Autonomous Indoor Navigation for Wheelchairs using Signboards

Ananthakrishnan D.S.
Department of Computer Science and Engineering
T.K.M. College of Engineering, Kerala
ananthakrishnands@gmail.com

Jishnu Prakash K.
Department of Computer Science and Engineering
T.K.M. College of Engineering, Kerala
jishnuastro@gmail.com

Renjith R.
Department of Computer Science and Engineering
T.K.M. College of Engineering, Kerala
renjithr201097@gmail.com

Dr. Ansamma John
Department of Computer Science and Engineering
T.K.M. College of Engineering, Kerala
ansamma.john@gmail.co

***Abstract*— Several studies have shown that people with disabilities benefit substantially from access to a means of independent mobility. While the requirements of many individuals with disabilities can be satisfied with manual or powered wheelchairs, a segment of the disabled community finds it difficult or impossible to use wheelchairs independently. This paper presents an autonomous indoor navigation system for wheelchairs based on sign board recognition. The system uses a deep learning model to detect signboards from surroundings and Azure Text Analytics API is used to extract the text from the signboard images. The system runs on a Raspberry Pi minicomputer and can be installed on any powered wheelchair. Experimental results and comparisons prove the efficiency of the proposed system.**

Keywords*— *Autonomous Indoor Navigation, Deep Learning

I. Introduction

According to the National Institute of Child Health and Human Development (NICHD), 2.2 million people in the United States depend on a wheelchair for day-to-day tasks and mobility. There is a constantly growing demand for mobility aids like wheelchairs in proportion to the senior citizen demographic. Mobility aside, people with disabilities rely on their relatives, nurses or caregivers to help them with the wheelchair. The current dependence of the disabled on their caretakers has challenged the R&D teams to develop smarter mobility aids with assistive technology aiming to return independent mobility back to the disabled.

George Klein invented the first powered wheelchair for people with quadriplegia injured in World War II while he was working as a mechanical engineer. While the needs of many individuals with disabilities can be satisfied with traditional manual or powered wheelchairs, a portion of the disabled community with visual acuity and lack of motor skills finds it arduous or impossible to use wheelchairs independently. To accommodate this population, several researchers have used technologies originally developed for mobile robots to create “smart wheelchairs”. A smart wheelchair typically consists of either a standard power wheelchair to which a computer and a collection of sensors have been added or a mobile robot base to which a seat has been attached. Recent works in smart wheelchairs include the head movement based [1], tongue-based human-machine interface [2], [3], brain-controlled wheelchairs [4] and voice-controlled wheelchairs [5]. The tongue-based control is motivated by the observation that the tongue has rich sensory and motor cortex representation. Most important evaluation factors for wheelchairs are safety and ease of operation. Providing autonomy is a way to improve both factors.

However, it is challenging to build an efficient smart wheelchair that people with all types of disabilities feel comfortable using. The system should be mountable on any model of powered wheelchair, easily removable for maintenance and travel.

In this paper we propose an autonomous indoor navigation system for wheelchair based on sign board recognition using deep learning. We utilize a digital camera that can be installed on the wheelchair to capture images of surroundings. The Kinect v1 sensor is used for obstacle detection. The input module is made into an application package interface and hence is reconfigurable supporting many input modes of operation. In driving the wheelchair, the user has to convey the destination place in any suitable input mode. The main advantage of our system is that it is more robust and increases ease of operation as very minimal human interaction is needed.

This paper is organized as follows. In Section II we discuss the related works in smart wheelchairs based on input modes and operating modes. In Section III we describe our proposed system and signboard detection model, while in Section 4 we describe the experiment details, model configuration and datasets. Results are presented in Section V and conclusions are drawn in Section VI.

