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

Anonymous VCIP Submission
Paper ID:

***Abstract*—Availability of global and scalable tools to assess disabled pedestrian level of service (DPLoS) is a real need but still a challenge in today's world. It is usually summarized by lacking of tools that can ease the measurement of a level of service adapted to disabled people, and the limitation concerns the availability of information regarding the existing level of service in real time as well. This paper aims to use advanced computer vision technologies and benefits from the prevalence of handheld devices in order to respond to those needs. Our approach allows the development of a navigation tool with crowdsourcing technologies that can help a disabled person to move around a city, and suggest the most adapted routes according to the person's disabilities. This solution provides the opportunity to get the up-to-date data with valuable field observations.**

I. Introduction

Creating the disabled pedestrian map-based routing is a challenge for both researchers and practitioners. Disabled pedestrian level of service (DPLoS) have been studied and discussed by more and more people. One of the most important reason is maps-based routing solution provides information about the available facilities of transitory obstacles to enable people with mobility disabilities to travel more easily. In fact, the demand for disabled pedestrian level of services is enormous. According to World Health Organization [1], over a billion people live with some form of disability. This correspond to about 15% of the world's population, and between 110 and 190 million adults have very significant difficulties in functions that include the wheelchair users. The interdisciplinary researches are working together 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 in replicable, objective, cost-effective systems to assess pedestrian infrastructure [2]. Therefore, two main limitations have been identified. On one hand, there is a lack of tools that can ease the measurement of a level of service (LoS) adapted to disabled people. These tools can help assess the level of accessibility of a pedestrian network. On the other hand, the missing of the availability of information regarding the existing level of service in real time causes problems. When moving through a network, a disabled person needs this kind of information to take a right decision.

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. This approach allows the development of a navigation tool with crowdsourcing technologies that can help a disabled person to move around a city suggesting the most adapted routes according to the person's disabilities. As illustrated in Figure 1 , the best route in our approach is based on the highest level of service of route and adapted to the user's handicap. Specifically, we give people with mobility impairments, mainly wheelchair users, a mean to travel safely and with ease through a modern urban area by applying the algorithms of computer vision. These algorithms use crowdsourced images taken by the users of a mobile application in order to determine a level of service adapted to disabled pedestrians.

This paper is organized as follows: section 2 presents a summary of the related works of 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 by applying computer vision and machine learning. 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*

Jean-Christophe Loubier

As sidewalks are one of the most important and secure places for pedestrian to walk on in urban areas, several studies have focused their researches on characterizing and assessing sidewalks accessibility, even based on percieved security from the pedestrian point of view [3]. But it's rather new that a PLoS has been specifically investigated for disabled people (DPLoS). Asadi et al. [4] have for example calculated 10 main feature indicators based on several studies and guidelines regarding the presence of facilities for disabled, like wheelchair users or blind people.

B. *Computer Vision based techniques to asses relevant street features for impaired*

Historically, computer vision tends to extract the most important features from images like color histograms, edges [5] or specific patterns. Hough transform is a well know line detector that may be used to extract street borders [6] as well as vanishing point extraction.

The slope of a street is an important feature for impaired. Making use of geolocation with odometric data – mainly GPS trajectory – Lu and Karimi [7] showed a simple approach to compute this value based on images taken perpendicular to the street axis.

Steroscopic vision provides at low cost the useful depth information as a disparity map. This is done by computing a 3D point cloud from stereo images pairs, which can help to better isolate relevant objects. Coughlan and Shen [9] make use of an embeded stereo camera to detect negative obstacles on the wheelchair user's path.

Hara et al. [10] show how computer vision may be used to determine accessibility from physical features of the real world based on Google Street View images coupled with machine learning techniques. Within the latter domain, convolutional neural networks (CNN) are especially designed to extract features from images [11]. Based on these techniques, field of autonomous driving also brings a solid background for street images segmentation [12]

Existing street scene datasets along with semantic segmentation and region-based methods [13] have proven the need of fine grained labeled images to suit a particular task. Using weights of low level features extracted by some CNN on these datasets may be a good starting point to build a more specific architecture dedicated to the DPLoS features extraction.

