RampNet

Curb Ramp Detection at Scale

JOHN S. O'MEARA
Issaquah High School
University of Washington

JARED HWANG
PhD Student
University of Washington

ZEYU WANG
PhD Student
University of Washington

MICHAEL SAUGSTAD
Research Scientist
University of Washington

JON E. FROEHLICH
Professor, Computer Science
University of Washington

DESIGN USE BUILD

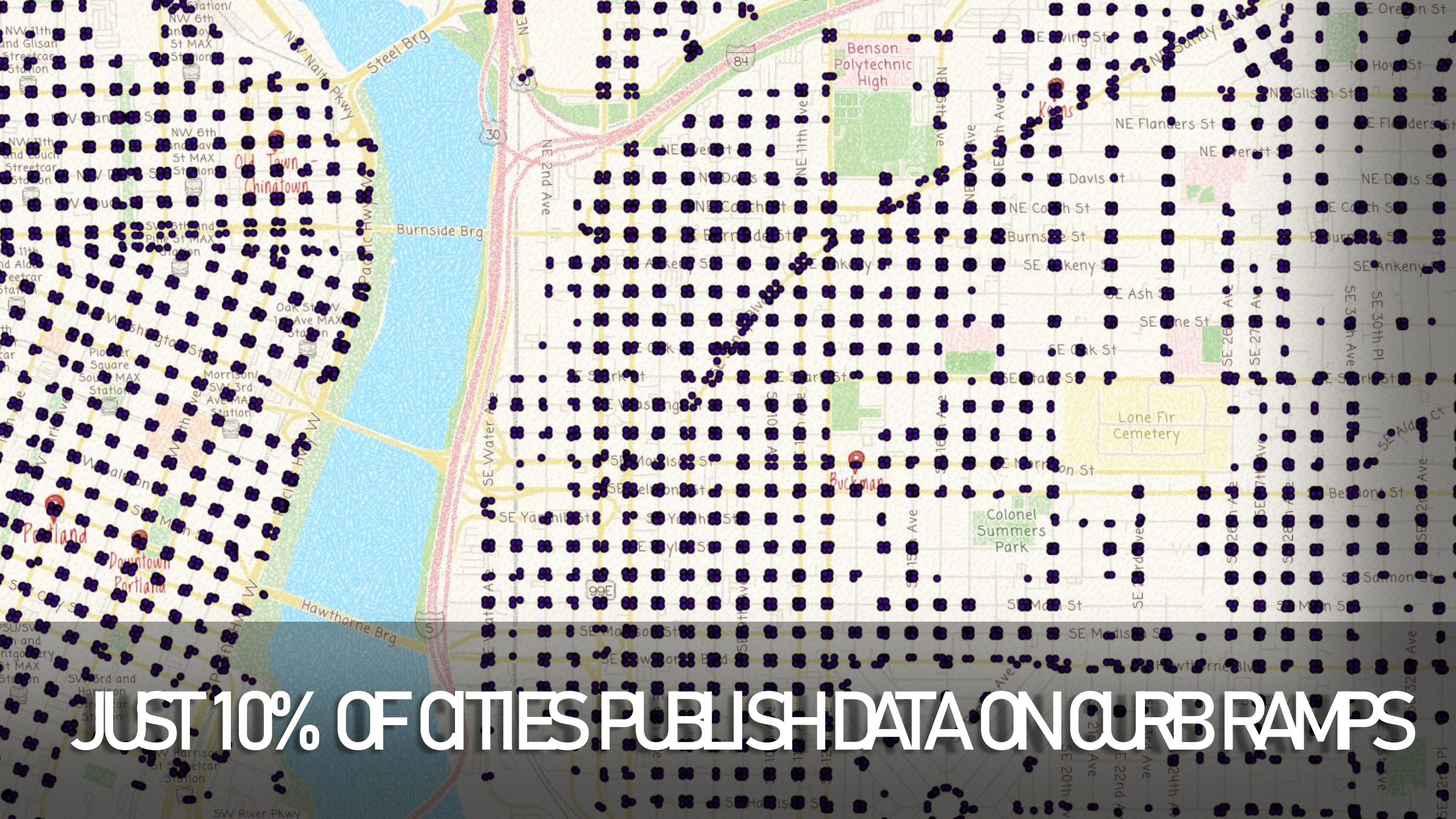
MAKEABILITY LAB

PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

UNIVERSITY of WASHINGTON


CURB RAMPS ARE
CRITICAL FOR URBAN
ACCESSIBILITY

CURB RAMP DATA IS CRUCIAL



JUST 10% OF CITIES PUBLISH DATA ON CURB RAMPS

Boots-On-The-Ground Collection

Past Computer Vision Approaches

Boots-On-The-Ground Collection

Slow.

Costly.

Boots-On-The-Ground Collection

Past Computer Vision Approaches

PROJECT SIDEWALK

Boots-On-The-Ground Collection

Past Computer Vision Approaches

Past Computer Vision Approaches

Scaling Crowd+AI Sidewalk Accessibility Experiments Examining Label Quality and on Performance

Michael Duan
Allen School of Computer Science, University of Washington
Seattle, USA

Shosuke Kiami
Allen School of Computer Science, University of Washington
Seattle, USA

Johnson Kuang
Allen School of Computer Science, University of Washington
Seattle, USA

Michael Saugstad
Allen School of Computer Science, University of Washington
Seattle, USA

Jon E. Froehlich
Allen School of Computer Science, University of Washington
Seattle, USA

ABSTRACT

Increasingly, crowds plus machine learning techniques are being used to semi-automatically analyze the accessibility of built environments; however, open questions remain about how to effectively combine the two. We present two experiments examining the effect of crowdsourced data in automatically classifying sidewalk accessibility features in streetscape images. In Experiment 1, we investigate the effect of validated data—which has been voted correct by the crowd but is more expensive to collect—compared with a larger but noisier aggregate dataset. In Experiment 2, we examine whether crowdsourced labeled data gathered in one city can be used as effective training data for another. Together, these experiments contribute to the growing literature in Crowd+AI approaches for semi-automatic sidewalk assessment and help identify pertinent challenges.

