Experimental Crowd+AI Approaches to Track Accessibility Features in Sidewalk Intersections Over Time

Ather Sharif
Allen School of Computer Science, University of Washington
asharif@cs.uw.edu

Paari Gopal
Allen School of Computer Science, University of Washington
paarig@cs.uw.edu

Michael Saugstad
Allen School of Computer Science, University of Washington
saugstad@cs.uw.edu

Shiven Bhatt
Allen School of Computer Science, University of Washington
shivbat@cs.uw.edu

Raymond fok
Allen School of Computer Science, University of Washington
rayfok@cs.uw.edu

Galen Weld
Allen School of Computer Science, University of Washington
gweld@cs.uw.edu

Kavi Dey
Seattle Academy High School

Jon E. Froehlich
Allen School of Computer Science, University of Washington
jonf@cs.uw.edu

ABSTRACT

How do sidewalks change over time? Are there geographic or socioeconomic patterns to this change? These questions are important but difficult to address with current GIS tools and techniques. In this demo paper, we introduce three preliminary crowd+AI (Artificial Intelligence) prototypes to track changes in street intersection accessibility over time—specifically, curb ramps—and report on results from a pilot usability study.

CCS CONCEPTS

• **Human-centered computing**; • **Accessibility**; • **Accessibility systems and tools**;

KEYWORDS

Mobility, disability, sidewalks, crowdsourcing, machine learning, change tracking

ACM Reference Format:

Ather Sharif, Paari Gopal, Michael Saugstad, Shiven Bhatt, Raymond fok, Galen Weld, Kavi Dey, and Jon E. Froehlich. 2021. Experimental Crowd+AI Approaches to Track Accessibility Features in Sidewalk Intersections Over Time. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21), October 18–22, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3441852.3476549

1 INTRODUCTION

In 1990, the US passed the *Americans with Disabilities Act* (ADA) [1] requiring that public infrastructure—including sidewalks and street crossings—be accessible. Yet, more than 30 years later, cities struggle to meet accessibility requirements, often only pursuing large-scale sidewalk renovations in response to civil rights litigation such as in New York [12], Seattle [9], and Los Angeles [18]. Observing these challenges and to help stimulate and structure ADA renovations and city planning, in 2015, the *US Federal Highway Administration* requested that local governments develop sidewalk ADA transition plans, including an inventory of accessibility barriers and a description of accessible renovations [25]. In a recent study of 401 municipalities, however, only 54 (13%) had published plans and only seven met the minimum ADA criteria [7].

Such findings reflect the challenge in making infrastructure accessible. Viable solutions require substantial political, economic, and technical investment—training, resources, community involvement, specialized tools, and the work and coordination of multiple governmental agencies [19]. And there is a lack of open tools, techniques, and datasets to track how urban infrastructure is becoming more or less accessible. To help understand *how* sidewalks are changing, *where* resources are being invested, and *whether* governments are acting on ADA requirements, our research group is developing new spatio-temporal tracking tools to analyze, visualize, and study changes in urban accessibility over time. While our current focus is on the US, tracking sidewalk accessibility is of interest to cities across the world [8]. With our tools, we hope to support overarching research questions such as: How does sidewalk infrastructure change over time? What are the spatiotemporal patterns of change? How do these changes correspond to socioeconomic and demographic factors? Where does inaccessibility persist? As a preliminary step towards addressing these questions, we introduce three new experimental crowd+AI (artificial intelligence) prototypes for semi-automatically tracking changes in street intersections, specifically curb ramps (or "curb cuts")—Figures 1-3. While curb ramps are only one part of accessible urban infrastructure, they are critical to mobility and required by the ADA [29]. Moreover, prior work has found that trained computer vision (CV) models can detect curb ramps at higher accuracy than surface degradations or sidewalk obstacles [11, 27], making curb ramps a good starting place for initial crowd+AI work.

