Contents

royalsocietypublishing.org/journal/rsta Phil. Trans. R. Soc. A 0000000

PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Research

Article submitted to journal

Subject Areas:
Computer vision, Image processing, Human-computer interaction

Keywords:
Accessibility, Participatory, Sidewalks, Visual AI, Mapping, Smartphone

Author for correspondence:
Diego Morra
e-mail: diego.morra@polimi.it
Fabio Duarte
e-mail: fduarte@mit.edu

Mapping sidewalks accessibility with smartphone imagery and visual AI: A participatory approach

Diego Morra[1,2], Xiaosheng Zhu[2,3], Chang Liu[2], Kyle Fu[2], Fábio Duarte[2], Simone Mora[2,4], Zhengbing He[2] and Carlo Ratti[2,5]

[1]Politecnico di Milano, Department of Electronics, Information and Bioengineering (DEIB), Milan, Italy
[2]Massachusetts Institute of Technology, Senseable City Lab, Cambridge, MA, USA
[3]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[4]Norwegian University of Science and Technology, Department of Computer Science, Trondheim, Norway
[5]Politecnico di Milano, ABC Department, Milan, Italy

Evaluating sidewalks' accessibility represents a manual and time-consuming task that requires specialized personnel. While recent developments in visual AI have paved the way for automating data analysis, the lack of sidewalk accessibility datasets remains a significant challenge. This study presents the design and validation of Sidewalk AI scanner, a web app that enables quick, crowd-sourced, and low-cost sidewalk mapping. The app enables a participatory approach to data collection through imagery captured using smartphone cameras. Subsequently, dedicated algorithms automatically identify sidewalk features such as width, obstacles, or pavement conditions. Though not a replacement for high-resolution sensing methods, this method leverages data crowdsourcing as a strategy to produce a highly-scalable, city-level dataset of sidewalk accessibility, offering a novel perspective on the city's inclusivity; fostering community empowerment and participatory planning.

THE ROYAL SOCIETY PUBLISHING

1. Introduction

Sidewalks are a critical component of urban infrastructure, providing a protected space for pedestrian movement and access to urban facilities. While some features such as minimal width or pavement material may go unnoticed (1), they can have significant impact on people with disabilities (2), affecting their emancipation (3), quality of life (4), and physical activity (5). With the increasing recognition of the importance of accessible urban infrastructure (6; 7; 8; 9), there is a growing interest in evaluating and enhancing sidewalk accessibility to meet the diverse needs of all city residents, including people with permanent or temporary impairments, such as people with disabilities, the elderly, and those with young children. In the United States, the Americans with Disabilities Act (ADA) establishes a comprehensive framework to guide the development and evaluation of accessible urban infrastructure (10). The ADA Accessibility Guidelines detail key features for the design and construction of safe sidewalks for physically impaired individuals. Still, most cities lack sidewalk accessibility inventories based on standardized evaluation metrics (2; 11). Traditionally, assessing sidewalk accessibility mainly relied on manual inspections (12), which are time-consuming, labor-intensive, and prone to human error (13; 14). Manual inspections also struggle to scale across larger urban areas, making it difficult to obtain a comprehensive view of accessibility issues.

Recent advancements in technology (such as LiDAR scanners and UAVs) are paving the way for more efficient and accurate data collection methods, enabling large-scale analysis with reduced human intervention. However, substantial drawbacks exist, such as high costs for data acquisition, which includes significant post-processing labor. These limitations contribute to the current lack of open and broadly available sidewalk datasets, being one of the main challenges for accessibility mapping in the cities (2; 7; 13). Participative methods that leverage crowdsourcing accessibility data have been investigated, engaging citizens in mapping their sidewalks showing promising results (15). Nevertheless, these methods require users to manually report sidewalk features through time-consuming tasks that impact the data re-sampling frequency. Furthermore, these manual approaches can also impact data reliability and quality (16).

This paper proposes Sidewalk AI scanner, a hybrid-crowdsourced method combining a participatory approach with Visual AI to face sidewalk accessibility data collection challenges. The paper presents the rationale, architecture, and validation of this method, which is based on a web app that: i) enables anyone to record and upload sidewalk-view videos registered through any smartphone; ii) employs Visual AI trained models to automatically identify features that can impact sidewalk accessibility such as width, pavement conditions or obstacles; iii) generate reports and maps that showcase the sidewalk accessibility level for each mapped city.

The paper contributes to the current sidewalk data collection state-of-the-art by:

- Proposing a framework that facilitates a participative collection of sidewalk data by enabling anyone to collect it automatically through a smartphone.
- Comparing different data collection technologies' time and cost efficiency to identify a low-cost and scalable solution for crowsourcing sidewalk accessibility mapping.
- Validating a Visual AI based model to detect sidewalk width from a smartphone-generated video.

The paper is organized as follows. Section 2 presents a literature review on current urban accessibility challenges and sidewalk data collection methods. Section 3 introduces the Sidewalk AI Scanner, and Section 4 validates the process by detailing models and prototypical tool adopted and their accuracy in inheriting an essential feature for sidewalk accessibility. Section 5 discusses the accuracy, limitations, and potential impact on urban accessibility and citizen's involvement. Finally, Section 6 summarizes the main contributions and outlines the next steps.

2. Literature review

The availability and accessibility of pedestrian pathways are critical to moving people away from motorized transportation. This is fundamental for people with permanent, temporary, or situational impairments, such as people with disabilities, the elderly, pregnant women, or guardians with strollers. Without accessible sidewalks, these people cannot access urban facilities, transport, and green areas—effectively; they are removed from full access to the city (2). Promoting inclusive cities has gained attention from governments, researchers, and practitioners (7; 9; 13; 17; 18). This is in line with the United Nations Sustainable Development Goals (SDGs) (19) "Sustainable Cities and Communities" (Goal 11), which emphasizes how inclusive urban planning can ensure that public spaces, buildings, and services are designed to accommodate diverse needs, fostering social inclusion and participation. In a similar vein, countries have enacted legislation and governmental regulations providing guidelines on the development of urban environments that respect the needs of people with disabilities, ensuring accessibility and inclusivity, as well as access to resources (10; 20; 21; 22) Furthermore, each city faces unique local challenges reflective of culture, policy, and resources that must be considered. For example, a recent study attempted to define a method to quantify the accessibility level of sidewalks in Brazil, considering different elements, such as the pavement condition, obstacles, slopes, or the actual walkable dimension (23). Another study in Mexico reports the preliminary use of a crowdsourced tool for collecting data on sidewalks (24). The authors also argue for new low-cost and scalable sidewalk tracking tools that support evidence-based advocacy and policymaking.

