Real-time Powered Wheelchair Assistive Navigation System Based on Intelligent Semantic Segmentation for Visually Impaired Users

Elhassan Mohamed
School of Engineering
University of Kent
Canterbury, Kent, UK
enrm4@kent.ac.uk

Konstantinos Sirlantzis
School of Engineering
University of Kent
Canterbury, Kent, UK
k.sirlantzis@kent.ac.uk

Gareth Howells
School of Computing
University of Kent
Canterbury, Kent, UK
w.g.j.howells@kent.ac.uk

***Abstract*—People with movement disabilities may find powered wheelchair driving a challenging task due to their comorbidities. Certain visually impaired persons with mobility disabilities are not prescribed a powered wheelchair because of their sight condition. However, powered wheelchairs are essential to the majority of these disabled users for commuting and social interaction. It is vital for their independence and wellbeing. In this paper, we propose to use a semantic segmentation (SS) system based on deep learning algorithms to provide environmental cues and information to visually impaired wheelchair users to aid with the navigation process. The system classifies the objects of the indoor environment and presents the annotated output on a display customised to the user's condition. The user can select a target object, for which the system can display the estimated distance from the current position of the wheelchair. The system runs in real-time, using a depth camera installed on the wheelchair, and it displays the scene in front of the wheelchair with every pixel annotated with distinguishable colour to represent the different components of the environment along with the distance to the target object. Our system has been designed, implemented and deployed on a real powered wheelchair for practical evaluation. The proposed system helped the users to estimate more accurately the distance to the target objects with a relative error of 19.8% and 18.4% for the conditions of a) semi-neglect and b) short-sightedness, respectively, compared to errors of 47.8% and 5.6% without the SS system. In our experiments, healthy participants were put in simulated conditions representing the above visual impairments using instruments commonly used in medical research for this purpose. Finally, our system helps to visualise, on the display, hidden areas of the environment and blind spots that visually impaired users would not be able to see without it.**

***Index Terms*—Convolutional neural networks, Practical implementation, Semantic segmentation systems, Visually impaired disabled users**

I. Introduction

Powered wheelchairs are essential tools for disabled users not only for the main reason, which is mobility but also for other factors such as productivity, leisure, and independence are involved [1]. However, disabled users with visual impairment cannot be prescribed a powered wheelchair because of their disability. Clinicians recommend against using powered wheelchairs for this category of users because visual impairment can result in serious accidents. Many injuries have been reported for users hurting themselves or falling from powered wheelchairs as they could not distinguish between pavement edges and car routes or walls and doors [2]–[4].

Clinicians report that they saw almost the same number of patients who cannot use a powered wheelchair as those who can [5]. Patients find it extremely difficult to manoeuvre powered wheelchairs indoors, especially in small areas and while negotiating doorways. Clinicians also report that 40% of powered wheelchair users find steering tasks difficult. At the same time, five to nine per cent find them impossible. On the other hand, the percentage of those who cannot use a powered wheelchair due to visual impairment, cognitive disorder or motor skills is 85%. An automated navigation system is believed to half this percentage [5].

We propose to use a smart system based on deep learning for computer vision tasks to act as a guide for the users. The system does not interfere with the navigation process (a non-intrusive system). However, it leads the navigation process of the user by classifying the surrounding environment into different classes, where every class has a distinguishable colour to guide and smooth the user's navigation process.

In addition, the system provides the distance to the target object. The distance estimation and the annotated environmental objects are shown on a display fitted according to the disabled user's need, where the information is shown and updated in real-time. If the user has short-sightedness, the display is fitted in front of the user to see far objects annotated clearly on the display. On the other hand, if the user is suffering from semi-neglect, the display is fitted in the visible location with respect to the user's condition (either centre-left or centre-right).

This paper presents the practical implementation of the proposed semantic segmentation (SS) system for safe navigation in the indoor environment. The contribution of this paper is twofold: we practically deployed the semantic segmentation system on a real powered wheelchair for the purpose of evaluation. Besides, healthy users are asked to try the system while simulating visual impairment to assess the system performance and its ability to help disabled users.

