

Building a Crowdsourcing based Disabled Pedestrian Level of Service routing application using Computer Vision and Machine Learning

Nicolas Blanc[†], Zhan Liu^{*}, Olivier Ertz[†], Diego Rojas^{*}, Romain Sandoz[†],
Maria Sokhn[§], Jens Ingensand[†], Jean-Christophe Loubier^{*}

^{*}University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis), Sierre, Switzerland
{zhan.liu, diego.rojas, jchristophe.loubier}@hevs.ch

[†]University of Applied Sciences and Arts Western Switzerland (HEIG-VD), Yverdon-les-Bains, Switzerland
{nicolas.blanc1, olivier.ertz, romain.sandoz, jens.ingensand}@heig-vd.ch

[§]University of Applied Sciences and Arts Western Switzerland (HEG Arc HES-SO), Neuchatel, Switzerland
maria.sokhn@he-arc.ch

Abstract—The availability of global and scalable tools to assess disabled pedestrian level of service (DPLoS) is a real need, yet still a challenge in today’s world. This is due to the lack of tools that can ease the measurement of a level of service adapted to disabled people, and also to the limitation concerns about the availability of information regarding the existing level of service, especially in real time. This paper describes preliminary results to progress on those needs. It also includes a design for a navigation tool that can help a disabled person move around a city by suggesting the most adapted routes according to the person’s disabilities. The main topics are how to use advanced computer vision technologies, and how to benefit from the prevalence of handheld devices. Our approach intends to show how crowdsourcing techniques can improve data quality by gathering and combining up-to-date data with valuable field observations.

I. INTRODUCTION

Creating a disabled pedestrian map-based routing is a challenge for both researchers and practitioners. In recent years the assessment of the pedestrian level of service has been discussed by an increasing number of researchers. Routing solutions provide information about the available facilities for surmounting transitory obstacles, so people with mobility disabilities and people with baby strollers are able to travel more easily. In fact, there is considerable demand for disabled pedestrian level of services. According to the World Health Organization [1], over a billion people, about 15% of the world’s population, live with some form of disability, and between 110 and 190 million adults, including wheelchair users, have very significant difficulties in mobility. Fortunately, research is becoming more and more interdisciplinary to make contributions in this domain.

Despite the existence of standards and regulations to increase the ability of both disabled people and persons with limited mobility to independently use pedestrian networks, there is still a lack of replicable, objective, cost-effective systems to assess pedestrian infrastructure [2].

Therefore, two main limitations have been identified. First, there is no model that can make ease the measurement of a level of service (LoS) adapted to people with mobility disabilities. Such a model could be used to assess the level of accessibility of a pedestrian network. Second, there are few tools and little information regarding the existing level of service in real time, and this lack causes problems. When moving through a road network, a disabled person needs these kinds of tools and information to make the best decisions.

In this study, we present a technical approach named CrowDPLoS, which uses advanced computer vision technologies and benefits from the prevalence of handheld devices in order to meet the challenges discussed above. The main and long-term objective is to develop a navigation tool with crowdsourcing techniques that can help people with disabilities to move around a city, by suggesting the most adapted routes according to the person’s disabilities. To summarize, the best route in our approach is based on the highest level of service in sections of the route, and it is adapted to the user’s handicap. Specifically, we give people with mobility disabilities, mainly wheelchair users, an innovative solution for travelling safely and with ease through a modern urban area by applying the algorithms of computer vision. These algorithms use crowdsourced data (e.g. images), which are taken by the users of a mobile application, to determine a level of service adapted to disabled pedestrians.

This paper is organized as follows: section 2 presents a summary of the related work regarding computer vision based techniques and crowdsourcing in disabled pedestrian level of service. In section 3, we explain our approaches to develop models that measure a crowdsourcing based disabled pedestrian level of service for pedestrian networks, which applies computer vision and machine learning technologies. Finally, we conclude with a summary of the current work, and we present suggestions for future research.