Nevertheless, tools designed to put the data collection results into the hands of people with disabilities are still lacking. This includes assistive location-based technologies that incorporate accessibility features into navigating, searching, and exploring the physical world. Accessibility scores, disability-sensitive navigation maps, and digital tools built on accurate data for visualizing and navigating urban environments can impact the ability of people with disabilities to evaluate and make informed and autonomous decisions about their movements (25). This is often impossible as the need for reliable and updated data represents the biggest feasibility challenge.

(a) Sidewalk data collection and mapping

(i) Mobile Remote Sensing and tree-dimensional Scene Reconstruction

Most recent methods are based on three-dimensional remote sensing technology like LiDAR to collect high-geometry georeferenced data. Ai, C., & Tsai, Y. (26) report on an experimental test conducted on the Georgia Institute of Technology campus in Atlanta. They showed accurate measurement results for the key features of the sidewalk and curb ramps from video log images and LiDAR point cloud. The proposed method automatically extracts key features regulated by the ADA. Similarly, Q., Hou, & C., Ai (27) proposes a network-level sidewalk inventory method using LiDAR data, deep neural network, and a stripe-based sidewalk extraction algorithm. The result of a case study shows that the proposed method can generate accurate and efficient means for network-level sidewalk inventory.

While the initial results reported by these two case studies and the recent diffusion of cheaper and portable LiDAR sensors are promising, the adoption of this technology is still inaccessible to most non-expert users, due to high costs and knowledge barriers.

(ii) Manual mapping and crowdsourced data generation

Numerous projects investigate qualitative approaches based on the involvement of stakeholders in the data collection and analysis process. Starting from 2013, Frackelton et al. (28), anticipating that the sidewalk assessment system would gain widespread national application, proposed a volunteer-based data collection system using an app that runs on a tablet attached to the base of a wheelchair. The video and vibration data are then processed to identify the sidewalk sections needing repair or reconstruction. Most recently, Biagi et al. (29) proposes a platform that enables residents to to report barriers directly to the governmental bodies and aspires to determine the best approach for mapping accessible routes and identifying obstacles in OpenStreetMap. Similarly, Mobasheri et al. (30) proposes a revised approach utilizing data mining methods to develop sidewalk geometries in OpenStreetMap using multiple GPS traces collected by wheelchair users.

Other projects focus instead on creating ad-hoc tool to enable the manual mapping of sidewalks features and barriers. Project Sidewalk (15) is a web-based tool that allows users to label accessibility problems by virtually walking through city streets in Google Street View. In a more recent publication (16), authors report that an 18-month deployment study resulted in 797 online users contributing 205,385 labels and auditing 2,941 miles of Washington DC streets, with remote users that managed to label 92% of accessibility problems. Based on the same concept, Maps for Easy Paths (MEP) (31) is an Android app that lets the user manually report obstacles and track the route while traveling, with the idea of mapping only accessible paths. In spite of the potential of the crowdsourced approach for data labeling, this could also represent a drawback, as resulted dataset can present reliability and quality issues due to human error (15; 16).

(iii) Semantic segmentation with satellite and street view image

Semantic segmentation categorizes each pixel of an image into a specific semantic label (32). This computer vision approach is among the most promising as it could rely on a large number of street image datasets, representing a cheaper and more accessible alternative to solutions based on technology like LiDAR.

Some works have investigated the application of computer vision models to satellite or aerial images to automatically map the features of sidewalks. Senlet & Elgammal (33) propose one of the first frameworks to construct sidewalk and crosswalk maps from satellite images. The model also addresses the challenge connected to sidewalks in satellite images occluded by trees. Recently, Hosseini et al. (16) proposed TILE2NET, an open-source scene classification model for pedestrian infrastructure from sub-meter resolution aerial tiles to generate pedestrian networks. While satellite mapping is effective for pathway network generation, the resulting data are not fine-grained enough for qualitative mapping. This aspect is critical in the urban accessibility context as static obstacles or other types of barriers reduce the actual width of a sidewalk (34).

Several studies investigated qualitative mapping possibilities by applying semantic segmentation on street view images (SVI). In addition to serving as a convenient source for extracting the features or traits of roads (35), SVI could be employed to map the individual featurs that make up a sidewalk. Furthermore, SVI is also freely available, making it possible to map sidewalks in large urban environments. Different services offer global images, such Google Street View (GSV), OpenStreetMap, Mapillary, and KartaView (36).

The multifaceted nature of walkability results in the exploration through semantic segmentation of various physical aspects, such as the distribution and network of sidewalks (37; 38), quality of sidewalks (34), and accessibility of walkways for the disabled individuals (39). For instance, (40) combined data of sidewalk presence and condition from a GSV application into a designed sidewalk walkability variable by the authors to show four possible levels: no, poor, fair, and good sidewalk conditions. Zhou et al. (41) utilized deep learning technologies to segment features from Baidu Map Street View. Although they achieve good results, few of these segment obstacles, such as holes, stairs, and damaged areas or extract dimensional features such as width.

(iv) Hybrid-crowsourced approach

Despite a considerable body of research proposing effective solutions to address the scarcity of sidewalk data, two main challenges continue to hinder the automation of sidewalk inventories at a large scale: (i) applying Visual AI models to datasets from satellite or SVI suffer from insufficient temporal resolution. This limitation arises because the frequency at which these image datasets are updated varies significantly, ranging from weekly to monthly intervals, depending on the urban area. This variability impacts the reliability and accuracy of the generated sidewalk inventories. Also, SVIs imaginery is mainly road-centric, so sidewalks could appear covered by parked cars, walls, or buildings. (ii) collecting sidewalks data on a city level requires fast and low-cost solutions to be sustainable on a long-term. For example, LiDAR-based approaches are promising but require investment and dedicated personnel to cover the full city sidewalks at least weekly. Not adequate periodic coverage could generate the same temporal resolution issues as SVI-based approach. Considering these two open challenges, none of the previous solutions seems sustainable in a long-term setting, where sidewalk data should be accurate and constantly updated, as any feature change could impact accessibility for a diverse category of people.