But there are still two main bottlenecks to build 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 amount of data is available, they may not be well enough labeled nor suited to a given specific task, hence well managed crowdsourcing may provide an invaluable help in both.

C. Crowdsourcing in disabled pedestrian level of service

Research on disabled pedestrian level of service (DPLoS) is relatively new. The analytical methods to estimate the DPLoS is not only considered as a narrow range of pedestrians, but also need to apply more diverse pedestrian populations with different characteristics, to ensure inclusive walking conditions. Therefore, a good quality of crowdsourced data from citizens is the key element to create the DPLoS. Liu et al. [14] identified three useful techniques for improving crowdsourced data quality when building accessibility maps, including qualification tests, reputation system, and aggregation techniques. Their findings highlighted that the accuracy rate has a significant increase after applying the intervention of quality control methods. Moreover, [15] and [16] shown how participatory Geographic Information System (GIS) using Free and Open Source Software (FOSS) may help provide high quality open-GIS data and make them available to the public using web technologies.

Other studies [17], [18] focused on mapping the level of accessibility of street networks using an end-user design approach, while the OpenSideWalk project also [19] investigate routing techniques for sidewalk graph analysis in order to propose the better paths for disabled people. The latter is now proposing a schema to better and more finely represents sidewalks as separate footways from the drivable roads in OpenStreetMap (OSM).

In this study, we follow the guidelines from literatures and apply selective recruitment and training of participants to provide the high-quality crowdsourcing data in addition to computer vision based techniques.

D. Coupling crowdsourcing and computer vision to evaluate the DPLoS model

Computer vision is of most interest in field of autonomous driving and objects detection but features extraction relies frequently on domain experts. Integrating crowdsourcing is not an easy task. A recent survey by Kovashka et al. [20] addressed some of the most important questions regarding computer vision and crowdsourcing and provide useful general guidelines. But only a few studies took advantages of both techniques for DPLoS assessment or application design. Hara and Froehlich [21] have shown the opportunity of integrating final users in the design of an application devoted to disabled pedestrian and to assess scaled street segments aptitude.

More recently, Hara et al. [10], [22] designed a specific tool which makes use of Amazon Mechanical Turk (AMT) to achieve rapidly and at low cost some detection tasks on images. Recall of AMT workers combined with computer vision algorithms is higher and faster than AMT workers alone. Both together also provide a better accuracy on detection than algorithms alone.

The drawbacks in such approach are data acquisition and labeling. The latter is one of the highest time and knowledge demanding operation. This is why it can lead to noisy results. Therefore, great care and efforts must be taken to acquaire high quality data while keeping crowdworkers motivation [23].

III. Proposed Approach

Based on the relevant literature, the research proposal is articulated in three main phases:

A) the creation of the DPLoS model
B) the establishment of a pedestrian network at a city-scale level
C) the design of a computer vision and machine learning architecture dedicated to features detections, coupled with crowdsourcing applications to collect data, digitize and label features and evaluate the detection results

Development of CV and ML algorithms have to be closely related to crowdsourcing application design because some of the user inputs may influence detection algorithms on one hand and on the other, detection results will have to be evaluated by crowd-workers and final users.

A. DPLoS model

Jean-Christophe Loubier

....

The DPLoS established model retains the 5 following features based on literature guidelines:

- slope
- sidewalks width and presence of obstacles
- coating quality (e.g. pavement, smooth concreet)

- crosswalks and security signals (e.g. pedestrian light, refuge island)
- presence of curb ramps to cross the road

A normalized score S_i for each feature i is then weighted to compute a global DPLoS score for any street edge as eq. 1 shows.

$$\frac{1}{5}\sum_{i=1}^{5} S_i \cdot w_i \tag{1}$$

Weights w_i have to be determined based on local conditions and street users perception.