II. RELATED WORKS

A physically disabled person on a wheelchair may find it difficult or impossible to maneuver a joystick based automatic wheelchair. Introduction of face gestures based semi-automatic wheelchairs may have solved the difficulties faced by a few but the navigation still required a major human effort by making chin or tongue gestures throughout the journey. While our computer vision-based control has many advantages over the existing ones, it is not the first-time computer vision technology has been adapted to smart wheelchairs. Smart Wheelchairs can be classified mainly into two categories based on Input modes and Operating modes

A. Input modes

i. *Computer Vision:* Kuno Y et al. [2] uses a system of 16 ultrasonic sensors and two video cameras to capture information from the environment. Their proposed system uses face directions to convey user's intentions to the system. Turning the wheelchair to any direction is initiated by turning the user's head intentionally to that particular direction. This system is affected by natural noisy motions where the users move their head when they do not intend to make turns. Li, H., Kutbi et.al. [1] proposed a similar system but with an egocentric camera worn by the user that collaborates with the robotic wheelchair by conveying the motion commands with head motions. Purwanto et.al. [7] set up a camera in front of a wheelchair user to capture the control information expressed through horizontal gaze direction for driving direction and eye blinking timing command for commands such as ready, backward movement and stop. The head-controlled and tongue-controlled navigation can be a feasible solution only when the user has good head or chin movement ability. The tongue-based solution also has the drawback that it interrupts the user's communication with other people. The computer vision-based methods have advantages compared to other approaches [5] and [6]. First, web cameras are low-cost devices making these systems more affordable. Second, minimal human effort is needed, thus improving the ease of operation. Third, it enables users to interact with others while operating the wheelchair.

ii. *Brain Controlled*: State of the art brain-controlled interface can even be used to monitor the user's emotional state [4], [5], such that when the user is frustrated the control unit will stop the wheelchair and wait for a new command from the user. In the other case, the control unit will continue executing the previously selected command.

iii. *Eye Movement:* In the Wheelesly robotic wheelchair system [8], the user can control the system by picking up a command icon on the CRT screen by gaze. Eye movements are measured through five electrodes. The system has the drawback that the user may move their eyeballs when they do not intend to pick a command. Usage of system can also be straining to eyes. Other works [9-11], include use of a calibration algorithm to find the direction of eye gaze in real-time

B. Operating Modes

Various operational modes are made available based on the user and the task employed, ranging from autonomous to semi-autonomous. For users with special needs, like limited ability to plan and execute a route, the autonomous system can be used to setup a controlled environment for ease of mobility given that the user spends considerable amount of time there. Some smart wheelchairs (e.g., TetraNauta, Kanazawa University [12]) function in a manner mimicking autonomous robots; the user's input is limited to his desired destination, while the smart wheelchairs plan and navigates the safest route to the target location. For this feature a complete blueprint of the area is typically required. Also, the surroundings where this wheelchair is supposed to operate must also be upgraded to facilitate its easy navigation and movement. The other class of smart wheelchairs confines their assistance to collision avoidance while the user is expected to navigate to the destination (e.g., NavChair, Tin-Man). In these cases, no prior knowledge of the surroundings is required.

Compared to the computer vision based controls [1], [6], [7], the camera setting in our system is different in that instead of having the camera installed on the wheelchair to focus on the user's face, we have a digital camera and Kinect sensor installed on the wheelchair focusing the environment in front of the user to support the vision-based control. The control is realized by tracking the signboard with the wearable camera and obstacles and floor detection with the Kinect sensor to generate motion commands. With our proposed autonomous navigation system, a person with disabilities can very easily navigate to his destination in an indoor environment with minimal human effort. The major advantage of our system is that the users can communicate with others while navigating on wheelchair and can be used by a person with limited physical movement. The modular design of the system helps to integrate additional functionalities easily.

III. PROPOSED METHOD

The proposed system consists of input module, sign board recognition system and hardware module as shown in Figure 1. The destination is provided to the wheelchair using the input module which is designed as an application package interface to the navigation system so that any modes of input can be incorporated into the system with ease. The sign board recognition system, hosted remotely on the cloud, searches for the signboard and processes the information, using image processing methods, for each frame captured by the digital camera. To detect obstacles a Microsoft Kinect v1 sensor is used. The control module coordinates the activities between each module and provides appropriate instructions to steer the wheelchair.