II. RELATED WORK

A. Models of disabled pedestrian level of service

The safety and accessibility of pedestrian infrastructure regulation developments have historically been supported by governmental institutions. For example, the Americans with Disabilities Act (ADA) has, since 1990, guided the implementation of regulations and specifications for pedestrians in the US.

The first formalization of sidewalks performance assessments was carried out by Fruin in 1971 [3]. According to Landis et al. [4], Fruin's approach was the only quantifying established methodology to measure the capacity of sidewalks. However, Fruin's methodology has practical limitations, and in fact only around three percent of the sidewalks could be effectively assessed.

Sidewalks are one of the main and secure places for pedestrians to walk in urban areas. Several studies have focused their research on characterizing and assessing sidewalk accessibility, sometimes including perceived security from the pedestrian's point of view [5]. According to Landis et al. [4], it is more difficult to assess the conditions of walking routes than vehicular roadways. Furthermore, research focused on PLoS for disabled people (DPLoS) is rather new. Reviewing several studies and guidelines, Asadi et al. [6] calculated 10 main feature indicators for the presence of facilities for disabled people, such as wheelchair users and blind people. Once the 10 DPLoS indicators were developed, they proceeded to establish 10 PLoS indicators with the same methodology. The combination of the DPLoS and PLoS is a General Pedestrian Level of Service (GPLoS), which assesses the inclusive walking conditions for both pedestrians and disabled users.

B. Computer Vision and machines learning based techniques for assessing relevant street features

Historically, computer vision tends to extract the most important features from images such as edges [7] and specific patterns. The Hough transform is a well known line detector that can extract street borders [8] as well as the vanishing point.

The slope of a street and the detection of obstacles are focus of research in the subject of accessibility. Using geolocation with odometric data, mainly GPS trajectory, is a common solution. Hara et al. [9] show how computer vision may be used to determine accessibility from physical features of the real world, by coupling Google Street View images with machine learning techniques. Within the latter domain, convolutional neural networks (CNN) are especially designed to extract features from images [10]. The field of autonomous driving is based on these techniques and it offers a solid background for street images segmentation [11]. Existing street scene datasets along with semantic segmentation and region-based methods [12] have proven the need for fine-grain labeled images to suit a particular task. Using the weights of low-level features (i.e. minor details of the image - lines or dots) that are extracted by CNN on these datasets could be a good starting point to build a more specific architecture dedicated to the DPLoS features extraction.

Moreover, stereoscopic vision provides, at low cost, the useful depth information as a disparity map. This is done

by computing a 3D point cloud from stereo image pairs, which can help to better isolate relevant objects. Coughlan and Shen [13] make use of an embedded stereo camera to detect negative obstacles on the wheelchair user's path.

But there are still two main bottlenecks to building a supervised computer vision classifier based on machine learning. These are 1) data acquisition and 2) the presence of high-quality associated labels. If nowadays huge amounts of data are available, such as Mapillary or Cityscapes Datasets [14], they may not be labeled well enough, nor suited to a given specific research topic. Hence, well managed crowdsourcing tasks may provide invaluable help for the topic of DPLoS.

C. Benefits of a crowdsourcing approach

One main benefit offered by crowdsourcing is the high potential number of actionable users. After quantity is attained, quality is added as an important control. To ensure inclusive walking conditions, DPLoS estimation must apply to a wide range of pedestrian populations, with different characteristics. Precisely crowdsourcing techniques offer the ability to reach a wider and more diverse panel of citizens, although targeting strategies may also be required to reach more specific audiences. Concerning the purpose of quality, Liu et al. [15] identified three useful techniques for improving the quality of crowdsourced data when building accessibility maps, including qualification tests, reputation systems, and aggregation techniques. Their findings highlighted that the accuracy rate has a significant increase after applying the intervention of quality control methods. Finally, they also address the third and overarching issue not to neglect which is to foster the motivation of the contributing users. While they explore a way to motivate actions by a desire for rewards, they also advise exploring how efficient the application of game theory elements can be.