ACM Reference Format:
Michael Duan, Shosuke Kiami, Logan Milandin, Johnson Kuang, Michael Saugstad, Maryam Hosseini, and Jon E. Froehlich. 2022. Scaling Crowd+AI Sidewalk Accessibility Assessments: Initial Experiments Examining Label Quality and Cross-city Training on Performance. In *The 24th International ... Conference on Computers and Accessibility (ASSETS '22), ... NY, USA, 5 pages.*

Deep Learning for Automatically Detecting Sidewalk Accessibility Problems Using Streetscape Imagery

Galen Weld[1], Esther Jang[1], Anthony Li[2], Aileen Zeng[1], Kurtis Heimerl[1], and Jon E. Froehlich[1]
[1]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, USA
[2]Department of Computer Science, University of Maryland, College Park, USA
{gweld, infrared, aileenz, kheimerl, jonf}@cs.washington.edu, antli@umd.edu

ABSTRACT

Recent work has applied machine learning methods to automatically find and/or assess pedestrian infrastructure in online map imagery (*e.g.,* satellite photos, streetscape panoramas). While promising, these methods have been limited by two interrelated issues: small training sets and the choice of machine learning model. In this paper, aided by the recently released Project Sidewalk dataset of 300,000+ image-based sidewalk accessibility labels, we present the first examination of deep learning to automatically assess sidewalks in Google Street View (GSV) panoramas. Specifically, we investigate two application areas: automatically *validating* crowdsourced labels and automatically *labeling* sidewalk accessibility issues. For both tasks, we introduce and use a residual neural network (ResNet) modified to support both image and non-image (contextual) features (*e.g.,* geography). We present an analysis of performance, the effect of our non-image features and training set size, and cross-city generalizability. Our results significantly improve on prior automated methods and, in some cases, meet or exceed human labeling performance.