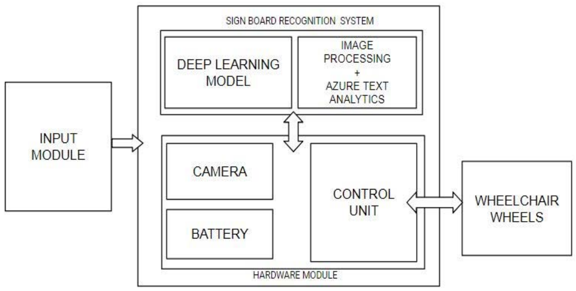

Figure 1. System Overview of Autonomous Navigation system using signboard

A. Signboard Design

For a four-way Intersection, the destinations that are straight ahead are marked on top of the sign board, and those on the left, right and in the direction opposite to the direction that the wheelchair is facing is marked in the left, right and bottom parts of the sign board respectively.

Figure 2. Design of Signboard

B. Signboard Recognition

The input of the camera stream is fed into a YOLO deep learning model for localizing the points of the sign boards in the image. YOLO v3 object detection CNN model is used to detect the Sign Boards within an image. YOLO (You Only Look Once) [13] is an object detection system which is much faster compared to other object detection methods making it suitable for real time object detection. The bounding box points obtained from the model are used to crop out the part of the image, containing the sign boards. The cropped-out image is then split into four regions. The top region is segmented from the top 30% portion of the image. Similarly, the left, right and bottom regions are segmented from the left-most 40%, the right-most 40% and the bottom-most 30% of the image respectively. These regions are further cropped out from the image and fed as input to Azure Text Analytics API for character recognition

IV. EXPERIMENT

The YOLO model used is pre-trained on the Microsoft COCO dataset containing more than 300000 images and then transfer learned on a dataset of 2859 images containing sign boards. The model was configured with a batch size of 2 images and an initial learning rate of 0.004. The dataset was collected by taking images of sign boards in different environmental and lighting conditions and at different camera angles and the images were annotated manually. The output of the signboard detection model is then processed to extract the four segments of the signboard and is given to Azure Text Analytics API for character recognition. These outputs are matched against the destination given as input to determine the direction.

V. RESULTS

For training, an HP Z230 Workstation with an Intel Core i5 CPU with a clock speed of 3.3 GHz and an Nvidia GeForce 1060 GPU supporting 8 GB of graphics memory was used. The model was trained for 22 hours with a batch size of 2 images per batch. The trained model produced a total loss of 11.1% and mean average precision of 22%.

The output of the model is then segmented into four regions and is fed to Azure text Analytics API to extract the textual information.

A comparison is drawn with other pre-trained models on the same dataset as shown in Table 1. and is found that the model pre-trained on COCO dataset model had better accuracy

