentially, each image is broken down into a series of codebook images and reoccurring elements are scanned and classified. Each of these codebook images is additionally clustered using k-means clustering. This portion of the model is used to eliminate features that occur with low frequencies in the training data. The remarkable aspect of this paper is that the model was able to achieve a 78% accuracy rate with such a low amount of supervision. While there are many applications of image classification models, the models outlined above provide a solid basis for our understanding of the evolution of image recognition and model application. Our next area of concentration has been on the specific use of neural networks for problem solutions in the image recognition and classification space.

An aspect of previous work that is of high importance for this work is the use and application of convolutional neural networks. As this is our model of choice, it was important for us to research the application of convolutional neural networks and their potential pitfalls. One such application of neural networks for image recognition was demonstrated by Goodfellow and a team from Google (5). In this work, Goodfellow applied the DistBelief method for neural networks combined with Google Streetview images to recognize multi-digit numbers, namely street addresses. In the model, the researchers first addressed training the model to identify house numbers. This was a very important step as many variables come into play with these image captures. For instance, lighting, obstructions, and changing conditions can provide potential issues when identifying numbers from images. Additionally, varying font sizes, colors, and styles can impact the ability of the algorithms to correctly identify an image. An important aspect of this type of recognition is that if a single digit is misidentified, the entire interpretation is irrelevant and meaningless. Once the model was trained on house numbers, a more complete Streetview dataset was used. The final approach involved subtracting the mean from each image. In the end, the researcher’s models were able to achieve a 97.84% accuracy with this approach, which was just short of the human benchmark of 98% for which the researchers were shooting. This piece of research and the approach acted as an important catalyst for our approach to identifying sidewalk obstructions and sidewalk grading.

Convolutional neural networks have also been used to improve the solutions submitted in the ImageNet Large-Scale Visual Recognition Challenge. In the work of Simonyan, Karen, and Zisserman (6), the team used convolutional neural networks combined with several other approaches to achieve one of the highest accuracy levels seen in the competition. Their application of multiple models to solve the problem provides a solid reference point for the problem that we solve in this model.

Another important piece of research was Clarke et al’s (13) audit of Streetview images as compared with an individual’s in-person audit. The study involved researchers in neighborhoods in Chicago walking each block from the inside to the outside, essentially walking the block twice, and assessing the quality of the sidewalks. This study found that subjective measures, such as sidewalk quality, has much lower consistency between Streetview and the observed condition. Essentially, the conclusion is that features requiring high levels of precision can be hard to attain via Streetview images. This poses an interesting aspect to our research. For instance, our model will need to stay informed of areas that have been treated in the previous time periods. Therefore, if a database of previously updated sidewalks does not exist, we will need to provide a means of storing projects within a database that allows for those items or coordinates to be referenced. This will prevent outdated Streetview images from being used in the classification and scoring process.

While it is important to see that our research problem can be solved via machine learning techniques, it is also important to see that there is indeed a reason for the application of these techniques to solve the task at hand. Therefore, it is important to see that improving sidewalk quality, coupled with other factors can lead to better health for society overall. In Haina et al (11), the researchers looked at signal data such as walkability of neighborhoods in relation to the overall health of the individuals in the area. The evidence used to provide insight into the improved environment of an area was sanitation practices and tobacco sales restriction. This coupled with increased walkability of an area leads to higher levels of physical activity and better health over time.

The fourth area of research for this project focused on the general health benefits of neighborhood walkability. Deehr and Shumann (7) provided work for five different neighborhoods in the Seattle area. Their research considered the incidence of pedestrian strikes by motorists, the health factors of walking, and the current modes of transportation that pedestrians were using. Their research led to the city adding additional walking paths, and trails. Additionally, much of the research sparked additional community involvement in the design of multi-model transportation infrastructure. Additionally, in Richardson,Troxel et al (10), the authors sought to understand whether factors such as green space and walkability resulted in “moderate to vigorous physical activity” for the residents of randomly selected neighborhoods in Pittsburg, PA. When controlling for factors such as crime, green space, and walkability in the selected targets, it was discovered that variables such as gender, age, education, and overall walkability of the neighborhood did play significant roles in the levels of physical activity for an area. This research helps us reaffirm that there is immense potential for identifying areas that need this sort of infrastructure. Ultimately the goals of helping people lead healthier and safer lives are potential outcomes of the modeling exercise laid out in this paper.

Overall, our research helped us layout the precedent for image recognition, understand the application of the specific type of model that we are attempting to build, evaluate the effectiveness of proper infrastructure, and provide statistical affirmation of the health and societal benefits of proper infrastructure. With this knowledge established, we can make the case for our model to be used for the stated application problem.

4 Algorithm Design and Solution

For this model, we will be applying a convolutional neural network to train and ultimately score a series of Google Streetview images of the Dallas metropolitan area. This project will help the City of Dallas better prioritize infrastructure projects with potential extensions beyond applications to sidewalks. For this model, we have considered several inputs: the training data, the testing data, and the scoring algorithm used to correctly classify images.