Author Keywords
Neural networks, accessibility, sidewalks, computer vision

ACM Classification Keywords

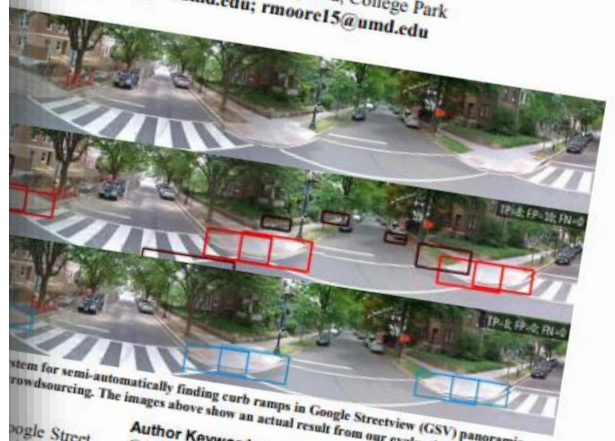
inspections by city transit departments or community volunteers. However, these audits are expensive, labor intensive, and infrequent. [1] Moreover, the resulting data is in disparate formats, is not typically open (*i.e.,* published online), and is not intended for end-user tools [23, 50]. To expand who can collect sidewalk data and to improve data granularity and freshness, researchers have introduced smartphone-based tools [15, 46, 52] as well as instrumented wheelchairs [35, 39, 51, 57], both of which capture sidewalk information *in situ* as it's experienced. However, these tools have been limited by low adoption, small geographic coverage, and high user burden (*e.g.,* requiring users to take out their phones, load an app, take a picture, annotate it, and upload it) [20, 23].

To partially address these scalability issues, researchers have begun developing automated methods for sidewalk assessment using machine learning and online imagery (*e.g.,* satellite photos [10, 8], panoramic streetscape imagery [31, 32, 59]). While still early, these complementary approaches promise to dramatically decrease manual labor and cost. However, they have been limited by two interrelated issues: small training sets and the choice in machine learning model—both of which negatively impact performance. In this paper, we attempt to address both of these issues.

We present the first examination of deep learning methods to

Tohme: Detecting Curb Ramps in Google Street View Using Crowdsourcing, Computer Vision, and Machine Learning

...in Sun, Robert Moore[1,2], David Jacobs, Jon E. Froehlich[1,2]
...bility Lab | [2]Human Computer Interaction Lab (HCIL)
...cience Department, University of Maryland, College Park
...insun, dwj, jonf}@cs.umd.edu; rmoore15@umd.edu



...stem for semi-automatically finding curb ramps in Google Streetview (GSV) panoramic ...rowdsourcing. The images above show an actual result from our evaluation.

...ogle Street ...ly collect ...present the ... machine ...n crowd ...y scenes. ...labeling ...n, which ...rmance. ...m four ...workers, ...g curb ...ne (F- ...on in ... curb ...based

Author Keywords
Crowdsourcing accessibility; computer vision; Google Street View; Amazon Mechanical Turk

ACM Classification Keywords
H.5.m Information interfaces and presentation; I.4.m Image processing and computer vision; K.4.2 Social Issues

INTRODUCTION

Recent work has examined how to leverage massive online map datasets such as Google Street View (GSV) along with crowdsourcing to collect information about the accessibility of the built environment [21–25]. Early results have been promising; for example, using a manually curated set of static GSV images, Hara *et al.* [23] found that minimally trained crowd workers in Amazon Mechanical ... (turkers) could find four types of ... problems with 81% ...

Past Computer Vision Approaches

Good in theory but...

Lackluster accuracy/performance

Impractical for fully-automatic detection

Key Problem: **Data**

Can we translate curb ramp **location** data into **image** data?