An hybrid-crowdsourced approach could address these challenges. This approach combines technological solutions, like computer vision, with crowdsourcing techniques. Only a few projects have investigated this approach, reporting positive results in the localization of zebra crosswalks in large image datasets (38) or in detecting curb ramps in GSV scenes, combining computer vision and custom user interfaces (42; 43). More recently, Weld et al. (44) propose a model for auto-validating and auto-labeling sidewalks in streetscape imagery previously manually labeled by Project Sidewalk participants (15). Authors of these hybrid-crowdsourced approaches report labeling performance comparable, or even more promising, than human performance, with a notable reduction in time cost. While these results address the challenges related to mapping sidewalks at a city level in a fast and scalable manner, they still need to solve the temporal resolution, as they are based on SVI datasets.

The method proposed in this paper aims to address both of the challenges, combining (1) an easy, fast, and broadly available tool for crow-sourcing sidewalk data collection (2) a visual AI method that automatically extracts all sidewalk features that could impact a sidewalk accessibility level. Together, these two aspects could ensure the granularity and precision of data collection and significantly impact the approach's scalability. This scalability allows the method to be extended to any city where citizens and stakeholders are willing to engage in a participatory data collection.

3. Sidewalk AI Scanner: Methodology

Sidewalk AI Scanner is a web app that proposes a participatory approach to sidewalk data collection empowered by a data-gathering technology that leverages the capabilities available in every smartphone. By offering a participatory approach in the form of a scalable data collection mechanism instead of a manual labeling task and by leveraging the high availability of smartphones in conjunction with visual AI algorithms able to automatize accessibility feature recognition, the Sidewalk AI Scanner builds upon previous studies of accessibility standards to craft a novel, hybrid-crowdsourced solution to address the critical gap in sidewalk mapping.

(a) Framework

The framework encompasses a web app that operates both as a *data collection tool* and *data visualization tool* (see figure 1). The data collection tool provides the user with all necessary functionalities and guidelines to correctly capture and upload a sidewalk video (as detailed in Section (d)). Subsequently, the visual AI model is applied to frames extracted from the user-uploaded video to automatically label width, pavement type, obstacles, and other sidewalk elements that impact accessibility (refer to Table 1), and generate georeferenced data from segmentation results, which are then made available through a dedicates interactive maps.

This framework required two steps. The first was to determine a list of main sidewalk features that should be identified to evaluate the accessibility level. The second step was to compare different data collection technologies, determining which allows for the least time-consuming and most cost-effective data collection process from citizens' point of view. These steps are presented and discussed in detail in the following sections.

Table 1. Elements that impact the accessibility of a sidewalk outlined from literature (22; 23; 25; 46)

Element name	Description
Width	The measure of space available for pedestrian traffic.
Length	The stretch of the sidewalk in terms of distance.
Slope	The incline level of the sidewalk.
Pavement type	The material and texture of the sidewalk surface.
Surface problem	Any issues with the surface that might hinder movement.
Obstructed path	Any blockage or barrier that reduce sidewalk width.
Curb Ramps	Sloped transitions between the sidewalk and the street.
Light poles	Presence of adequate illumination source.
Crosswalks	Presence and pedestrian crossing areas on the road.
Visual Contrast	Color and texture contrast for better visibility.
Visual information	Signs or visual clues that support orientation.
Non-visual information	Tactile or auditory cues for the visually impaired.
Overcrowding	The amount of pedestrian traffic.
Temporary obstacles	Non-permanent elements that reduce sidewalk width.

(b) Sidewalk accessibility features

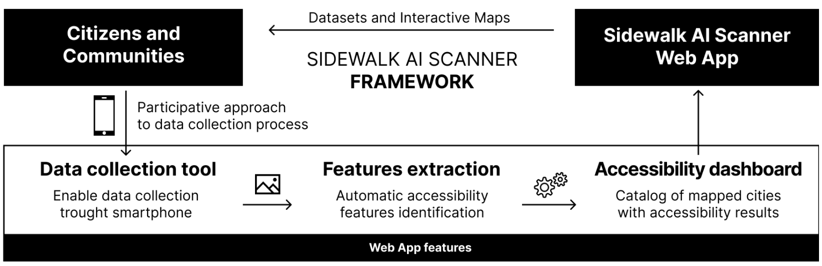
Diverse features can impact the accessibility of sidewalks. Individuals with disabilities encounter multiple barriers that complicate or prevent mobility within urban settings. An accessible sidewalk should ensure accessibility that meets the needs of all citizens beyond merely physical impairments (1; 23; 45). Through a comprehensive analysis of literature and previous case studies, we catalogued a set of features that are crucial to ensure citizens' access and safety irrespective of their physical, sensory, or cognitive limitations, whether temporary or permanent. Table 1 lists the chosen accessibility elements.

(c) Data collection technologies

We started by considering all the data formats already investigated in previous studies. The formats considered were 3D-point clouds, depth maps, and photographic images/videos. The comparison of technologies takes into account the Sidewalk AI scanner's primary user, i.e., any citizen interested in collecting sidewalk data without prior knowledge and experience in urban data collection.

As shown in Table 2, the most accessible data format for non-expert users is represented by photographic images/videos. This data format acquisition also requires the slightest prior knowledge. This is significant as methods based on crowdsourced approaches have reported that user training is a critical aspect (16). Although cheap compact and action cameras also allow video recording, the availability of such a device is lower than that of smartphones from an opportunistic data collection point of view. This represents a critical evaluation aspect, as carrying a camera during city movements is less common than carrying a smartphone. Moreover, most of

Figure 1. Sidewalk AI scanner framework



the cheapest cameras do not equip basic sensors like GPS or accelerometers, which are instead basic features of any smartphone. About the 3D-point cloud, while this data format appears to be the most promising and accurate in literature, it is also the most expensive and least accessible when considering citizens as users. In fact, in addition to the high cost of LiDAR sensors, which can reach hundreds of thousands of dollars, this is compounded by an extremely high knowledge cost, as these systems do not present a user-friendly interface and require specific knowledge and dedicated software to be utilized.

Table 2. Data format acquisition requirements. Access Cost refers to the financial investment needed for equipment and software; Required Knowledge refers to the technical expertise or training required to operate the technology to collect the data; Computational Power indicates the hardware requirements needed to process the data.