D. Combination of crowdsourcing and computer vision to evaluate the DPLoS model

Computer vision is a useful technology for autonomous driving and objects detection, but its features extraction relies frequently on domain experts. Integrating crowdsourcing is not an easy task. A recent survey by Kovashka et al. [16] addressed some of the most important questions about computer vision and crowdsourcing, which provided useful general guidelines. However, only a few studies took advantage of both techniques for DPLoS assessment and the application design. Hara and Froehlich [17] showed the value of integrating final users in (1) the design of an application devoted to disabled pedestrians, and (2) the assessment of adaptive street segments aptitude.

More recently, Hara et al. [9], [18] designed a specific tool which makes use of Amazon Mechanical Turk (AMT) to provide rapidly, and with lower costs on image detection tasks. Recall of AMT workers combined with computer vision algorithms was higher and faster than AMT workers alone. Therefore, a combination of both technologies could provide higher accuracy on objects detection.

III. METHODOLOGY AND PRELIMINARY RESULTS

Based on the relevant literature, our research methodology is divided into three main phases: 1) the creation

of the DPLoS model; 2) the establishment of a pedestrian network at a city-scale level; and 3) the combination of computer vision algorithms and crowdsourcing techniques to collect data, digitize and label features. We also discuss how to increase user engagement in Crowdsourcing works and finally evaluate the detection results.

A. DPLoS model

In this section we synthesize the development process of the DPLoS model: the chosen indicators and the mathematical approach used to implement the model.

The DPLoS established model contains the 5 following features based on the literature and on the data availability, in order to ease the replication process: 1) slope; 2) sidewalk width(s) and presence of obstacles; 3) crosswalks and security signals (e.g. pedestrian light, refuge island); 4) presence of curb ramps to cross the road; and 5) coating quality (e.g. pavement, smooth concrete).

A normalized score S_i for each feature i is then weighted to compute a global DPLoS score for any street edge, as shown in equation (1).

$$\frac{1}{5} \sum_{i=1}^5 S_i \cdot w_i \quad (1)$$

PLOs weights w_i calculations are based on local conditions and street users' perceptions. A percentage expression is used to facilitate the interpretation of the DPLoS score. To obtain this percentage value, Assadi et al. [6] proposed calculating a ratio between the observed DPLoS and the ideal DPLoS, which they called a DPLoS percentage. We have adapted their equation for the five chosen indicators, as shown in the following two equations.

$$DPLoS = \sum_{i=1}^5 c_i DI_i \quad (2)$$

Where DPLoS = disabled pedestrian level of service; i = indicator number; c = coefficient of disabled pedestrian indicator; DI = disabled pedestrian indicator score. The value (c) represents the priority and importance given to the indicator in the literature and DI is the actual score of a street when compared by the standards. DI is a value between 0 and 1, where a score equal to 1 represents ideal conditions. The final DPLoS result will be used in the equation (3) as follows:

$$DPLoS_{percentage} = 100 \cdot \frac{DPLoS}{\sum_{i=1}^5 C_i} \quad (3)$$

The DPLoS percentage is classified in different categories that represent the street quality conditions. The DPLoS percentage rating $A = 80 - 100$ means the highest quality. $B = 60 - 79$ means high quality. $C = 40 - 59$ represents average quality. $D = 20 - 39$ means low quality. $E = 1 - 19$ represents the lowest quality and $F = 0$ that there is not a standard disabled pedestrian infrastructure.

B. Pedestrian network as an input

The DPLoS computation for street edges requires a pedestrian network as an input. In order to build it, we relied on existing open data such as the National Map Agency (NMA) Cadastral Survey, OSM streets and footways. In order to obtain the high-quality vector data to draw the pedestrian network, we used segmentation-based classification methods to extract a road network from the high-resolution satellite data.