B. Pedestrian graph as an input

DPLoS computation for street edges requires a pedestrian graph as an input. In order to build this graph, one will rely on existing data such as National Map Agency (NMA) Cadastral Survey or vector data if available, OSM streets and footways or third party road networks.

It is also conceivable to use high resolution satellite data to extract a road network using segmentation-based classification methods in regions where high quality vector data does not exist.

For this study, OSM street pathways were extracted over a European small town (fig. 1) using a dedicated tool [24].

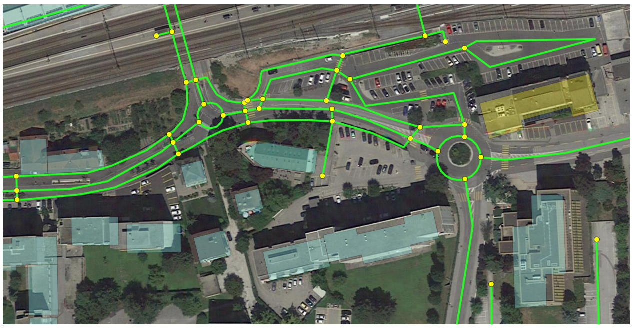

Figure 1. Part of the pedestrian network extracted from OpenStreetMap (OSM) data (light) in a downtown area

C. Coupling algorithms and crowdsourcing

It was decided to use one different tool for each of the four identified goals. These are, by logical order:

1) Footways network update
2) Data acquisition: raw images + geolocation
3) Images features digitizing and labeling
4) End-user route choice and update abilities

1) Footways network update: The initial pedestrian network may not be totally fulfilled (fig. 1). A tool will be proposed to give crowd-workers a pleasant way to complete or make it fit the actual reality more precisely based on a recent orthophoto background.

2) Data acquisition and geolocation: The second application, would serve as a feeding source for raw images.

Everyone can take a picture with a handheld mobile device and upload it to a database. Along with existing street pictures sources like Google Street View these images will be used as raw data input for the digitizing stage. If position and orientation (azimuth) values are know from a GPS and compass, they are stored along the picture. Otherwise, a geolocation tool will propose to the user to geolocate and orient its photograph before uploading it.

3) Feature digitizing and characterizing: The main goal of the third stage is to detect one or more of the 5 identified DPLoS features on raw images and marking them with a score based on the DPLoS model. If the former may be done by a well trained CNN (fig. 2), the latter would be evaluated by crowd-workers.

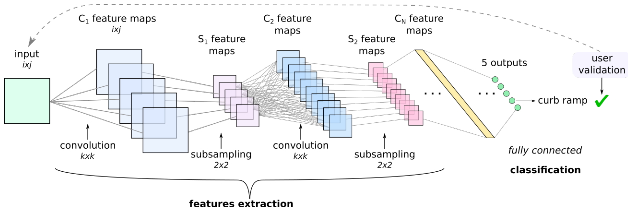

Figure 2. A CNN detecting a curb ramp. User checking may disable the label associated with this image if the result is a false positive.

E.g., if someone sees a curb ramp on a picture, he will be able to digitize it more accurately and marking it as a more or less friendly curb ramp (fig. 3). An other example would let the user annotate a sidewalk surface width as "no possible way for a single wheelchair" based on the perceived width on the given picture. If depth information is also available for the given image, a measuring tool may also be provided.

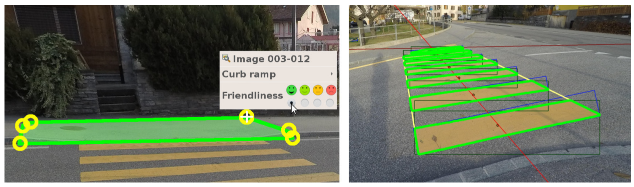

Figure 3. Image examples of a manually digitized curb ramp with a friendliness score selection toolbox (left) and a true positive automatically detected crosswalk (right).