Curb Ramp Locations

Latitude	Longitude
40.840397	-73.81662
40.840287	-73.81657
...	...

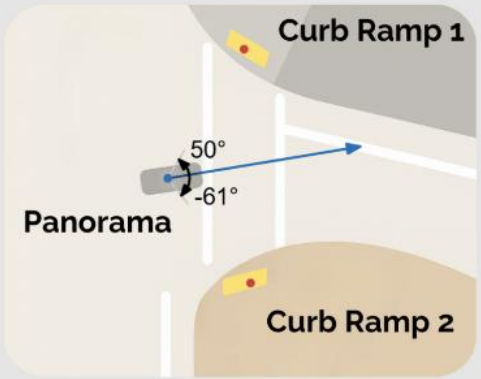
Input

Google Street View Panorama

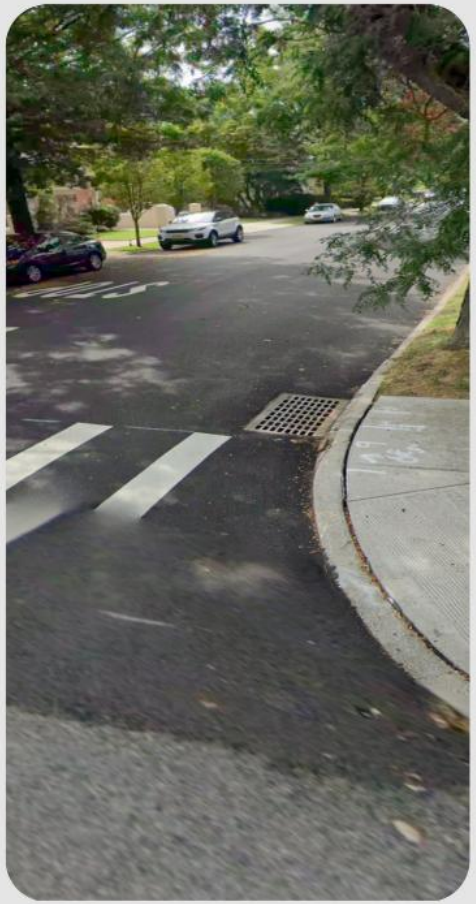
Curb Ramp Locations

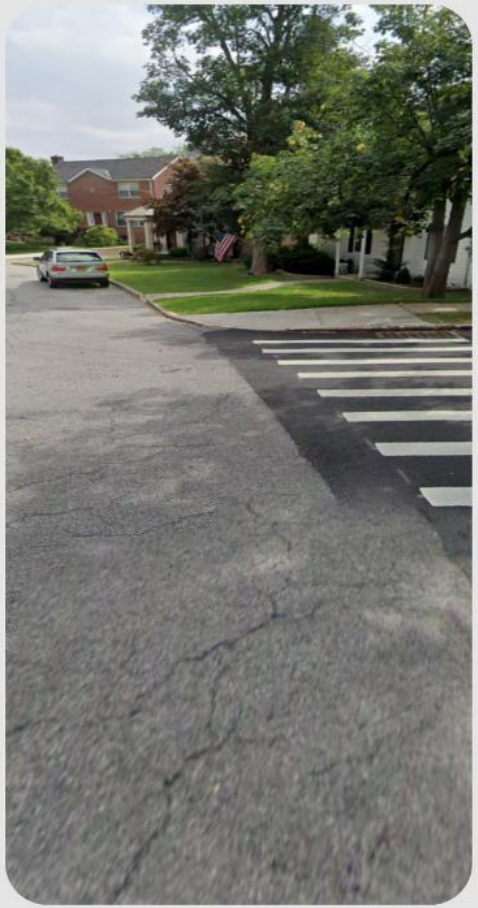
Latitude	Longitude
40.840397	-73.81662
40.840287	-73.81657
...	...