The paper is organised as follows: section 2 explores related studies. Section 3 presents the system design. Experimental protocol is presented in section 4. Results are discussed in section 5. Conclusion are highlighted in section 6.

II. Background studies

Many powered wheelchair systems have been introduced to help disabled users with navigation [6]–[8]. However, these systems take the control away from the user, such as autonomous obstacle avoidance systems where surrounding object is categorised as an obstacle that needs to be avoided. This might impact the user experience when trying to approach a specific location or an object that the smart system has classified as an obstacle. Also, user preference might be overridden if the algorithm decides on a specific path to the target object because it is the optimum one, while the user wants to follow a different path. The only solution, in this case, is to disable the system to be able to approach the object or follow a preferred path.

Yoder et al. [9] propose a wheelchair system to guide people with severe disabilities. The system is used to track manually taught paths using optical encoders mounted on the wheelchairs' wheels and visual beacons placed throughout the wheelchair surrounding environment. The main disadvantages of such a system are as follows: it needs visual cues to be deployed in the wheelchair environment. It needs manually taught reference paths. Most importantly, the system needs different setups for different environments. This means that if the environment changes, new reference paths are needed to be learned by the system.

In contrast to the fully autonomous systems that take full control away from the users, the proposed method provides environmental cues to aid the user with navigation. However, it keeps the users in full control. Powered wheelchair systems that provide collision avoidance support such as [10], [11] may not be suitable for drivers who are unable to determine their location and cannot navigate to a specific location, especially visually impaired users.

Although the proposed system is not used for path tracking, it is capable of detecting visual cues automatically using deep learning methods, which means no need to add visual beacons to the environment. Moreover, the proposed system can detect the distance to a specific object using the Intel®RealSense depth camera. Also, our system provides all the information to the user on a display. This allows visually impaired users to understand their surroundings and approach the target object with the aid of the distance estimation capability. Thus, the used system can be seen as an in-between system that can provide environmental cues (scene understanding) while giving full control to the users to avoid both limitations of non-autonomous and fully autonomous systems. However, it can be integrated with autonomous ones, and the users can decide the level of assistance.

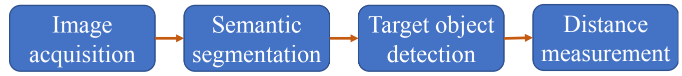

Fig. 1: System building blocks.

III. System design and implementation

The proposed system comprises a depth camera for image acquisition and distance estimation followed by a semantic segmentation system for pixels classification and a post-processing step for distance estimation to the target object. Fig 1 shows the main building blocks of the proposed system.

The system works as follows: the depth camera capture images from the indoor environment. Then the semantic segmentation system processes the images by classifying each pixel into one of the nine predefined classes. Finally, the system estimates the distance to the pixel in the centre of gravity of the target object.

The user can choose the target object to approach. However, an image may contain several objects of the same category. The system is designed to estimate the distance to one object at a time to avoid confusion. Thus, we exclude far objects depending on the size since far objects are smaller than near ones. Also, some pixels can be classified wrongly to the target object category. Consequently, these pixels can confuse the system. Applying a pixel score threshold value to the pixels' confidence scores can elevate this issue [12].

Another challenge for distance estimation is to determine a specific pixel on the target object that is not on its borders. Indoor objects of interest have irregular shapes. Consequently, the pixel in the object's centre of gravity is the best choice. The depth camera can then estimate the target object's distance with respect to the powered wheelchair. The user can monitor the real-time update of the distance on a display customised to the user's need while approaching the highlighted target object. Besides, other components of the indoor environment are semantically segmented to provide guiding cues to the user. This facilitates the navigation process of visually impaired users who cannot estimate the target object's distance or location without the proposed system due to their disability.

In this study, the SS system introduced in [13] is adopted. The SS system is based on DeepLabv3+ [14] to classify indoor objects such as doors, fire extinguishers, different kinds of door handles, switches, and walls. Basically, the objects a powered wheelchair user needs to interact with or manoeuvre in a daily life base [15]. The encoder-decoder SS system has proved high performance and accurate segmentation compared to other state-of-the-art SS systems [12]. The system is deployed on a laptop with a GeForce RTX 2070 Graphics Card. The Laptop is connected to an Intel®RealSense depth camera to estimate the distance to the target object. A separate display is fitted according to the user's need to display the objects' annotations and the estimated distance to the target object.