C. Combination of algorithms and crowdsourcing

In this section, we first discuss how to increase user engagement in crowdsourcing works, and then we use four steps to explain how the computer vision algorithms and crowdsourcing techniques could be combined together. These steps contain: 1) footways network update; 2) data acquisition; 3) image features digitizing and labeling; and 4) end-user route choice and update abilities.

In order to make crowdsourcing project work, it faces many challenges, range from get users participation to keep user participation, from control data quality to decide on the legal aspects [19]. Literature [20] emphasizes the importance of motivations. A recent study conducted by Zeng et al. [21] found that user motives can be enhanced with the integration of game elements. In this study, we intend to integrate reward, recognition and progress related game elements where many missions are available in relation to the above needs. From the moment of "boarding", users will be prompted to accomplish tasks such as identifying crossroads on satellites images or identify images with a given feature (crossroad and curb ramp).

1) *Footways network update*: considering OpenStreetMap as the initial pedestrian network, which may not be totally fulfilled due to its voluntary basis and the interest of local citizen to contribute geographic data. A crowdsourcing tool could then be relevant to complete features important for accessibility and make the network fit the actual reality more precisely, based on a recent orthophoto background with a 30 cm minimal resolution and a sub-meter accuracy.

2) *Data acquisition and geolocation*: in this step, a second crowdsourcing tool may be relevant to serve as a feeding source for raw images. While applying crowdsourcing motivation methods [15], this tool invites application users to take pictures with their mobile device. These will be used as raw data input for the digitizing stage. If the picture's position and orientation (azimuth) values are recognized from a GPS and compass, then it can be stored directly in the database. Otherwise, a function will be proposed to the user to geolocate and orient his or her pictures before uploading them.

3) *Feature digitizing and characterizing*: the main goal of the third step is to detect one or more of the five identified DPLoS features on raw images, and mark them with a score based on the DPLoS model. If the former can be done by a well-trained CNN (fig. 1), the latter can be evaluated through crowdsourcing approach.

Let us think about the following examples. If a mobile user sees a curb ramp on a picture, he/she will be able to digitize it more accurately then mark it as a more friendly curb ramp (fig. 2). Another example would be to let the

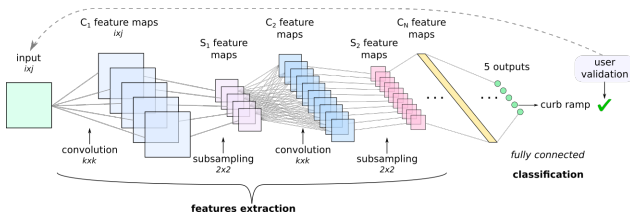


Figure 1. A CNN detecting a curb ramp. User checking could disable the label associated with this image if the result is a false positive.

user annotate a sidewalk surface width as "impossible for a single wheelchair", based on the perceived width on the given picture. If depth information is also available, a distance-measuring tool may also be provided to allow the user to directly measure the sidewalk width from the image.

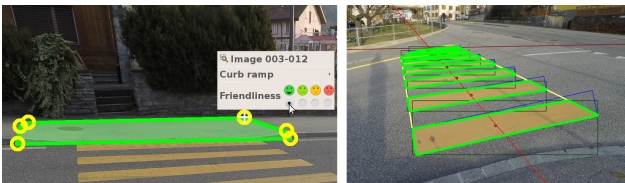


Figure 2. Image examples of a manually digitized curb ramp with a friendliness score selection toolbox (left) and a true positive automatically detected crosswalk (right).

This stage will also give the user a tool to evaluate the features that were detected by algorithms as shown in fig. 3. This important step provides two advantages. It will first classify automatically detected features according to a standard confusion matrix. And then, machine learning detection algorithms could be refined with these new positively and negatively labeled images. Training and testing results will be tracked to avoid overfitting, in addition to some optimizations.