Angle Calculation

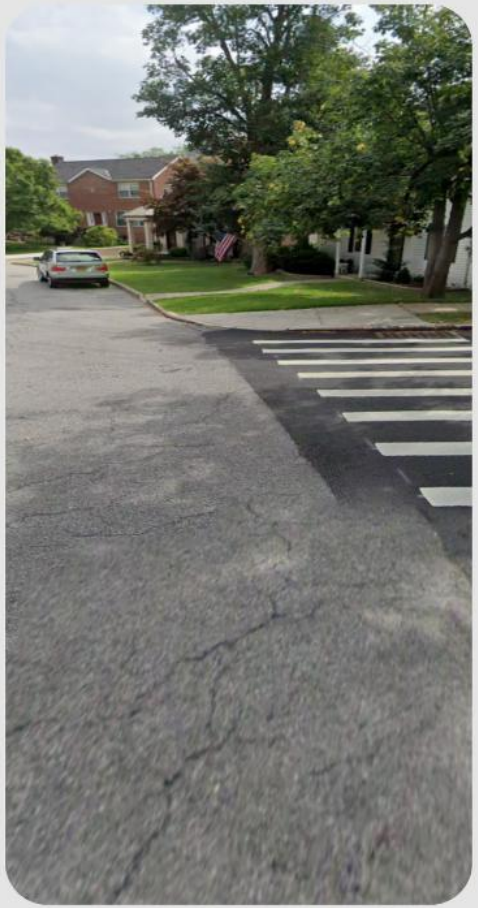


Perspective Crop Extraction





pective Crop Extraction



Auto-Translation Model

Auto-Labeled Crops

beled Crops

Fully-Labeled Panorama

Output

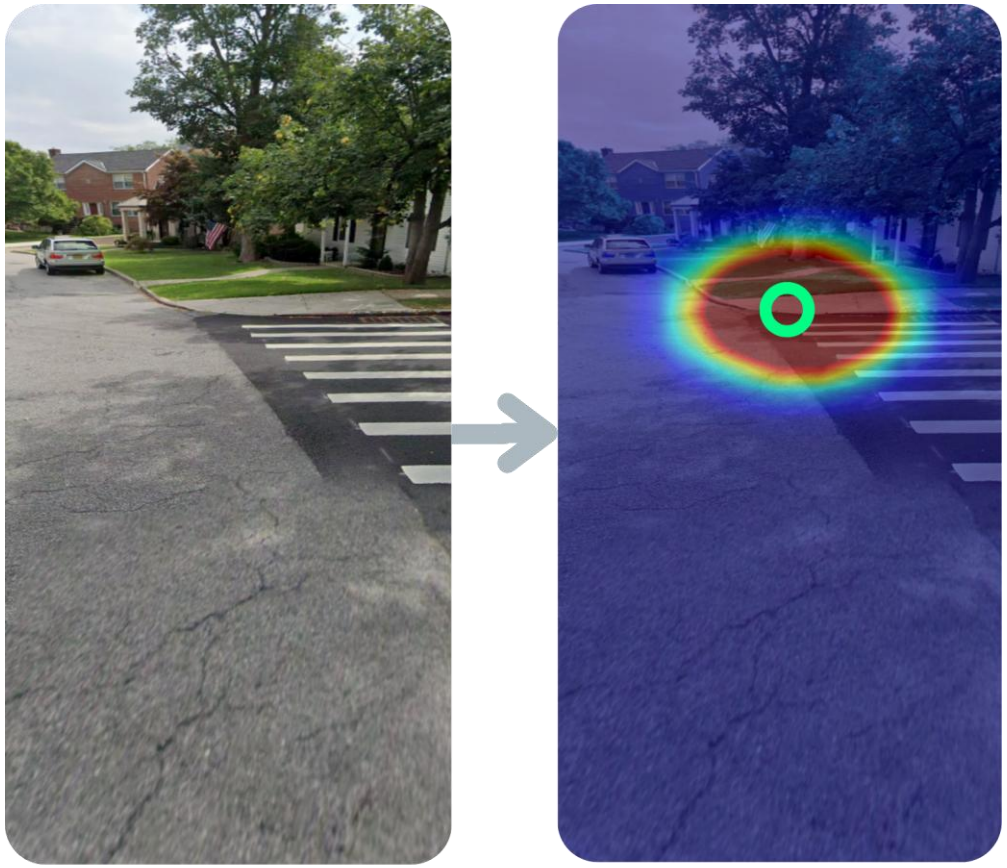
Auto-Translation Model

ConvNeXt V2

Finds the pixel position of a curb ramp in a small, directional crop.

Pretrained on 20,698 Project Sidewalk panoramas.

Second round of training on smaller, manually-labeled dataset of 312 panoramas.



Input Output

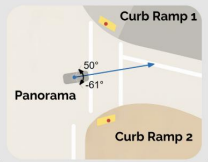
Input

Google Street View Panorama

Curb Ramp Locations

Latitude	Longitude
40.840397	-73.81662
40.840287	-73.81657
...	...