(a) Glasses to simulate users with semi-neglect.

(b) Glasses to simulate users with short-sightedness.

Fig. 2: Visual impairment simulation.

IV. Experimental protocol

To assess the system, we conduct the experiments with healthy users. We choose to conduct the initial trials with healthy users to avoid any hazards that might impact the disabled user's experience and collect initial results until we receive the ethical approval to experiment with visually impaired disabled users. However, the used techniques to simulate visual impairments are robust and have been approved by institutions such as the university of Cambridge [16], [17].

Visual impairment is simulated using two glasses (Fig. 2) to mimic typical use-case scenarios. The first pair of glasses has a single covered lens to simulate users with semi-neglect (Fig.e 2a). The second pair of glasses are used to introduce vision acuity and contrast loss which simulate users with short-sightedness (Fig. 2b). The Cambridge simulation glasses[1] used in the experiments provide insights into the effects of vision loss on product use. Also, the glasses can be used to examine the visual accessibility of products and services. Acuity and contrast loss glasses have five levels. In the experiments, level three is used as a median of the five levels. Ten users are asked to drive a powered wheelchair in a controlled environment with one objective. The goal is to approach a moveable door handle. The target moveable door handle is approximately 12 m away from the starting point. Fig 3 shows a simple sketch of the experiment route. Participants need to stop the powered wheelchair half a meter (50 cm) from the target object with and without the SS system. Then, the actual distance is measured from the camera to the object using a measuring tape. Lastly, the users need to continue driving for another 12 m, rotate in place, and return to the starting point.

Initially, the user is given a couple of minutes to try the powered wheelchair without any objective to familiarise and understand the controlling process. Then, each user needs to repeat the experiment five times. First, users need to approach the target object with their bare eyes. Second, users need to approach the target object using the one-eye-covered lens glasses with and without the SS system. Third, the user needs to use the acuity and contrast loss glasses to approach the target with and without the SS system. The SS is used in the second and third trials, where the user is asked to drive the powered wheelchair to the target object depending only on the display. Fig 4a shows the system setup used in the trials. The display and the SS system are not used in the first trial. The display is customised to the user's needs (Fig 4b).

[1] http://www.inclusivedesigntoolkit.com/csg/csg.html

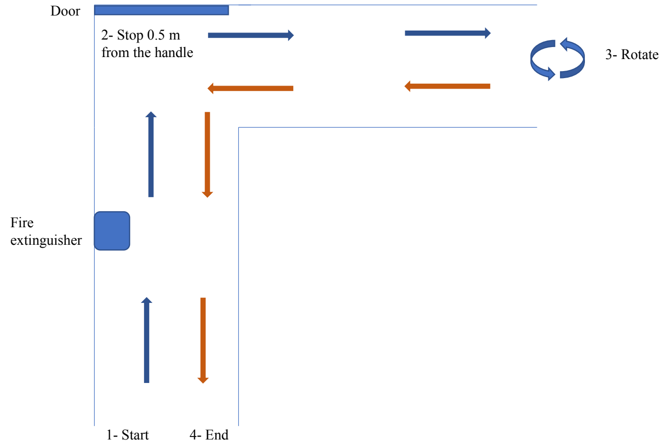

Fig. 3: Experiment route.

As an example, if the user is using the one-eye-covered lens glasses to simulate semi-neglect disorder on the right side of the environment, the display is fitted on the user's left side and vice versa. The display is placed 30 cm from the user's face at all times. The display shows the indoor environment objects classified on the pixel level where each object is annotated with a different colour. Also, the system shows the distance to the target object in the yellow box (Fig 4c).

Users' information, such as user ID, age, powered wheelchair driving experience, and use of glasses, are recorded prior to the experiments. During the experiments, trial time, the actual distance to the target object, and the number of collisions are recorded. Users' details are shown in Table I.

TABLE I: Users' details.