The training data used in this model came from the University of Maryland’s Project Sidewalk. The University has agreed to share the data collected through this project with the SMU Walknet team. Project Sidewalk is part of a crowd-sourced approach to classifying images for potential issues in the image. This tool can be thought of as an application of the same methodology used by the Wayz application for smartphones, with the only difference being that users are not physically in the environment that they are assessing, rather they assess the sidewalk from their home computers. Users are essentially dropped into a Google Streetview environment and then told to identify and grade a potential feature in the given image. The grading scale falls from a best value of “passable” or 1 to a worst value of “not passable” or 5. Of course, this grading schema is based off of subjective, visual inspection and not actual user experience feedback. Understandably, this does introduce the potential for some bias in the training data. However, it is believed that the data cleaning methods in place from the University of Maryland are sufficient for our model. Additionally, our model is more concerned with the obstacle identification in the actual image rather than the grade given. Our hope is that we can incorporate grading into the actual algorithm based on some sort of dispersion measure amongst the pixels in the image itself. The data has been shared via a Box repository on the cloud where a series of panoramic images have been dropped for our consumption into the algorithm. **UPDATE** – we are currently working on a piece of code that is able to ingest these images and understand which pixel values constitute obstacles or sidewalk features that we are interested in for this study.

The test data that we are using will be an area of focus on the South side of the City of Dallas. This is an area where a great deal of funding and regentrification has been focused. In order to ensure that the neighborhoods are getting the correct amount of funding, we will be focusing in on this area of the city to start. We will use the Google Streeview API in Python in order to download the proper number of images to test for the area. We have currently not decided on image spacing as we are still working on the process for loading the images into the environment appropriately. Once we have the data downloaded into the appropriate repository we will be able to test the data collected.

The algorithm that we are using is a convolutional neural network. Convolutional neural networks have been applied to this sort of problem on many occasions and are basically the go-to approach to solving this sort of problem. These algorithms are especially strong when the training data is of the level of quality that we have from the University of Maryland. While the convolutional neural network approach will be the approach that we start with, we may take an approach using different deep learning approaches such as Deep Believe Networks or stacked denoising autoencoders. However, we are still some time off with this approach for our project. We are still working on the best approach for ingesting images into memory in an efficient manner.

5 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

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)

References

1. Kotaro Hara, Shiri Azenkot, Megan Campbell, Cynthia L. Bennett, Vicki Le, Sean Pannella, Robert Moore, Kelly Minckler, Rochelle H. Ng, and Jon E. Froehlich. 2015. Improving public transit accessibility for blindriders by crowdsourcing bus stop landmark locations with Google Street View: An extended analysis. ACM Trans. Access. Comput. 6, 2, Article 5 (March 2015), 23 pages.
2. Kotaro Hara, Jin Sun, Robert Moore, David Jacobs, Jon E. Froelich. 2014. Tohme: Detecting Curb Ramps in Google Stret View Using Crowdsourcing, Computer Vision, and Machine Learning. UIST ’14, October 5-8 2014. Honolulu, HI.
3. Claus Bahlmann, Ying Zhu, Visvanathan Ramesh, Martin Pelkofer, Thorsten Koehler. “A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion.” IEEE Xplore, July 2005.
4. Fei-Fei, Li, and Pietro Perona. "A bayesian hierarchical model for learning natural scene categories." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 2. IEEE, 2005.
5. Goodfellow, Ian J., et al. "Multi-digit number recognition from street view imagery using deep convolutional neural networks." *arXiv preprint arXiv:1312.6082* (2013).
6. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
7. Deehr, Rebecca C., and Amy Shumann. "Active Seattle: achieving walkability in diverse neighborhoods." *American journal of preventive medicine* 37.6 (2009): S403-S411.
8. Knopp, Jan, Josef Sivic, and Tomas Pajdla. "Avoiding confusing features in place recognition." *Computer Vision–ECCV 2010* (2010): 748-761.
9. Le, Quoc V. "Building high-level features using large scale unsupervised learning." *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013.
10. Richardson, Andrea S., et al. "One size doesn’t fit all: cross-sectional associations between neighborhood walkability, crime and physical activity depends on age and sex of residents." *BMC public health* 17.1 (2017): 97.
11. Hajna, Samantha, et al. "Neighborhood walkability: field validation of geographic information system measures." *American journal of preventive medicine* 44.6 (2013): e55-e59.
12. Hosler, Akiko S., et al. "Relationship between objectively measured walkability and exercise walking among adults with diabetes." *Journal of environmental and public health* 2014 (2014).
13. Clarke, Philippa, et al. "Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument." *Health & place* 16.6 (2010): 1224-1229.
14. Kim, Hyungjin et al. “A Probabilistic Feature Map-Based Localization System Using a Monocular Camera.” Sensors 2015
15. Kelly, Cheryl et al. “The Assoication of Sidewalk Walkability and Physical Disorder with Area-Level Race and Poverty.”
16. Sheppard, Stephen R.J., Cizek, Petr. “The Ethics of Google Earth: From Spatial Data to Landscape Visualization.”