ASSETS '21, October 18–22, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8306-6/21/10.
https://doi.org/10.1145/3441852.3476549

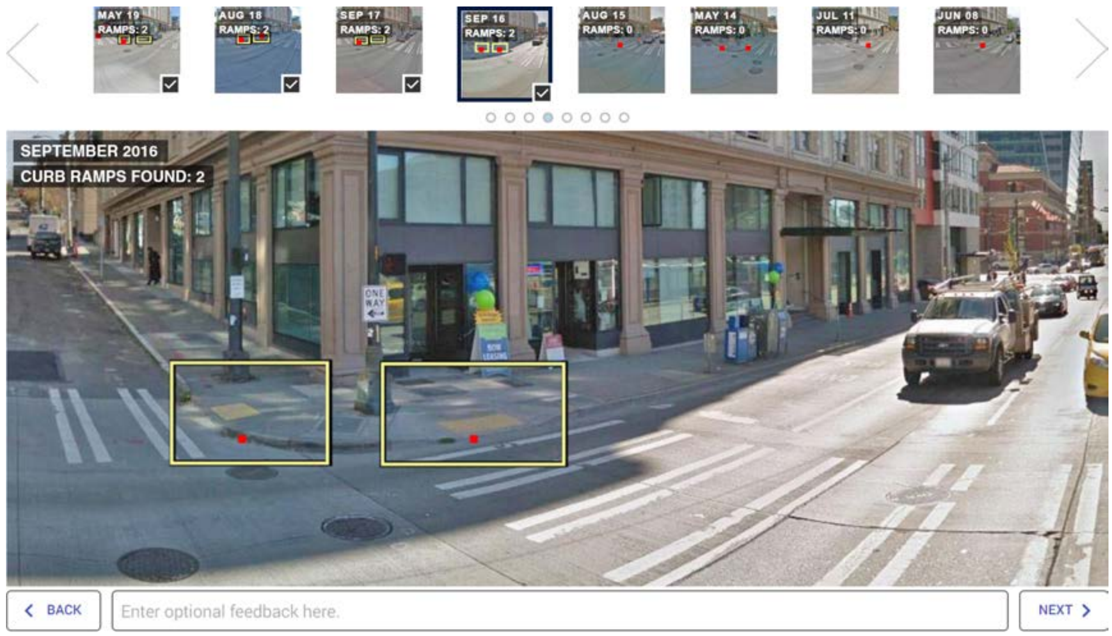

Figure 1: With P1: Single View, users label time-series images of individual street corners (in this case, from May 2019 to June 2008). The thumbnail menu shows available time-series images at the selected corner, which update in real-time as the user draws bounding-box labels with their mouse. Checkboxes indicate a completed (labeled) time snapshot. Automatically detected ramps are indicated with small red squares, which we can be turned on/off. We plan on conducting experiments to examine the potential benefits of these automatic detections, particularly given that they are not always accurate.

Studying and characterizing spatiotemporal patterns of urban change from remote imagery is a longstanding thread in the urban- and geo-sciences [13, 23, 28]. Recent developments in CV, specifically deep learning, and the widespread availability of historic street-level imagery have enabled new urban change detection techniques [3, 5, 15, 21, 26]. However, limited work exists on applying these techniques to urban accessibility to characterize *how* and *where* sidewalks are changing. Below, we describe design considerations for tracking accessibility-related changes in street intersections, three preliminary crowd+AI prototypes, and results from a pilot usability study with five users. During the ASSETS demo session, we will provide interactive demonstrations of our prototypes, solicit feedback from attendees, and guide discussion about open challenges and the future of sidewalk "change tracking" tools.

2 UI DESIGN CONSIDERATIONS FOR TRACKING CHANGES IN SIDEWALKS

In brainstorming and working on initial prototypes, we developed the following design considerations:

Humans struggle with change detection. Studies in perceptual psychology have consistently found that humans perform poorly in identifying differences between images [17, 22]. How can we create tools that help humans identify and label accessibility features in time-series imagery while mitigating these perceptual effects [22]?