User ID	Age	Gender	Height (cm)	Weight (kg)	Use of glasses	Exp*
1	40	M	170	95	Yes	1
2	22	F	160	48	No	1
3	24	F	170	70	No	1
4	23	F	164	52	No	1
5	30	M	178	75	No	4
6	32	M	175	100	No	1
7	29	M	178	61	No	1
8	26	M	180	80	Yes	1
9	32	M	180	78	No	1
10	28	M	187	115	Yes	2

* Powered wheelchair driving experience on a scale from 1 to 5, where 1 represents the least experience.

After the experiments, users are asked three questions: Did you feel any stress or screen fatigue? Does the SS system help in the navigation process, and How? Lastly, if they have any comments or concerns?

V. Results and discussion

Table II shows the experiment results for all users. User 3 could not participate in the second experiment. It can be seen that the duration of all experiments under different conditions for all users is very similar, which means that the SS system

(a) System installation. (b) System in use. (c) Distance to the target object.

Fig. 4: Practical implementation of the system.

does not slow down the navigation process. There is a slight increase in collisions when simulating visual impairment while using the SS system. All the collisions happen when the users rotate before going back to the starting point. The slight increase in the total number of collisions while simulation visual impairment conditions (4 and 3 collisions for semi-neglect and short-sightedness, respectively, with the SS system compared to 2 and 3 collisions without the SS system) can be attributed to the restricted view shown on the display. While users can rotate the head to check all the sides of the powered wheelchair, it is challenging to check the sides and the back of the powered wheelchair from the display. Also, all users except user 5 have limited experience in driving powered wheelchairs.

Interestingly, the SS system has helped the users to approach the target object accurately. While simulating the semi-neglect condition, users could approach the target object with a mean distance of 59.9 cm with the aid of the SS system compared to 73.9 cm without the SS system. For the short-sightedness case, the mean distance using the SS system is 59.2 cm compared to 52.8 cm without the SS system. The estimated distance of the short-sightedness case without the SS system is better than that with the SS system and even better than the estimated distance in the normal case. It was observed that the users kept approaching the target object until it becomes distinguishable with the acuity and contrast loss glasses, which results in better distance estimation. The consistency in the results of the SS system is an indication of the system's robustness. In addition, the SS system results are significantly better than the normal case without any glasses to simulate visual impairment, which indicates the system's efficiency in estimating the accurate distance to the target object.

Fig 5 shows the impact of applying different threshold values to the pixels' confidence scores on the accuracy of the SS system. As the threshold value increases, the number of the rejected pixels increases, which results in discarding low confidence noisy pixels. For a robust system, low confidence pixels represent objects of non-interest. Rejecting these pixels can result in a better definition of object shapes. However, high values of threshold might result in discarding pixels of interest. A threshold of 0.7 is used in the experiments to achieve better object boundary definition without discarding pixels of interest.

The post-experiment interviews reveal some interesting observations:

Simulation of semi-neglect condition

- 50% of the users felt some level of stress when they have one eye covered, which is normal as they are not used to that situation. As a result, they could not estimate the distance correctly without the SS system. The one-eye-covered lens glasses have also introduced limited vision.
- 80% of the users found the SS system with the display helpful in estimating the distance because the system displayed the distance to the object in real-time and highlighted the target object. In addition, the system helped the users to see the whole scene on the display, especially the blind spots due to the usage of the one-eye-covered lens glasses.
- User 8 and 9 suggested to use a high-resolution display and a faster processing rate to achieve fast estimations.
- User 2 suggested that the display's limited vision in tight places may result in collisions.
- Without the SS system, user 10 was worried about hitting any object on the covered side by the glasses. However, the SS system greatly helps by showing these blind spots and highlighting the components of the environment in different colours. In contrast, user 9 suggests highlighting only the target object and not all the environmental cues.

Simulation of short-sightedness condition

- 40% of the users felt stress because of the blurred and foggy vision of the acuity and contrast loss glasses.
- 8 out of 10 users found the system helpful in understanding the surroundings and estimating the distance to the target object. One user said the experience was easy with and without the system because the path became familiar (User 4). In contrast, user 9 said the system was not helpful for distance estimation because the refresh rate is slow but was helpful for segmentation.
- 6 out of 10 users found the display hard to use while turning due to the limited vision of the sides and back of the powered wheelchair.
- User 1 found the system very helpful in understanding the surroundings and estimating the distance to the target

TABLE II: Experimental results.