Data Format	Access Cost	Required Knowledge	Computational Power
Point Cloud	High	High	High
Depth Map Camera	Low	Medium	Medium
Photographic Images/Videos	Low	Low	Low

Table 3. Comparison of iPhone LIDAR and Smartphone Camera for data collection. Diffusion indicates the technology's market penetration; Ease of Use assesses the user-friendliness and accessibility; Collection time measures the efficiency in gathering spatial data over a 40-meter sidewalk; Edge computational Power evaluates the technology's demand on device resources.

	iPhone LiDAR	Smartphone Camera
Diffusion	Low	High
Ease of Use	Medium	High
Collection time (40m)	3 min	<1 min
Edge Computational Power	Medium	Low

In recent years, Apple has commercialized smartphones equipped with micro LiDAR sensors. This sensor, available in all Pro and Max versions of the iPhone starting from the 12 models, can generate the same data format as professional LiDAR sensors. Once it was determined that the smartphone is the most suitable technology to involve citizens in participatory sidewalk data collection, a second evaluation was performed to determine which between the LiDAR sensor and smartphone camera represents the most accessible and fast method (As seen in Table 3). The comparison results emphasize that an approach based on data collection with smartphone video cameras represents the most accessible, low-cost, and user-friendly method. In addition to ease of use, video recording while walking on a sidewalk is significantly faster than LiDAR. Furthermore, the iPhone does not provide a proprietary application that allows the use of the LiDAR sensor, which instead requires the download of third-party applications. These applications, which usually only provide access to a limited set of LiDAR-based functionalities, require a payment subscription to unlock all functionalities. At the same time, the widespread global proliferation of smartphones allows our approach to be highly inclusive and scalable to any citizens worldwide with access to a smartphone with a camera, not being limited to Apple devices.

(d) Web App

The Sidewalk AI Scanner is a web app organized into two main sections. The *data collection tool* enables users to access data collection functionalities through smartphones. The *Sidewalk accessibility dashboard* provides access to a catalog of cities that have already been mapped using this approach, in the form of interactive maps.

Data collection tool The data collection tool section provides an interface for users to collect video through their smartphones in a standardized manner. Following the recording phase, the tool permits the uploading of the registered video. Comprehensive guidelines and best practices for video recording are provided to ensure the quality and consistency of the data collected. Visual aids, including illustrative images and detailed video tutorials, assist users in adhering to these specifications. These include directives on optimal camera framing, orientation, and positioning.

Sidewalk accessibility dashboard The sidewalk accessibility dashboard provides a catalog of currently mapped cities, displaying a global and a city-specific accessibility level. The accessibility level is derived from the average level of accessibility of the mapped sidewalks. From this Section, users can access the individual city map, where, for each mapped sidewalk segment, an accessibility level is given along with a list of identified or missing accessibility features. Additionally, an indication of the currently mapped surface area of the selected city is provided. A legend displays the average daily number of sidewalks mapped by all contributors and an value indicating the current data's recency. The color coding of the sidewalk segments represents the level of accessibility determined by the latest scan. While a green color denotes an accessible sidewalk, a red color indicates a sidewalk that is inaccessible to one or more disabilities. The opacity of the sidewalk representation indicates the recency of the scan. The less recent the scan, the lower the opacity of the color displayed. This feature aids users in identifying which sidewalks within the city have been updated most recently and which require a new scan to verify the displayed accessibility data.

4. Validation

The validation focuses on demonstrating the feasibility of the proposed data collection tool in a real-world setting. The width of a sidewalk is one of the most crucial factors impacting sidewalk accessibility (34). While detecting sidewalk features from a sidewalk image, such as obstacles, pavement type, or curbs, is feasible by deploying a dedicated semantic segmentation model, sidewalk width estimation from an image is a more complex problem. For this reason, the validation process has focused on assessing the feasibility of estimating width from sidewalk images, regardless of the type of smartphone used to capture it.

This Section reports the systematic process adopted to address and validate a visual AI-based sidewalk width calculation method. The validation process was designed to address the performance of the two main models that must be combined to make the method work: a semantic segmentation model and a width estimation model. Both models were tested in a preliminary case study conducted in Cambridge, MA, USA.

(a) Semantic Segmentation Model

The first step to calculating the width of a sidewalk is to identify its boundaries within the image. Several pre-trained segmentation models already encompass the sidewalk class (47; 48; 49; 50). From preliminary tests conducted with these models on sidewalk images, it has emerged that the accuracy is significantly lower compared to SVI results. This discrepancy may be attributed to the road-centric perspective of the training dataset.

Out of all the tested models, PSPUNet (50) shows prosiming performance when run on sidewalk images. PSPUNet is the combination of Pyramid Scene Parsing Network (PSPNet) (51) and UNet (52). The detailed architecture was introduced in (50). The PSPUNet model tests 22 types of objects that obstruct walking on sidewalks. The details are shown in Table A.1. The performance of PSPNet in segmenting sidewalk boundaries in sidewalk images led to the selezion of this model for the case study.

(b) Width Estimation Model

One approach to determining the distance of an object in an image collected by a monocular camera is to use a reference object within the scene that is static and whose dimensions are already known (53). A more complex task is to measure the distance of an object in sidewalk images. In

royalsocietypublishing.org/journal/rsta Phil. Trans. R. Soc. A 0000000

such cases, using a static reference object with known dimensions is inadequate, as lens distortion and perspective impact the result accuracy. The camera's position in space or calibration are both information that could be used to increase the estimation accuracy (54).

However, a challenge arises when the dimensions and distance of an object have to be estimated in an image from a camera that has not been calibrated and which position in the space is undefined. A secondary challenge arises from the variations in lens parameters across different smartphone models. These variations result in differences in camera distortions and focal lengths, leading to significant uncertainties and errors if relying solely on the original photos for size estimation in three-dimensional space.

To face these challenges, we investigate a user-end-based approach: since the user must use the smartphone to collect the sidewalk video, the smartphone can be used to embed within the recorded video markers that a computer vision algorithm can use to estimate the distance between any two points on the sidewalk. This approach minimizes user disruption by directly optimizing data on the user's end employing smartphone-available sensors and visually embedding information within the submitted material.

The following Section outlines the design and technology behind the data collection tool prototype developed to collect sidewalk video to test the feasibility and accuracy of the proposed width estimation model.

(i) Data Collection Tool Prototype

We developed an iOS app prototype to simulate the data collection tool. The app functionalities include performing camera calibration and optimization on the user's device automatically and embedding computed spatial markers into the registered video in the form of a red grid.