Leverage temporal similarity. Unlike general street scene labeling tasks [6, 16], we are interested not just in identifying objects in a single snapshot but tracking these objects over time. How can we leverage structural similarities in time-series photography to create efficient and accurate labeling interfaces?

Combine AI + human labeling. Similarly, how can humans + machine learning work together to maximize labeling efficiency and accuracy [4, 14]? How should AI-based detections and uncertainty be represented to humans? Can the underlying ML model also leverage similarities across time-series images?

Interactive training. Ultimately, to scale our approach, we will deploy our interfaces to crowdworkers who likely have minimal experience with sidewalks, curb ramps, and advanced labeling interfaces. How can we develop interactive training UIs that allows our users to quickly learn and perform accurately in our tasks?

3 THREE CROWD+AI INTERFACES FOR TRACKING CURB RAMPS OVER TIME

Towards these considerations, we have developed three early-stage interactive prototypes for tracking changes in street intersections over time, which differ in the amount of simultaneous time-series imagery shown, how labels propagate from one time-series snapshot to the next (using a derivation of linked editing [24]), and how we incorporate a deep learning model for automatic curb ramp detection (from [27]). Rather than ask users to detect *changes*, users find and label curb ramps in each image. To improve labeling efficiency, we leverage similarities across time shots to auto-propagate

Figure 2: With P2: Grid View, thumbnails are larger and presented in a grid allowing us to show up to 9 time-series images simultaneously. Unlike P1, P2 uses linked editing [24] to leverage structural similarities across time. When the user draws or edits a bounding box on the most recent time snapshot (always shown as the top-left thumbnail, which in this case is May 2019), these annotations are auto-propagated to all previous time shots using x,y pixel location similarity. If CV detections are turned on (red squares), we attempt to auto-align propagated boxes based on inferred curb ramp locations; however, these alignments are not always accurate due to noise in the ML model (e.g., there are two incorrect CV detections on the July 2011 time shot above). The user can make micro-edits or deletions, as necessary, on the propagations.

Figure 3: Similar to P2, P3: Panorama View also includes linked editing [24] and auto-propagation of labels across time; however, unlike P1 and P2, which serve segmented intersections cropped into four individual images (one corner per image), P3 presents full panorama views. Benefits of panoramas include greater context for the user and potentially faster overall labeling; however, the curb ramps themselves are small and only ~3-4 time-series panoramas can fit on a laptop screen, so users need to scroll to access older images. To help users more closely examine panorama parts, we have an always-available zoom inset of the mouse location (shown currently at the May 2019 panorama above). In this particular example, the intersection was renovated between June 2008 (bottom pano) and July 2011 with ADA compliant ramps and three ramp additions. These changes are identified with our techniques.

labels through linked editing and CV. Each prototype begins with a step-by-step tutorial to train users on the task and the interface. Prototype details with example screenshots are shown in Figures 1-3.

For our historic street scene dataset, we use Google Street View's "time machine" feature, which provides high-resolution street-level panoramas dating back to 2007 captured every ~1-3 years. Our test set consists of 100 intersections drawn from Washington DC and

Seattle (50 each). The DC dataset contains an average of 6.4 time-series images per location (*SD*=1.7) while Seattle has 7.8 (*SD*=2.6). The first capture dates are 2008 and the last are 2019 (while our research is ongoing, this initial test dataset was created in 2019).

4 PILOT USABILITY STUDY

To assess the usability and understandability of our prototypes and to prepare for larger web-based deployments, we conducted an in-person "think aloud" usability study with five participants (ages 20-45; all had technical backgrounds). Sessions were ~50 minutes. To simulate the experience of using the prototypes in an online deployment, we provided limited instruction and, instead, asked participants to follow the interactive tutorials.