User ID	Normal vision			Semi-neglect						Short-sightedness					
				Without SS system			With SS system			Without SS system			With SS system		
	T	C	D	T	C	D	T	C	D	T	C	D	T	C	D
1	2:13	1	80	2:00	1	64	2:24	0	53	2:12	3	62	2:18	1	60
2	2:13	0	60	2:10	0	75	2:27	1	50	1:58	0	41	1:53	0	28
3	2:08	1	80	1:54	0	75	2:11	1	45	-	-	-	-	-	-
4	2:05	0	90	1:54	0	92	2:13	0	73	1:49	0	65	1:52	0	75
5	1:45	0	75	1:42	0	70	1:44	0	66	1:45	0	62	1:39	0	55
6	2:25	0	88	2:07	0	103	2:21	1	80	1:52	0	83	2:03	1	64
7	2:00	0	41	1:49	0	62	1:57	0	60	1:44	0	46	1:49	0	66
8	2:01	0	92	2:05	0	67	2:14	0	42	1:42	0	33	1:51	0	60
9	1:50	0	81	1:45	0	70	2:00	0	55	1:51	0	36	1:46	1	60
10	1:47	0	77	1:59	1	61	1:56	1	75	1:41	0	48	1:51	0	65
Mean	2:02	0.2	76.4	1:56	0.2	73.9	2:08	0.4	59.9	1:52	0.3	52.8	1:56	0.33	59.2
Std	± 12	± 0.42	± 15.3	± 9	± 0.42	± 13.5	± 13	± 0.51	± 13.1	± 11	± 1	± 16.2	± 12	± 0.5	± 12.9

T : Time (min : sec), C : No# of collisions, D : Actual distance to the target (cm), Std: Standard deviation in seconds for time measurement.

object. With the blurred vision simulation, the user can recognise the door but not the handle, the wall but not the fire extinguisher. However, the SS system can give accurate information and allow the recognition of the door handle and the fire extinguisher.

- Three users highlighted that the distance estimation would disappear from the display if you closely approach the target object. This is understandable because the camera has a theoretical minimum range of 28 cm.

Qualitatively, the SS system helped the users to visualise blind spots covered by visual impairment simulation glasses. Also, it allows the users to better comprehend the surrounding environment.

Quantitatively, the SS system helped the users to better estimate the distance to the target object with a relative error of 19.8% and 18.4% during the simulation of semi-neglect and short-sightedness conditions, respectively, compared to a relative error of 47.8% and 5.6% without the SS system. The overall all system speed is approximately 3 FPS. Thus, the post-processing steps (discarding far objects and finding the pixel in the centre of gravity) have a minor impact on the system speed. However, they are important to achieve an accurate segmentation.

VI. Conclusion

This paper presents a practical implementation of a semantic segmentation system for pixel classification and distance estimation to enhance the visually impaired disabled user's independence and wellbeing by allowing them to understand and interact with their surroundings.

It can be concluded from the quantitative and qualitative results that the proposed system can help visually impaired users to estimate the distance to the target object accurately. Besides, it can help the users to understand the surrounding environment. However, the system's display limits the vision of the powered wheelchair sides in narrow areas. This can be compensated using proximity sensors to alert the user when turning. Also, increasing the processing speed of the proposed system can enhance the overall experience of the system.

Acknowledgment

This work was supported by the Assistive Devices for empowering dis-Abled People through robotic Technologies (ADAPT) project. ADAPT was selected for funding by the INTERREG VA France (Channel) England Programme which is co-financed by the European Regional Development Fund (ERDF). The European Regional Development Fund (ERDF) is one of the main financial instruments of the European Unions (EU) cohesion policy.

References

[1] L. Jensen, "User perspectives on assistive technology: a qualitative analysis of 55 letters from citizens applying for assistive technology," *World Federation of Occupational Therapists Bulletin*, vol. 69, no. 1, pp. 42–45, 2014.