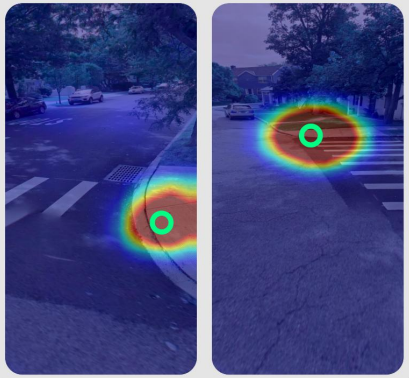
Angle Calculation



Perspective Crop Extraction

Auto-Translation Model

Auto-Labeled Crops



Fully-Labeled Panorama

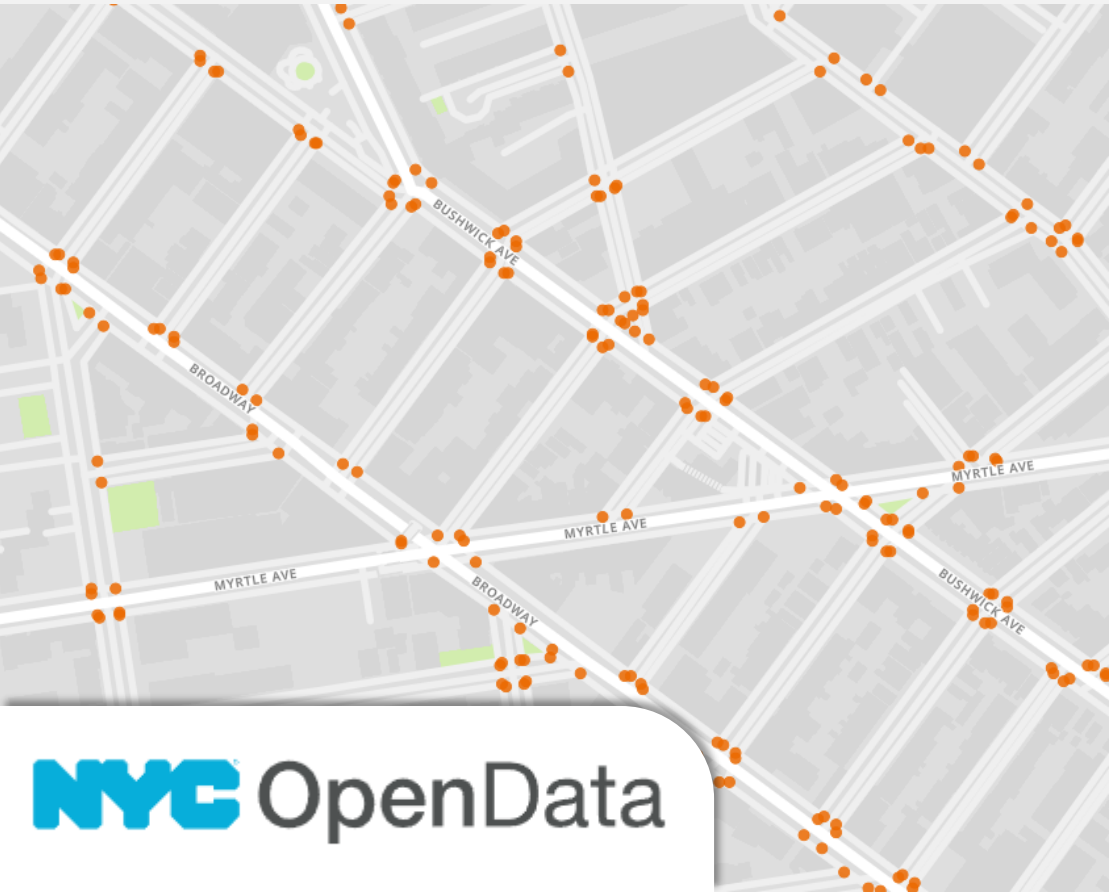
Output

City-Provided Location Data

New York City

Portland, OR

Bend, OR



Generated Image Dataset

Generated Image Dataset

214,388 panoramas

849,904 labels

Key Problem: **Data**

Stage Two: Curb Ramp Detection

Model Inference

Unseen Panoramas

Curb Ramp Detection Model

ConvNeXt V2

Model Predictions

0.924 AP

Curb Ramp Detection Model

ConvNeXt V2

Trained on the dataset we generated in Stage 1.

Input: 4096 × 2048px panorama

Output: 1024 × 512px heatmap

We evaluated against a **manually labeled** dataset of 1,000 panoramas.

	Weld *et al.* [45]	RampNet Stage 2
Input Type	Depth+Image+Geo Data	Image Data
Dataset Used	Project Sidewalk [35]	Auto-Generated
Avg. Precision	0.3803	0.9236

TRY IT!