Findings. While users were appreciative of the step-by-step tutorials, some aspects of label propagation, and the promise of CV-assisted labeling, we found important areas for future work. First, participants wanted more information on how they should *label*—the size of their bounding boxes, pixel-level precision, *etc.* Second, participants were confused about *label propagations*—should they trust them or modify them? Because auto-propagations only work in one direction (labeling is propagated backwards but *not* forwards through time) and because only some operations are supported (additions but not deletions), users did not have a strong understanding or confidence in this feature. Finally, though the automatic CV detections (visualized as red squares) were deemed helpful in drawing attention to curb ramps, participants felt that it was too often incorrect and thus distracting (though one participant enjoyed "outperforming" the AI).

5 CONCLUSION AND FUTURE WORK

In this demo paper, we introduced three novel crowd+AI tools aimed at rapidly labeling and tracking changes in sidewalk accessibility features over time. In addition to addressing results from our usability study, we aim to support richer qualitative labels about how curb ramps are changing (*e.g.,* tactile strips, flares, steepness) and other accessibility-related labels for crosswalks [2], accessible pedestrian signals [10], and street/sidewalk surfaces [20]. We also plan to conduct a larger-scale deployment study to further assess our tools and progress towards public deployment, like Project Sidewalk [20], for tracking changes in urban accessibility infrastructure across cities and creating open "change tracking" datasets.

ACKNOWLEDGMENTS

This work was partially funded by the USDOT Pacific Northwest University Transportation Center (PacTrans) and the UW Center for Research and Education on Accessible Technology and Experiences (CREATE).