[2] A. Carlsson and J. Lundälv, "Acute injuries resulting from accidents involving powered mobility devices (PMDs)—Development and outcomes of PMD-related accidents in Sweden," *Traffic Injury Prevention*, vol. 20, no. 5, pp. 484–491, jul 2019. [Online]. Available: https://doi.org/10.1080/15389588.2019.1606910

[3] R. P. Gaal, N. Rebholtz, R. D. Hotchkiss, and P. F. Pfaelzer, "Wheelchair rider injuries: Causes and consequences for wheelchair design and selection," *Journal of Rehabilitation Research and Development*, 1997.

[4] W. Y. Chen, Y. Jang, J. D. Wang, W. N. Huang, C. C. Chang, H. F. Mao, and Y. H. Wang, "Wheelchair-related accidents: relationship with wheelchair-using behavior in active community wheelchair users," *Archives of Physical Medicine and Rehabilitation*, 2011.

[5] L. Fehr, W. E. Langbein, and S. B. Skaar, "Adequacy of power wheelchair control interfaces for persons with severe disabilities: A clinical survey," *Journal of rehabilitation research and development*, vol. 37, no. 3, pp. 353–360, 2000.

[6] R. H. Wang, A. Korotchenko, L. H. Clarke, W. B. Mortenson, and A. Mihailidis, "Power mobility with collision avoidance for older adults: User, caregiver and prescriber perspectives," *Journal of rehabilitation research and development*, vol. 50, no. 9, p. 1287, 2013.

[7] P. Viswanathan, J. J. Little, A. K. Mackworth, and A. Mihailidis, "Navigation and obstacle avoidance help (noah) for older adults with cognitive impairment: A pilot study," in *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, ser. ASSETS '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 43–50. [Online]. Available: https://doi.org/10.1145/2049536.2049546

[8] H. Grewal, A. Matthews, R. Tea, and K. George, "Lidar-based autonomous wheelchair," in *2017 IEEE Sensors Applications Symposium (SAS)*, 2017, pp. 1–6.

[9] J.-D. Yoder, E. T. Baumgartner, and S. B. Skaar, "Initial results in the development of a guidance system for a powered wheelchair," *IEEE Transactions on Rehabilitation Engineering*, vol. 4, no. 3, pp. 143–151, 1996.

[10] P. Viswanathan, J. Boger, J. Hoey, and A. Mihailidis, "A comparison of stereovision and infrared as sensors for an anti-collision powered wheelchair for older adults with cognitive impairments," in *2nd International Conference on Technology and Aging, Toronto*. Citeseer, 2007.