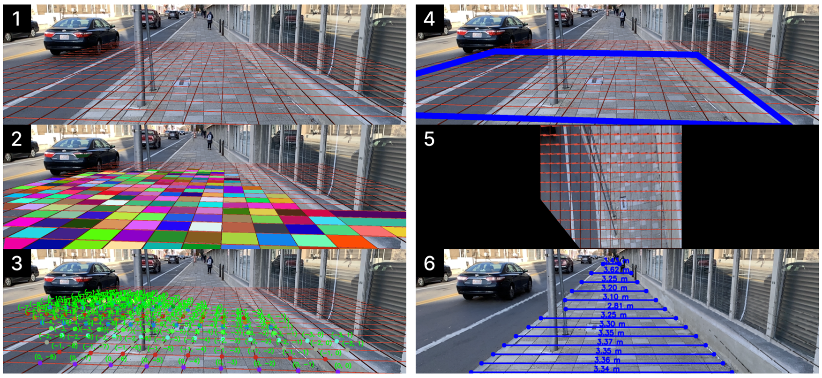
The prototype app has been developed using Flutter, a cross-platform framework developed by Google. The app's core functionality is powered by Apple's ARKit framework, which provides advanced augmented reality capabilities and three-dimensional scene understanding. Upon launching the app, the device's gyroscope and camera depth sensing capabilities, if available, are utilized to construct a three-dimensional representation of the area visible through the camera lens. The app then employs algorithms to selectively extract the ground plane from the reconstructed 3D scene. Once the ground plane is identified, a fixed-size red grid pattern with a consistent absolute length of 25 cm is dynamically generated and overlaid onto the extracted plane. This grid pattern adapts to the image perspective and extends according to the user's movements. When the user initiates video recording, the app captures the camera feed, embedding the red grid pattern within the video frames. The resulting video, complete with the integrated grid, is automatically saved upon completion of the recording. In addition to the grid-embedded video, the app also saves a separate, clear version of the recorded video. This clear version ensures that sufficient original image information is preserved for processing regions within the video frames, even if the grid pattern partially obscures certain areas.

(ii) Width Detection Pipeline

A frame is extracted from both the embedded-grid and the clear video at regular intervals. Color thresholding is applied to detect the grid by isolating it from the rest of the frame. The grid squares are identified by looking for square-shaped contours, as shown in Figure 2. Subsequently, four corner points are extracted for each detected grid square, and duplicate points are merged using DBSCAN (55). Adjacent grid squares are identified by looking for which corner points were merged and which were part of the same grid square. The locations of adjacent corner points are finally employed to generate line equations for each grid line.

The identified grid lines are used to provide a top-down view of the image using a perspective transform, also known as a homography (56). A chosen rectangle in the grid with a known metric width and height is mapped to a rectangle in the top-down view. The homography is performed by finding the matrix M that maps points in homogeneous coordinates $(x, y, 1)$ on the original image to the correct coordinates $(x', y', 1)$ in the top-down view, where $x' = Mx$ and $y' = My$.

Figure 2. Width detection pipeline: (1) Embedded-grid frame; (2) Grid identification through square-shaped contours; (3) Corner points extractions; (4) Homography perspective transformation; (5) Resulted Top-down view; (6) Width estimation result.



The top-down view is aligned with the grid, so the Euclidean distance between any two points in the image can be used to estimate their real-world distance.

It is possible to extrapolate using the inverse of the homography matrix M 1 as this matrix represents the transformation from a flat plane to the original image. For example, a landing position on the original image of the points (-20 m, -20 m) and (20 m, 20 m) can be identified and used to generate a new, larger top-down view to measure more distances.

By using the perspective transform along with the sidewalk segmentation results, the width of the sidewalk at each part of the image can be determined. Measuring sidewalk width can be performed by drawing horizontal lines across the sidewalk, with the lines evenly spaced vertically, and using the distance between these lines to get the sidewalk width.

(c) Results

The accuracy of the width estimation model in detecting sidewalk width in the preliminary case study in Cambridge shows promising results. The ground-truth comparison shows an accuracy of $\pm5\%$, which increases to $\pm10\%$ when less than 3 meters away from the camera point of view (POV) (see Figure 3). The accuracy starts to lower significantly as the distance from the camera's POV exceeds 5 meters. This accuracy reduction is influenced by grid resolution and perspective, outlining a high-accuracy range that falls between 0.50 and 5 meters from the camera POV.

After validating the initial accuracy of the model, we expanded our testing by running three more case studies, further evaluating the model's accuracy in estimating sidewalks width in different urban settings. This further evaluation focused on assessing width estimation performance inside the high-accuracy range. A total of 30 images were collected and tested, with 10 from each city: Hong Kong, China, London, UK and Milan, Italy.

As visible in Table 4, the model exhibits an overall high accuracy in estimating sidewalk width, with an average error margin of 0.30 meters compared to the ground truth reference. Notably, a primary limitation arose from the model's deployment capabilities across diverse urban settings; it successfully analyzed 63% of the images, while the remaining could not be scanned. This discrepancy was uniformly caused by the performance of the tPSPUnet model, whose accuracy vastly decreased in images where the sidewalk and the road share similar materials or colors, a scenario commonly observed in Italy.

5. Discussion

Results demonstrate that the proposed method enables the collection of one of the fundamental features for urban accessibility, namely sidewalk width, through frames extracted from smartphone videos. As the app prototype exploits sensors available in most smartphones and does not require the use of complex technologies like LiDAR, it is possible to assert that this method enable a particpatory, low-cost solution for urban accessibility mapping through a hybrid-crowdsourced approach that involves citizens.

Limitations exist and will be addressed in future studies to ensure that the proposed method delivers precise and reliable performances. The segmentation model, selected from those already available in this study, showed highly variable results. Although sufficient for this study's scope, its outcome significantly impacts the scalability of the approach. The web app deployment will require a dedicated segmentation model to ensure high precision in segmenting sidewalk boundaries combined with the capacity to identify all accessibility features listed in Table 1.

The accuracy of the width estimation model also requires further development. Results demonstrate that while a high-accuracy area is identifiable between 0.5 and 5 meters from the camera POV, the precision decreases outside of this range. While this result is acceptable for this first study, an inaccuracy greater than $\pm 10\%$ can significantly affect the reliability of the collected data and, consequently, the trustworthiness of the derived accessibility value. Although the reduction in precision can be improved by fine-tuning the model, the loss of precision outside a certain distance range is a factor that might persist due to the resolution of the grid or the quality of the smartphone camera. This can be addressed by defining a distance limit beyond which the elements present in the frame should not be considered. Because the model is based on video recordings, sampling of frames at predetermined intervals is possible. This ensures sidewalk coverage at least every 5 meters, preventing distance reliability loss to impact results. Future work will focus on improving the width estimation model by a process of stitching the results of each 5-meter frame analysis together, reconstructing a digital twin of the scanned sidewalk.