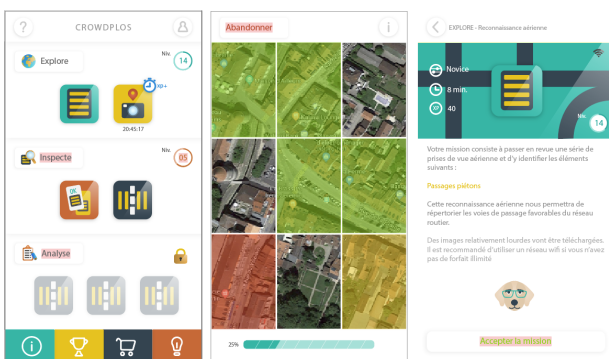


Figure 3. Evaluation tool

4) *End-user route choice assessment and update:* in the final step, our system will suggest the route with the best DPLoS value between two points, based on the weighted sum of the DPLoS of each n edges composing the route as defined by the equation (4).

$$DPLoS_{route} = \frac{1}{n} \sum_{n} DPLoS_{edge\ n} \quad (4)$$

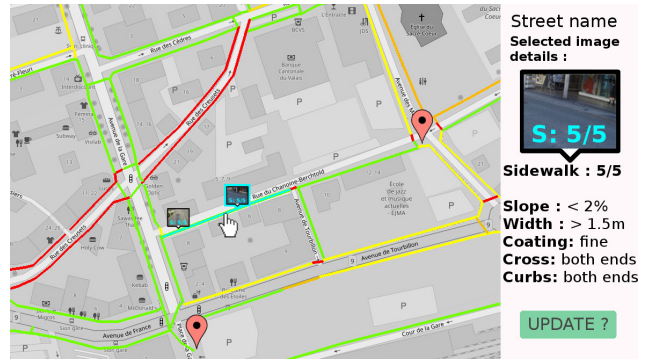


Figure 4. The prototype of the final application with edges of the pedestrian graph colored according to their DPLoS value (red=low value, green=high value). The selected edge shows that two different images contributed to its score. The features' details for this edge and the selected image are shown on the right.

Information about the edge as well as an update button (fig. 4) will offer the user the possibility to give his/her feedback on a selected segment, for example if he or she sees a mismatch between the real world and a digitized feature. This will continuously improve the pedestrian network and update it as the urban environment changes.

IV. CONCLUSION AND FUTURE WORK

In this paper, we describe how to build a crowdsourcing-based disabled pedestrian level of service routing application that uses computer vision and machine learning. We analyze the existing technical algorithms for image processing, and explain their advantages and drawbacks. More importantly, we present the combination of these algorithms with crowdsourcing techniques to improve the quality and the specificity of features detection. Our methodology contains a four-step process, which gives a detailed view to the practitioners about this combination. The preliminary results highlight that the accuracy of the features detection could be improved by applying the technique of crowdsourcing. Given that crowdsourcing is a centerpiece in this project, some future work must be dedicated to motivation methods, so as to engage citizens to participate in the achievement of the several tasks required by our approach. In the continuation of Liu et al. [15], and with the same focus on boosting the motivation of participants and data quality control, we plan to investigate how to use a gamification conceptual framework to help understand how human motivations can be directed by game elements (e.g. Octalysis Framework [22]) and to develop incentives to execute tasks.

REFERENCES

- [1] W. H. Organization. (2017) 10 facts on disability.
- [2] A. Frackelton, A. Grossman, E. Palinginis, F. Castrillon, V. Elango, and R. Guensler, "Measuring walkability: Development of an automated sidewalk quality assessment tool," *Suburban Sustainability*, vol. 1, no. 1, p. 4, 2013.
- [3] J. J. Fruin, "Pedestrian planning and design," Tech. Rep., 1971.
- [4] B. Landis, V. Vattikuti, R. Ottenberg, D. McLeod, and M. Guttenplan, "Modeling the Roadside Walking Environment: Pedestrian Level of Service," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1773, pp. 82–88, Jan. 2001.