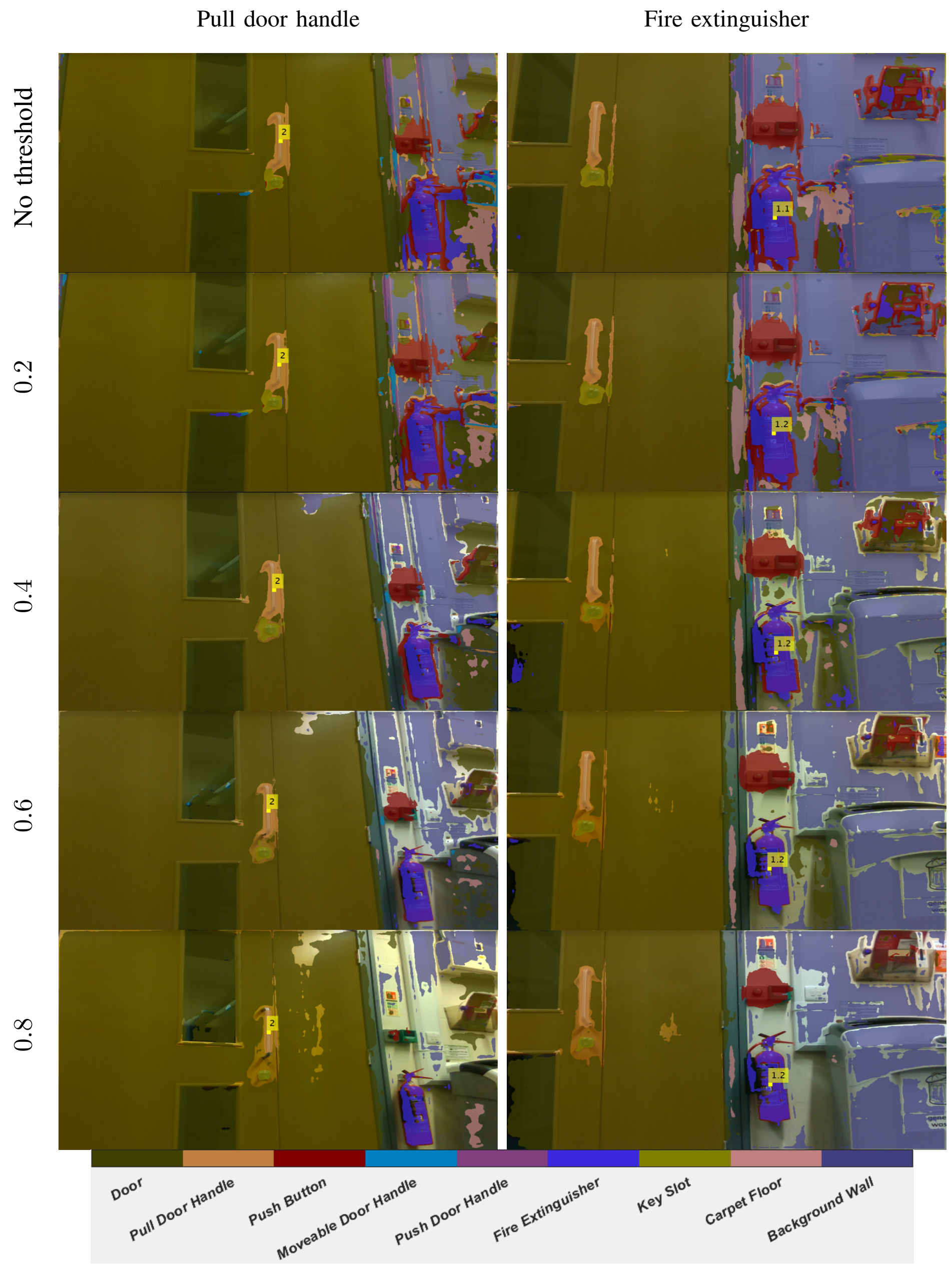

Fig. 5: Impact of applying different threshold values to the pixels' confidence scores on the quality of the segmentation. Estimated distance values are shown in the yellow boxes in meters.

[11] S. Chatzidimitriadis, P. Oprea, M. Gillham, and K. Sirlantzis, "Evaluation of 3d obstacle avoidance algorithm for smart powered wheelchairs," in *2017 seventh international conference on emerging security technologies (EST)*. IEEE, 2017, pp. 157–162.

[12] E. Mohamed, K. Sirlantzis, and G. Howells, "Incorporation of rejection criterion - a novel technique for evaluating semantic segmentation systems," in *2021 14th International Conference on Human System Interaction (HSI)*, 2021, pp. 1–7.

[13] ——, "Indoor/outdoor semantic segmentation using deep learning for visually impaired wheelchair users," *IEEE Access*, vol. 9, pp. 147 914–147 932, 2021.

[14] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.

[15] E. Mohamed, K. Sirlantzis, and G. Howells, "A pixel-wise annotated dataset of small overlooked indoor objects for semantic segmentation applications." *Data in Brief*, vol. 40, p. 107791, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352340922000038

[16] J. Goodman-Deane, S. Waller, A.-C. Collins, and J. Clarkson, "Simulating vision loss: what levels of impairment are actually represented?" in *Contemporary Ergonomics and Human Factors 2013: Proceedings of the international conference on Ergonomics & Human Factors 2013, Cambridge, UK, 15-18 April 2013*. CRC Press, 2013, p. 347.

[17] M. Zallio, S. Waller, C. Chivaran, and J. Clarkson, "Visual accessibility and inclusion. an exploratory study to understand visual accessibility in the built environment." in *SMART ACCESSIBILITY 2021 The Sixth International Conference on Universal Accessibility in the Internet of Things and Smart Environments*. ThinkMind, 2021.