RampNet in Action

https://tinyurl.com/ps-vancouver

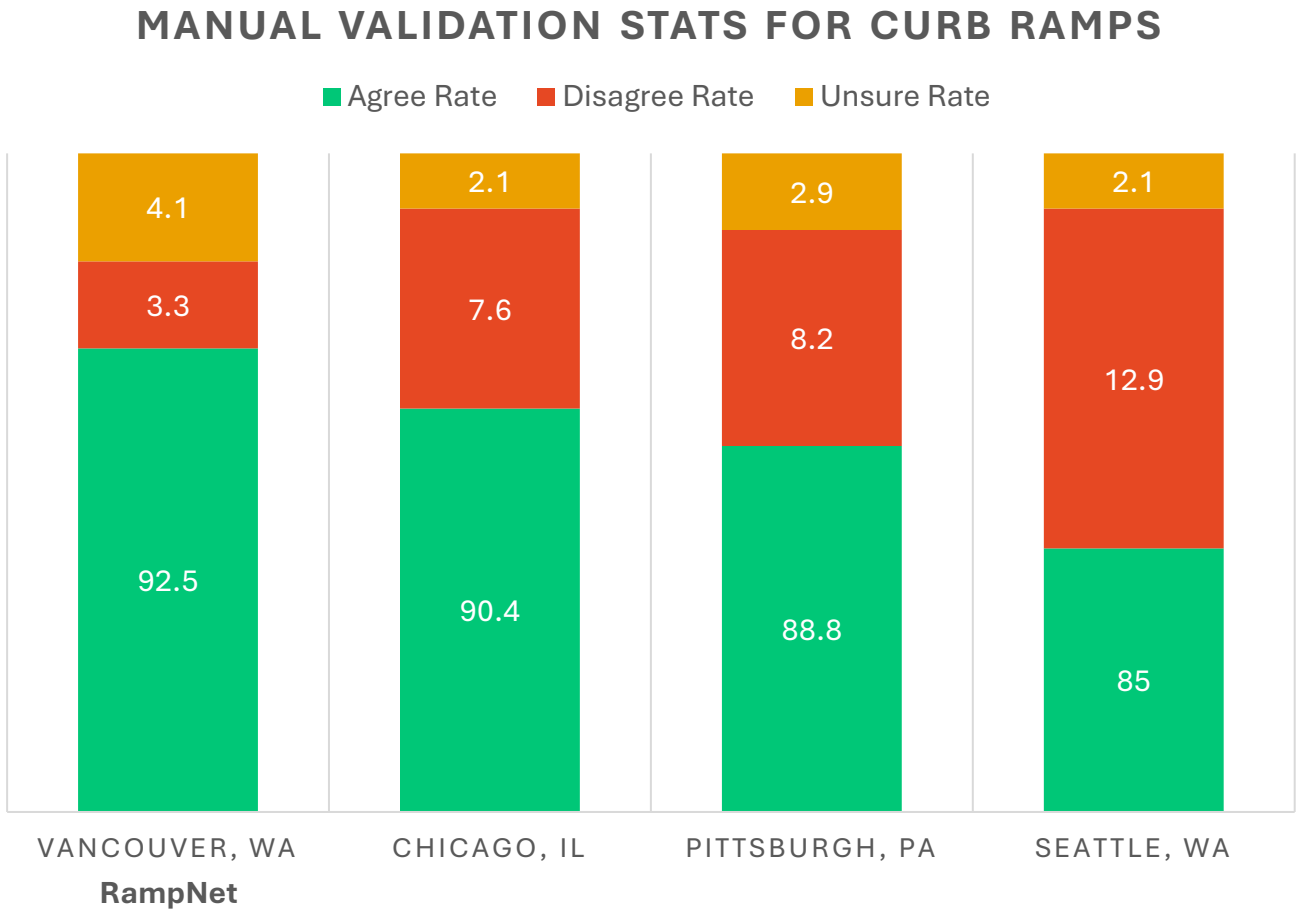
PROJECT SIDEWALK

TRY IT!

RampNet in Action

https://tinyurl.com/ps-vancouver

PROJECT SIDEWALK



TRY IT!

RampNet in Action

https://tinyurl.com/ps-vancouver

PROJECT SIDEWALK

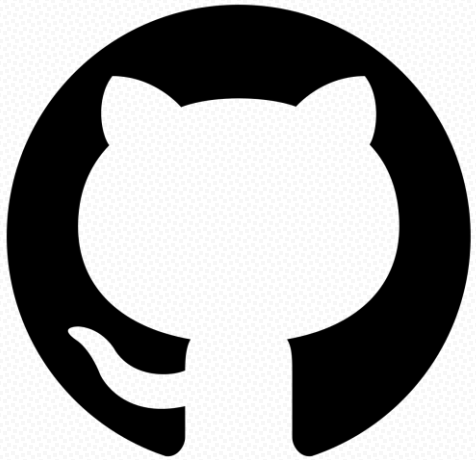
- Less than 48 hours
- Less than $20
- More accurate than crowdsourced labels
- Near infinite ability to scale

New translation technique

Largest curb ramp image dataset

SOTA detection model

MAP, ASSESS, & VISUALIZE EVERY SIDEWALK IN THE WORLD



https://github.com/ProjectSidewalk/RampNet

This work was supported by NSF SCCIRG #2125087 and NSF OAC #2411222.

RampNet

Curb Ramp Detection at Scale

JOHN S. O'MEARA
Issaquah High School
University of Washington

JARED HWANG
PhD Student
University of Washington

ZEYU WANG
PhD Student
University of Washington

MICHAEL SAUGSTAD
Research Scientist
University of Washington

JON E. FROEHLICH
Professor, Computer Science
University of Washington

DESIGN USE BUILD

MAKEABILITY LAB

PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

UNIVERSITY of WASHINGTON