This stage will also give the user a tool to check the features that were detected by algorithms. 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 positive and negative labeled images. Training and testing results will be tracked to avoid overfitting in addition to some optimizations.

4) End-user route choice and update: The final application will directly target the impaired pedestrian street users. It will propose 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 define by eq. 2.

$$\text{DPLoS}_{\,route} = \frac{1}{n}\sum \text{DPLoS}_{\,edge\ \ n} \tag{2}$$

Informations about the edge as well as an update button (fig. 4) will offer the user the possibility to give its feedback on a selected segment, for example if he sees an inadequate between the real world and a digitized feature.

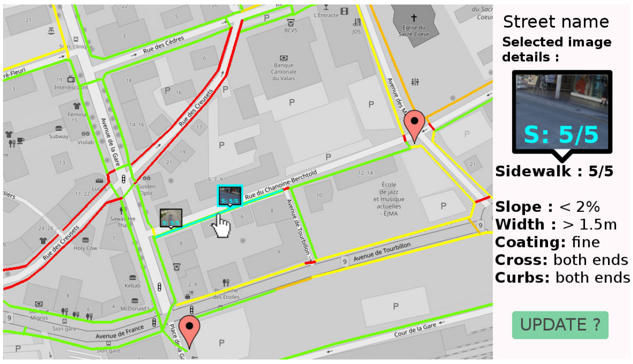

Figure 4. Prototype of the final application with edges of the pedestrian graph colored according to their DPLoS value (red=low value, green=high value). The selected edge shows that two different images participated to its score. Features' details for this edge and the selected image are shown on the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

This will continuously improve the pedestrian network and update it as urban environment is changing.

IV. Conclusion and Future Work

This project is in its early stages. The first version of the DPLoS model is ready. The setting up of a database to store images metadata like geolocation, as well as algorithms and crowd-workers labeling results is taking place. This same database will serve the four different applications through APIs that need to be developed. It will finally need further investigating and tests regarding coupling of CV and ML algorithms. Database populating and labeling will be the most resource demanding 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] 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.

[4] 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.

[5] 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.

[6] N. Kiryati, Y. Eldar, and A. M. Bruckstein, "A probabilistic Hough transform," *Pattern Recognition*, vol. 24, no. 4, pp. 303–316, Jan. 1991.

[7] Y. Lu and H. A. Karimi, "Real-Time Sidewalk Slope Calculation through Integration of GPS Trajectory and Image Data to Assist People with Disabilities in Navigation," *ISPRS International Journal of Geo-Information*, vol. 4, no. 2, pp. 741–753, 2015.

[8] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] M. A. Brovelli, M. Minghini, and G. Zamboni, "Web–based Participatory GIS with data collection on the field – A prototype architecture," *OSGeo Journal*, vol. 13, no. 1, pp. 29–33, Apr. 2014.

[16] M. Minghini, "Multi-dimensional GeoWeb platforms for citizen science and civic engagement applications," Tesi di dottorato, Mar. 2014.

[17] S. Comai, D. Kayange, R. Mangiarotti, M. Matteucci, S. Ugur Yavuz, and F. Valentini, "Mapping city accessibility: review and analysis," *Stud. Health Technol. Inform*, vol. 217, pp. 325–331, 2015.

[18] A. Gharebaghi, M.-A. Mostafavi, G. Edwards, P. Fougeyrollas, P. Morales-Coayla, F. Routhier, J. Leblond, and L. Noreau, "A confidence-based approach for the assessment of accessibility of pedestrian network for manual wheelchair users," in *International Cartographic Conference*. Springer, 2017, pp. 463–477.

[19] "OpenSidewalks."

[20] 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.

[21] 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.

[22] 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.

[23] L. Von Ahn, R. Liu, and M. Blum, "Peekaboom: a game for locating objects in images," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 55–64.

[24] G. Boeing, "OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks," *Computers, Environment and Urban Systems*, vol. 65, pp. 126–139, Sep. 2017, arXiv: 1611.01890.