The app prototype, combined with the segmentation and width estimation model, has confirmed that the smartphone-based approach could be an easy, rapid, and relatively economical sidewalk data collection method compared to alternatives proposed by previous studies. This participative-based solution raises new possibilities and challenges. The accuracy results from

Figure 3. On the top: PSPUNet segmentation results. On the bottom: width estimation model accuracy results, ground truth dimension (yellow) and width estimation results from the model (blue).

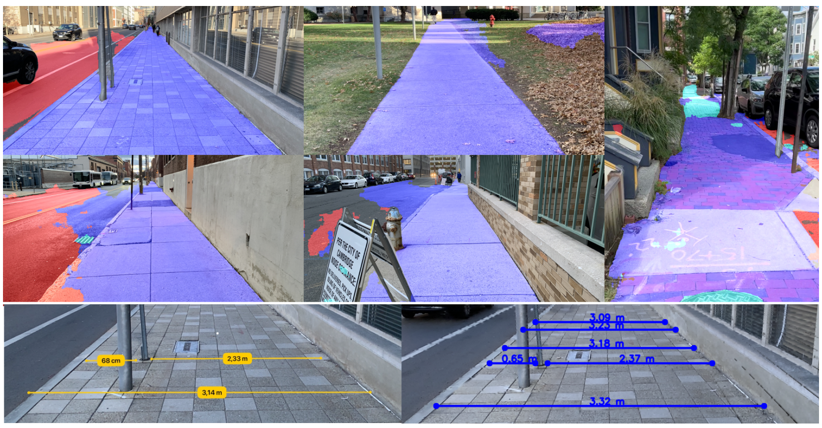


Table 4. Width estimation model result from performed case studies. Estimation success refers to the number of cases where the model doesn't fail to calculate sidewalk width. The average error refers to the discrepancy between the model's width estimations and the actual sidewalk widths (ground truth data), expressed in meters.

	Hong Kong	London	Milan	Average
Estimation Success (%)	70	80	36.36	**63.33**
Average Error (m)	0.14	0.58	0.05	**0.30**

the London case study serve as an example. Specifically, London exhibited a higher average error than the other two cities. Notably, London was also the only case study in which images and ground truth data were collected by an independent volunteer. This highlight how guidelines for user participation will be necessary, and it will be essential to further investigate which factors influence the accuracy of the result and how to provide the users with guidance on how to record the video (e.g. using the phone in landscape mode, walking at a fixed speed).

Overall, the results confirm smartphone's videos are sufficient to provide information to study sidewalk accessibility, allowing the data collection approach to be accessible regardless of the device's manufacturer, model, and performance. This opens new veins in urban accessibility mapping, where citizens are not asked to perform time-consuming and complex tasks but instead are empowered to contribute to data collection simply by recording videos of the sidewalks they walk. To our knowledge, this innovative approach, proven to be feasible through this study, has never been explored before. Considering the ease with which final users can access it, this method could have a huge impact once made available through a web app like Sidewalk AI Scanners.

Finally, the user-centered approach proposed in this paper would require to further investigate what triggers and supports participation within communities. Strategies to increase participation, such as gamification, community outreach programs, and data collection workshops are all methods worth consideration. At the same time, engaging with fragile communities, like people with disabilities, could represent a way to involve users who are also driven by a direct personal return other than just interest or the desire to contribute.

6. Conclusion & future works

This paper presents Sidewalk AI Scanner, a novel hybrid-crowdsourced approach for sidewalk data collection. The method is founded on a web app that enables citizens to upload sidewalk videos registered with a smartphone and then generate interactive accessibility maps starting from the automatic analysis of the user-generated videos. This participatory method represents a scalable solution for municipalities to fill the lack of sidewalk datasets, essential to generating accessible maps and assisting people with disabilities in urban navigation. Through case study conducted in various cities, the method's capability to address the sidewalk width estimation challenge in an image without the need for prior camera calibration has been evaluated. The results, showing high accuracy compared to ground truth, confirm the web app's potential to provide a low-cost tool to enable citizen participation in urban data collection experiments.

By providing a scalable solution, the Sidewalk AI Scanner approach aligns with broader efforts to intervene on barriers faced by people with disabilities in urban environments. The participatory model exemplifies the power of participative action in urban data collection and highlights the transformative potential of visual AI technology in fostering more accessible and equitable urban landscapes, advancing the conversation around urban inclusivity.

Future work will concentrate on giving access to the web app to communities and on training a semantic segmentation model that can identify all the crucial accessibility features necessary for estimating a accessibility level in a sidewalk images. The goal is to deliver a tool to enable worldwide participation in accessibility mapping efforts. Such a tool could significantly empower individuals to actively contribute to enhancing the accessibility of urban landscapes in their communities, supporting the push towards a more inclusive society.

If further result will be promising, the hybrid-crowsourced approach could be extended to other domains that share similar goals and challenges as the sidewalk accessibility one. For example, aspects such as the accessibility of public buildings, stations, subways, or green areas could all be domains to which a similar approach be applied, providing data for the design of targeted and effective urban accessibility interventions, contributing towards the goal of making future cities more accessible and inclusive for everyone.