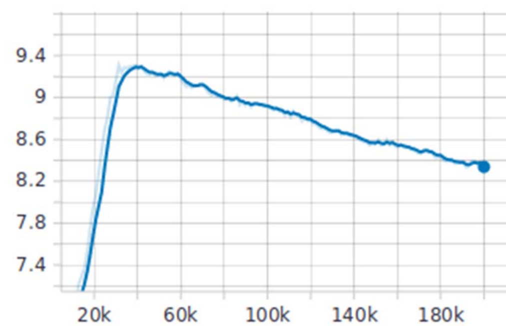

Figure 3. Classification loss of the model

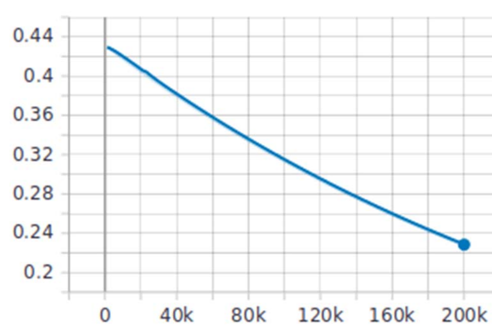

Figure 4. Regularization Loss of the model

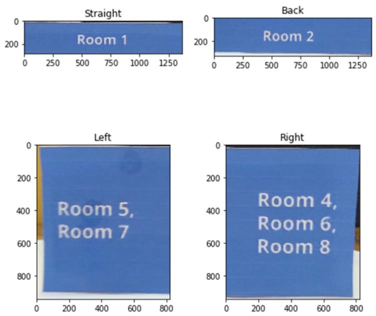

Figure 5. Segmented images of signboard

TABLE 1. COMPARISON OF RESULTS WITH OTHER MODELS

Model	***Total Loss***
YOLO SSD_MobileNet_v2_COCO	11.1
SSD_Inception_v2_COCO	15.9
Faster_RCNN_COCO	17.4

VI. CONCLUSION AND FUTURE SCOPE

This paper has proposed an indoor autonomous navigation system based on signboards. The YOLO v3 CNN model detects signboards from the surroundings and Azure Text Analytics OCR extracts the textual information from the signboard. The modular design of the system helps to integrate additional functionalities easily. Experimental results show the signboard detection CNN model has a better accuracy than other models. The system can be further enhanced by adding trackers for continuous health monitoring and can be extended to IoT applications.

REFERENCES

[1] Li, H., Kutbi, M., Li, X., Cai, C., Mordohai, P. and Hua, G., 2016, October. “An egocentric computer vision based co-robot wheelchair” In Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on (pp. 1829-1836). IEEE

[2] M. E. Lund, H. V. Christiensen, H. Caltenco, E. R. Lontis, B. Bentsen, L. N. Struijk, et al.“ Inductive tongue control of powered wheelchairs” In Engineering in Medicine and Biology Society (EMBC), 2010.

[3] M. Mace, K. Abdullah-Al-Mamun, R. Vaidyanathan, S. Wang, and L. Gupta. Real-time implementation of a non-invasive tongue-based human-robot interface. In IROS, 2010.

[4] T. Carlson and J. d. R. Millan. Brain-controlled wheelchairs: a robotic architecture.IEEE Robotics and Automation Magazine, 20(EPFL-ARTICLE-181698):65–73, 2013.

[5] M. Fezari and M. Bousbia-Salah. Speech and sensor in guiding an electric wheelchair. Automatic Control and Computer Sciences, 41(1):39–43, 2007.

[6] Kuno Y, Shimada N, Shirai Y: A robotic wheelchair based on the integration of human and environmental observations. IEEE robotic & Automation Magazine 2003, 26-34.

[7] Purwanto, D., Mardiyanto, R. and Arai, K., 2009. Electric wheelchair control with gaze direction and eye blinking. Artificial Life and Robotics, 14(3), p.397.

[8] Yanco, Holly A. "Wheelesley: A robotic wheelchair system: Indoor navigation and user interface." In Assistive technology and artificial intelligence, pp. 256-268. Springer, Berlin, Heidelberg, 1998.

[9] Chern-Sheng Lin, Chien-Waho, Wen-Chen, Chuang-Chien Chiu, Mav-Shiun Yeh,” Powered wheelchair controlled by eye tracking system”, Optica Applicata, Vol. XXXVI, no.2-3, pp. 401-412, 2006

[10] Poonam S. Gajwani, Sharda A. Chhabria,” Eye motion tracking for wheelchair control”, International Journal of Information Technology and Knowledge Management, Vol. 2, No. 2, pp. 185-187, July-December 2010

[11] Gunda Gautam, Gunda Sumanth, Karthikeyan K C, Shyam Sundar, D.Venkataraman,” Eye Movement Based Electronic Wheel Chair For Physically Challenged Persons”, International Journal Of Scientific & Technology Research, Vol. 3, Issue 2, pp. 206-212, February 2014

[12] Seki H, Kobayashi S, Kamiya Y, Hikizu M, Nomura H .Autonomous/ semi-autonomous navigation system of a wheelchair by active ultrasonic beacons. Robotics and Automation: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA); 2000 Apr 24–28; San Francisco, CA. Piscataway (NJ): IEEE; 2000. p. 1366–71.

[13] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp.779-788)