- [5] D. Tan, W. Wang, J. Lu, and Y. Bian, "Research on Methods of Assessing Pedestrian Level of Service for Sidewalk," *Journal of Transportation Systems Engineering and Information Technology*, vol. 7, no. 5, pp. 74–79, Oct. 2007.
- [6] Asadi-Shekari Zohreh, Moeinaddini Mehdi, and Zaly Shah Muhammad, "Disabled Pedestrian Level of Service Method for Evaluating and Promoting Inclusive Walking Facilities on Urban Streets," *Journal of Transportation Engineering*, vol. 139, no. 2, pp. 181–192, Feb. 2013.
- [7] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [8] N. Kiryati, Y. Eldar, and A. M. Bruckstein, "A probabilistic Hough transform," *Pattern Recognition*, vol. 24, no. 4, pp. 303–316, Jan. 1991.
- [9] K. Hara, J. Sun, R. Moore, D. Jacobs, and J. Froehlich, "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*, ser. UIST '14. New York, NY, USA: ACM, 2014, pp. 189–204.
- [10] Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.
- [11] J. M. Alvarez, T. Gevers, Y. LeCun, and A. M. Lopez, "Road scene segmentation from a single image," in *European Conference on Computer Vision*. Springer, 2012, pp. 376–389.
- [12] G. Neuhold, T. Ollmann, S. R. Bulò, and P. Kontschieder, "The mapillary vistas dataset for semantic understanding of street scenes," in *Proceedings of the International Conference on Computer Vision (ICCV), Venice, Italy, 2017*, pp. 22–29.
- [13] J. M. Coughlan and H. Shen, "Terrain Analysis for Blind Wheelchair Users: Computer Vision Algorithms for Finding Curbs and other Negative Obstacles." in *CVHI*, 2007.
- [14] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [15] Z. Liu, S. Shabani, N. G. Balet, M. Sokhn, and F. Cretton, "How to motivate participation and improve quality of crowdsourcing when building accessibility maps," in *2018 15th IEEE Annual Consumer Communications Networking Conference (CCNC)*, Jan 2018, pp. 1–6.
- [16] A. Kovashka, O. Russakovsky, L. Fei-Fei, K. Grauman *et al.*, "Crowdsourcing in computer vision," *Foundations and Trends® in Computer Graphics and Vision*, vol. 10, no. 3, pp. 177–243, 2016.
- [17] K. Hara, V. Le, and J. Froehlich, "A Feasibility Study of Crowdsourcing and Google Street View to Determine Sidewalk Accessibility," in *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, ser. ASSETS '12. New York, NY, USA: ACM, 2012, pp. 273–274.
- [18] K. Hara and J. E. Froehlich, "Characterizing and Visualizing Physical World Accessibility at Scale Using Crowdsourcing, Computer Vision, and Machine Learning," *SIGACCESS Access. Comput.*, no. 113, pp. 13–21, Nov. 2015.
- [19] J. Ingensand, S. Composto, O. Ertz, D. Rappo, M. Nappetz, T. Produit, M. Oberson, I. Widmer, and S. Joost, "Keys to successful scientific vgi projects," PeerJ Preprints, Tech. Rep., 2016.
- [20] M. Robinson, A. Fischer, A. Ahuja, E. Lesser, and H. Maniates, "Roles of "wanting" and "liking" in motivating behavior: gambling, food, and drug addictions," in *Behavioral neuroscience of motivation*. Springer, 2015, pp. 105–136.
- [21] Z. Zeng, J. Tang, and T. Wang, "Motivation mechanism of gamification in crowdsourcing projects," *International Journal of Crowd Science*, vol. 1, no. 1, pp. 71–82, 2017.
- [22] Y. Chou, *Actionable Gamification: Beyond Points, Badges, and Leaderboards*. Createspace Independent Publishing Platform, 2015.