References

1 Evcil AN. 2012 Raising Awareness about Accessibility. *Procedia - Social and Behavioral Sciences* **47**, 490–494. Cyprus International Conference on Educational Research (CY-ICER-2012)North Cyprus, US08-10 February, 2012 (https://doi.org/10.1016/j.sbspro.2012.06.686)
2 Labbé D, Eisenberg Y, Snyder D, Shanley J, Hammel JM, Froehlich JE. 2023 Multiple-Stakeholder Perspectives on Accessibility Data and the Use of Socio-Technical Tools to Improve Sidewalk Accessibility. *Disabilities* **3**, 621–638. (10.3390/disabilities3040040)
3 Mitchell C. 2006 Pedestrian mobility and safety: a key to independence for older people. *Topics in Geriatric Rehabilitation* **22**, 45–52.
4 Kirchner CE, Gerber EG, Smith BC. 2008 Designed to Deter: Community Barriers to Physical Activity for People with Visual or Motor Impairments. *American Journal of Preventive Medicine* **34**, 349–352. Theme Issue: Active Living in Diverse and Disadvantaged Communities (https://doi.org/10.1016/j.amepre.2008.01.005)
5 Christensen K, Holt JM, Wilson JF. 2010 Effects of perceived neighborhood characteristics and use of community facilities on physical activity of adults with and without disabilities. *Prev Chronic Dis. 7(5)*.
6 Pineda VS, Meyer S, Cruz JP. 2017 The Inclusion Imperative. Forging an Inclusive New Urban Agenda. *The Journal of Public Space* **2**, 1–20. (10.5204/jps.v2i4.138)
7 Froehlich JE, Eisenberg Y, Hosseini M, Miranda F, Adams M, Caspi A, Dieterich H, Feldner H, Gonzalez A, De Gyves C, Hammel J, Kirkham R, Kneisel M, LabbÉ D, Mooney SJ, Pineda V, PinhÃO C, RodrÍGuez A, Saha M, Saugstad M, Shanley J, Sharif A, Shen Q, Silva C, Sukel M, Tokuda EK, Zappe SF, Zivarts A. 2022 The Future of Urban Accessibility for People with Disabilities: Data Collection, Analytics, Policy, and Tools. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* ASSETS '22 New York, NY, USA. Association for Computing Machinery. (10.1145/3517428.3550402)
8 Efthimis Kapsalis NJ, Hale J. 2022 Disabled-by-design: effects of inaccessible urban public spaces on users of mobility assistive devices – a systematic review. *Disability and Rehabilitation: Assistive Technology* **0**, 1–19. PMID: 35984675 (10.1080/17483107.2022.2111723)
9 Zając AP. 2016 City Accessible for Everyone – Improving Accessibility of Public Transport Using the Universal Design Concept. *Transportation Research Procedia* **14**, 1270–1276. Transport Research Arena TRA2016 (https://doi.org/10.1016/j.trpro.2016.05.199)
10 . 1990 Americans with Disabilities Act of 1990. Pub. L. No. 101-336, § 2, 104 Stat. 328 (1990). U.S. Department of Justice, Civil Rights Division, Disability Rights Section.
11 National Council on Disability. 2007 The Impact of the Americans with Disabilities Act: Assessing the Progress Toward Achieving the Goals of the ADA. https://www.aucd.org/docs/ADA%20Impact%20Report.pdf. [Accessed: 2024-01-22].
12 Clifton KJ, Livi Smith AD, Rodriguez D. 2007 The development and testing of an audit for the pedestrian environment. *Landscape and Urban Planning* **80**, 95–110. (https://doi.org/10.1016/j.landurbplan.2006.06.008)
13 Froehlich JE, Brock AM, Caspi A, Guerreiro Ja, Hara K, Kirkham R, Schöning J, Tannert B. 2019 Grand challenges in accessible maps. *Interactions* **26**, 78–81. (10.1145/3301657)
14 Rundle AG, Bader MD, Richards CA, Neckerman KM, Teitler JO. 2011 Using Google Street View to Audit Neighborhood Environments. *American Journal of Preventive Medicine* **40**, 94–100. (https://doi.org/10.1016/j.amepre.2010.09.034)
15 Saha M, Saugstad M, Maddali HT, Zeng A, Holland R, Bower S, Dash A, Chen S, Li A, Hara K, Froehlich J. 2019 Project Sidewalk: A Web-based Crowdsourcing Tool for Collecting Sidewalk Accessibility Data At Scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* CHI '19 p. 1–14 New York, NY, USA. Association for Computing Machinery. (10.1145/3290605.3300292)
16 Hosseini M, Saugstad M, Miranda F, Sevtsuk A, Silva CT, Froehlich JE. 2022 Towards Global-Scale Crowd+AI Techniques to Map and Assess Sidewalks for People with Disabilities. .
17 DIAUD & CBM. 2016 The Inclusion Imperative: Towards Disability-inclusive and Accessible Urban Development. https://www.cbm.org/fileadmin/user_upload/Publications/The-Inclusion-Imperative....pdf. [Accessed 2024-03-01].

18 Frías-López E, de Llano JQ. 2020 Methodology for 'reasonable adjustment' characterisation in small establishments to meet accessibility requirements: A challenge for active ageing and inclusive cities. Case study of Madrid. *Cities* **103**, 102749. (https://doi.org/10.1016/j.cities.2020.102749)
19 United Nations. 2015 Transforming our world: the 2030 Agenda for Sustainable Development. https://sdgs.un.org/2030agenda. [Accessed: 2024-02-28].
20 United Nations. 2016 Good Practices in Accessible Urban Development. https://www.un.org/disabilities/documents/desa/Good_Practicies...pdf. [Accessed 2024-02-12].
21 Imrie R. 2000 Disabling Environments and the Geography of Access Policies and Practices. *Disability and Society* **15**, 5 – 24. Cited by: 81 (10.1080/09687590025748)
22 City of Boston. 2019 Boston Complete Streets Guidelines. https://www.boston.gov/sites/default/files/file/2019/12/BCS_Guidelines.pdf.
23 Ferreira MA, da Penha Sanches S. 2007 PROPOSAL OF A SIDEWALK ACCESSIBILITY INDEX. *Journal of Urban and Environmental Engineering* **1**, 1–9.
24 Froehlich JE, Saugstad M, Martínez E, de Buen Kalman R. 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*.
25 Hara K, Chan C, Froehlich JE. 2016 The Design of Assistive Location-based Technologies for People with Ambulatory Disabilities: A Formative Study. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* CHI '16 p. 1757–1768 New York, NY, USA. Association for Computing Machinery. (10.1145/2858036.2858315)
26 Ai C, Tsai YJ. 2016 Automated Sidewalk Assessment Method for Americans with Disabilities Act Compliance Using Three-Dimensional Mobile Lidar. *Transportation Research Record* **2542**, 25–32. (10.3141/2542-04)
27 Hou Q, Ai C. 2020 A network-level sidewalk inventory method using mobile LiDAR and deep learning. *Transportation Research Part C: Emerging Technologies* **119**, 102772. (https://doi.org/10.1016/j.trc.2020.102772)
28 Frackelton A, Grossman A, Palinginis E, Castrillon F, Elango V, Guensler R. 2013 Measuring walkability: Development of an automated sidewalk quality assessment tool. *Suburban Sustainability* **1**, 4.
29 Biagi L, Brovelli MA, Stucchi L. 2020 Mapping the accessibility in openstreetmap: A comparison of different techniques. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **43**, 229–236.
30 Mobasheri A, Huang H, Degrossi LC, Zipf A. 2018 Enrichment of OpenStreetMap Data Completeness with Sidewalk Geometries Using Data Mining Techniques. *Sensors* **18**. (10.3390/s18020509)
31 Comai S, De Bernardi E, Salice F, Vali A. 2019 Maps for easy paths (MEP): a Mobile application for City accessibility. *Mobile Solutions and Their Usefulness in Everyday Life* pp. 105–125.
32 Ning H, Ye X, Chen Z, Liu T, Cao T. 2022 Sidewalk extraction using aerial and street view images. *Environment and Planning B: Urban Analytics and City Science* **49**, 7–22.
33 Senlet T, Elgammal A. 2012 Segmentation of occluded sidewalks in satellite images. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)* pp. 805–808.
34 Coppola NA, Marshall WE. 2021 Sidewalk Static Obstructions and Their Impact on Clear Width. *Transportation Research Record* **2675**, 200–212. (10.1177/0361198121991833)
35 Weld G, Jang E, Li A, Zeng A, Heimerl K, Froehlich JE. 2019 Deep learning for automatically detecting sidewalk accessibility problems using streetscape imagery. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* pp. 196–209.
36 Biljecki F, Ito K. 2021 Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning* **215**, 104217.
37 Hosseini M, Sevtsuk A, Miranda F, Cesar Jr RM, Silva CT. 2023 Mapping the walk: A scalable computer vision approach for generating sidewalk network datasets from aerial imagery. *Computers, Environment and Urban Systems* **101**, 101950.
38 Ahmetovic D, Manduchi R, Coughlan JM, Mascetti S. 2017 Mind Your Crossings: Mining GIS Imagery for Crosswalk Localization. *ACM Trans. Access. Comput.* **9**. (10.1145/3046790)