REFERENCES

[1] United States Department of Justice Civil Rights Division Ada.gov. Introduction to the ADA.
[2] Dragan Ahmetovic, Roberto Manduchi, James M. Coughlan, and Sergio Mascetti. 2015. Zebra Crossing Spotter: Automatic Population of Spatial Databases for Increased Safety of Blind Travelers. In The 17th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2015).
[3] Pablo F. Alcantarilla, Simon Stent, Germán Ros, Roberto Arroyo, and Riccardo Gherardi. 2018. Street-view change detection with deconvolutional networks. Autonomous Robots 42, 7: 1301–1322. https://doi.org/10.1007/s10514-018-9734-5
[4] Saleema Amershi, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, Eric Horvitz, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, and Paul N Bennett. 2019. Guidelines for Human-AI Interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19, 1–13. https://doi.org/10.1145/3290605.3300233
[5] Kuan-Ting Chen, Fu-En Wang, Juan-Ting Lin, Fu-Hsiang Chan, and Min Sun. 2016. The world is changing: Finding changes on the street. In Asian Conference on Computer Vision, 420–435.
[6] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[7] Yochai Eisenberg, Amy Heider, Rob Gould, and Robin Jones. 2020. Are communities in the United States planning for pedestrians with disabilities? Findings from a systematic evaluation of local government barrier removal plans. Cities 102: 102720. https://doi.org/10.1016/j.cities.2020.102720
[8] Jon E Froehlich, Mikey Saugstad, Edgar Martínez, and Rebeca de Buen Kalman. 2020. Sidewalk Accessibility in the US and Mexico: Policies, Tools, and A Preliminary Case Study. In CSCW2020 Workshop on Civic Technologies: Research, Practice, and Open Challenges.
[9] David Gutman. 2017. Seattle may have to spend millions making sidewalks more accessible to people with disabilities. The Seattle Times.
[10] Richard Guy and Khai Truong. 2012. CrossingGuard: exploring information content in navigation aids for visually impaired pedestrians. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12) (CHI '12), 405–414. https://doi.org/10.1145/2207676.2207733
[11] Kotaro Hara, Jin Sun, Robert Moore, David Jacobs, and Jon Froehlich. 2014. Tohme: detecting curb ramps in google street view using crowdsourcing, computer vision, and machine learning. In Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14, 189–204. https://doi.org/10.1145/2642918.2647403
[12] Winnie Hu. 2017. For the Disabled, New York's Sidewalks Are an Obstacle Course. The New York Times.
[13] Wei Ji, Jia Ma, Rima Wahab Twibell, and Karen Underhill. 2006. Characterizing urban sprawl using multi-stage remote sensing images and landscape metrics. Computers, Environment and Urban Systems 30, 6: 861–879. https://doi.org/10.1016/J.COMPENVURBSYS.2005.09.002
[14] Vikram Mohanty, David Thames, Sneha Mehta, and Kurt Luther. 2019. Photo Sleuth: Combining Human Expertise and Face Recognition to Identify Historical Portraits. In Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19), 547–557. https://doi.org/10.1145/3301275.3302301
[15] Ladan Najafizadeh and Jon E. Froehlich. 2018. A Feasibility Study of Using Google Street View and Computer Vision to Track the Evolution of Urban Accessibility. In Poster Proceedings of ASSETS'18.
[16] Gerhard Neuhold, Tobias Ollmann, Samuel Rota Bulo, and Peter Kontschieder. 2017. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).
[17] Ronald A. Rensink. 2002. Change Detection. Annual Review of Psychology 53, 1: 245–277. https://doi.org/10.1146/annurev.psych.53.100901.135125
[18] Emily Alpert Reyes. 2015. L.A. agrees to spend $1.3 billion to fix sidewalks in ADA case. Los Angeles Times.
[19] Manaswi Saha, Devanshi Chauhan, Siddhant Patil, Rachel Kangas, Jeffrey Heer, and Jon E. Froehlich. 2020. Urban Accessibility as a Socio-Political Problem: A Multi-Stakeholder Analysis. In Proceedings of the ACM on Human Computer Interaction (PACM HCI); To Be Presented at CSCW2020.
[20] Manaswi Saha, Michael Saugstad, Hanuma Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, and Jon E. Froehlich. 2019. Project Sidewalk: A Web-based Crowdsourcing Tool for Collecting Sidewalk Accessibility Data at Scale. In Proceedings of CHI 2019.
[21] Ken Sakurada, Daiki Tetsuka, and Takayuki Okatani. 2017. Temporal city modeling using street level imagery. Computer Vision and Image Understanding 157: 55–71. https://doi.org/10.1016/j.cviu.2017.01.012
[22] Daniel J. Simons and Michael S. Ambinder. 2005. Change Blindness. Current Directions in Psychological Science 14, 1: 44–48. https://doi.org/10.1111/j.0963-7214.2005.00332.x
[23] Xiao-Peng Song, Joseph O. Sexton, Chengquan Huang, Saurabh Channan, and John R. Townshend. 2016. Characterizing the magnitude, timing and duration of urban growth from time series of Landsat-based estimates of impervious cover. Remote Sensing of Environment 175: 1–13. https://doi.org/10.1016/J.RSE.2015.12.027
[24] M Toomim, A Begel, and S L Graham. 2004. Managing Duplicated Code with Linked Editing. In 2004 IEEE Symposium on Visual Languages - Human Centric Computing, 173–180. https://doi.org/10.1109/VLHCC.2004.35
[25] US Department of Transportation Federal Highway Administration. 2015. ADA Transition Plan.
[26] Jan D. Wegner, Steve Branson, David Hall, Konrad Schindler, and Pietro Perona. 2016. Cataloging Public Objects Using Aerial and Street-Level Images — Urban Trees. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6014–6023. https://doi.org/10.1109/CVPR.2016.647

[27] Galen Weld, Esther Jang, Anthony Li, Aileen Zeng, Kurtis Heimerl, and Jon E Froehlich. 2019. Deep Learning for Automatically Detecting Sidewalk Accessibility Problems Using Streetscape Imagery. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '19), 196–209. https://doi.org/10.1145/3308561.3353798

[28] 2Jianguo Wu, G. Darrel Jenerette, Alexander Buyantuyev, and Charles L. Redman. 2011. Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States. *Ecological Complexity* 8, 1: 1–8. https://doi.org/10.1016/J.ECOCOM.2010.03.002

[29] 1993. *Kinney v. Yerusalim, 9 F.3d 1067 (3d Cir. 1993).* Retrieved from https://casetext.com/case/kinney-v-yerusalim