39 Koh CY, Hendawi A. 2023 Towards Disability-Aware Sidewalk Routing. In *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems* pp. 1–2.

40 Plascak JJ. 2019 Sidewalk conditions in northern New Jersey: using Google Street View imagery and ordinary kriging to assess infrastructure for walking. *Preventing chronic disease* **16**.

41 Zhou H, He S, Cai Y, Wang M, Su S. 2019 Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning. *Sustainable Cities and Society* **50**, 101605. (https://doi.org/10.1016/j.scs.2019.101605)

42 Hara K, Sun J, Moore R, Jacobs D, Froehlich J. 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 p. 189–204 New York, NY, USA. Association for Computing Machinery. (10.1145/2642918.2647403)

43 Hara K, Froehlich JE. 2015 Characterizing and visualizing physical world accessibility at scale using crowdsourcing, computer vision, and machine learning. *SIGACCESS Access. Comput.* p. 13–21. (10.1145/2850440.2850442)

44 Weld G, Jang E, Li A, Zeng A, Heimerl K, Froehlich JE. 2019 Deep Learning for Automatically Detecting Sidewalk Accessibility Problems Using Streetscape Imagery. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* ASSETS '19 p. 196–209 New York, NY, USA. Association for Computing Machinery. (10.1145/3308561.3353798)

45 Axelson P. 1999 *Designing sidewalks and trails for access*. US Department of Transportation, Federal Highway Administration.

46 Maarten S, Cláudia P. 2021 Using Artificial Intelligence to measure accessibility. https://www.amsterdamintelligence.com/posts/using-artificial-intelligence-to-measure-accessibility. [Accessed: 2024-02-29].

47 Cheng B, Misra I, Schwing AG, Kirillov A, Girdhar R. 2022 Masked-attention mask transformer for universal image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* pp. 1290–1299.

48 Cai H, Gan C, Han S. 2022 Efficientvit: Enhanced linear attention for high-resolution low-computation visual recognition. *arXiv preprint arXiv:2205.14756*.

49 Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo WY et al.. 2023 Segment anything. *arXiv preprint arXiv:2304.02643*.

50 SolITLab. 2020 Codes for Road Segmentation. https://github.com/JunHyeok96/Road-Segmentation. [Accessed: 2023-09-30].

51 Zhao H, Shi J, Qi X, Wang X, Jia J. 2017 Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp. 2881–2890.

52 Ronneberger O, Fischer P, Brox T. 2015 U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* pp. 234–241. Springer.

53 Zeng J, Pei X, Xi W. 2022 The monocular ranging based on stable target detection and tracking for intelligent vehicle. In *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)* pp. 1–5. IEEE.

54 Xiong J, Zhang K. 2022 Monitoring multiple geometrical dimensions in WAAM based on a multi-channel monocular visual sensor. *Measurement* **204**, 112097.

55 Bi F, Wang W, Chen L. 2012 DBSCAN: density-based spatial clustering of applications with noise. *J. Nanjing Univ* **48**, 491–498.

56 OpenCV Developers. 2023 Geometric Transformations of Images — OpenCV 3.4 Documentation. https://docs.opencv.org/3.4/da/d6e/tutorial...html. [Accessed: 2024-02-10].

A. Additional Tables

Table A.1. PSPUnet labels of obstacle identification

Class	Detailed item	Label
Background	Background	"background"
Bike lane	Normal bike lane	"bike_lane_normal"
	Asphalt sidewalk	"sidewalk_asphalt"
	Urethane sidewalk	"sidewalk_urethane"
Caution zone	Grating	"caution_zone_grating"
	Manhole	"caution_zone_manhole"
	Repairing zone	"caution_zone_repair_zone"
	Stairs	“caution_zone_stairs"
	Tree zone	"caution_zone_tree_zone"
Crosswalk	Crosswalk of alley	"alley_crosswalk"
	Crosswalk of roadway	"roadway_crosswalk"
Braille guide blocks	Normal block	"braille_guide_blocks_normal"
	Damaged block	"braille_guide_blocks_damaged"
Roadway	Normal roadway	"roadway_normal"
	Normal alley	"alley_normal"
	Speed dump of alley	"alley_speed_bump"
	Damaged alley	"alley_damaged"
Sidewalk	Blocks	"sidewalk_blocks"
	Cement	"sidewalk_cement"
	Soil or stone	"sidewalk_soil_stone"
	Damaged	"sidewalk_damaged"
	Other	"sidewalk_other"