Iraqi Journal of Electrical and Electronic Engineering
Volume 15, No.2 – 2019

المجلة العراقية للهندسة الكهربائية والالكترونية
المجلد 15 العدد 2، 2019

Indoor Low Cost Assistive Device using 2D SLAM Based on LiDAR for Visually Impaired People

Heba Hakim
Computers Engineering Department
Basrah University
Basrah-Iraq
hebah.hakem@gmail.com

Ali Fadhil
Electrical Engineering Department
Basrah University
Basrah-Iraq
alifm60@gmail.com

Abstract *Many assistive devices have been developed for visually impaired (VI) person in recent years which solve the problems that face VI person in his/her daily moving. Most of researches try to solve the obstacle avoidance or navigation problem, and others focus on assisting VI person to recognize the objects in his/her surrounding environment. However, a few of them integrate both navigation and recognition capabilities in their system. According to above needs, an assistive device is presented in this paper that achieves both capabilities to aid the VI person to (1) navigate safely from his/her current location (pose) to a desired destination in unknown environment, and (2) recognize his/her surrounding objects. The proposed system consists of the low cost sensors Neato XV-11 LiDAR, ultrasonic sensor, Raspberry pi camera (CameraPi), which are hold on a white cane. Hector SLAM based on 2D LiDAR is used to construct a 2D-map of unfamiliar environment. While A* path planning algorithm generates an optimal path on the given 2D hector map. Moreover, the temporary obstacles in front of VI person are detected by an ultrasonic sensor. The recognition system based on Convolution Neural Networks (CNN) technique is implemented in this work to predict object class besides enhance the navigation system. The interaction between the VI person and an assistive system is done by audio module (speech recognition and speech synthesis). The proposed system performance has been evaluated on various real-time experiments conducted in indoor scenarios, showing the efficiency of the proposed system.*

***Index Terms*—Indoor navigation system, recognition system, visually impaired people, path planning, LiDAR.**

I. Introduction

People with healthy vision obtain their orientation in their environment, navigate from one place to another and recognize places and things with ease. Unfortunately, visually impaired and blind people face many problems in their life. These problems can be divided into three components, avoidance of obstacle, path finding and object recognition. Avoidance of obstacle deals with a local problem in walking safely without colliding into anything or falling. While path finding is a global problem of planning and following a path from one position to another till reach the goal. Therefore, many computer and electronic technologies in the market have been adopted to help VI person and increase their self-confidence. These technologies are costly and complex to be use. So, the blind/VI people have mostly been utilizes white cane due to its low cost. However, a white cane has a limited-range that cannot provide full information about its surroundings including the location information.

A survey [1] demonstrated that visually impaired people need an assistive device with: (1) low cost of technology construction; (2) simple to use; (3) enough information of their surrounding; (4) light weight; and (5) accurate. To satisfy these requirements, many research studies have been proposed in recent years [2, 3]. These researches have been implemented into two fields: navigation/guidance system or object/obstacle recognition system. A very few of them interested to implement both fields (navigation and recognition) in a single system. Merging recognition and navigation functions can improve the visually impaired person's daily navigating and perceiving. For example, a chair and a door will be understand as obstacles without a recognition technology, when in fact a chair is assisting a VI person to sit and a closed door can be open to enter the room.

According to achieve of the above needs, an assistive device that integrates recognition and navigation functions in a single system is proposed in this paper. The object recognition system allows the visually impaired person to distinguish and categorize the objects being in his/her way and also it is used to provide the location of the recognized object that is considered as a goal to a navigation system. There are many object detection methods of deep-learning that have become most popular in robotics and computer vision fields since the major success of Alex in 2012 [4]. Several researches adopted Convolution Neural Network in their works to aid the VI people to recognize obstacles and perceive their surrounding independently [5-7]. In this work, the object recognition module uses CNN based object detection technique YOLO v2 [8] because it is fast and suitable for real time use. Table I shows the comparison of YOLO v2 with the others methods of object detection.

On another hand, a navigation system is used to plan an optimal path and guide a visually impaired person movement within an unknown indoor environment. Simultaneously and location mapping (SLAM) is utilized to construct a map of the surrounding environment and localizing the VI person within a map. To acquire a 2D view of the unknown environment, Many SLAM techniques have been proposed with various sensors such as sonar sensors, laser scanners and cameras. In last years, a laser sensor has frequently been used in SLAM techniques for robotic system (robotics applications) due to its high accuracy and speed. Since these laser sensors have a high cost and most of visually impaired people with limited budget, Neato xv-11 LiDAR with low cost is used in this paper to collect the information from an indoor environment as a 2D plane. The main objective of the proposed system presented in this work is to assist VI person to understand his/her environment by recognizing the obstacles and guide him /her to final destination with accurate and low cost assistive device.

TABLE I
THE COMPARISON OF THE OBJECT DETECTION METHODS ON PASCAL VOC DATASET.

Object detection method	Accuracy mAP	Speed FPS
Fast R-CNN	70.0	0.5
Faster R-CNN VGG-16	73.2	7
Faster R-CNN Reset	76.4	5
YOLOv1	63.4	45
SDD300	74.3	46
YOLOv2 416×416	76.8	67

II. THE SYSTEM OVERVIEW

The proposed navigation system shown in Fig. 1 includes both of a Neato XV-11 LiDAR, an ultrasonic sensor and Raspberry Pi camera, are attached to a white cane. All of these sensors are controlled by Raspberry Pi 3 B+ which is a low cost, small size and low power consumption hardware platform. This micro-computer handles the data processing such as object recognition, SLAM, path planning, path following and audio processing which are implemented with python programming language. LiDAR sensor is adopted for mapping and localization module to provide an indoor 2D map. The ultrasonic sensors are utilized for detecting temporary obstacles that appear during VI person way on the optimal path. To capture a scene in front of VI person, a raspberry pi camera is utilized. Object recognition method is applied on the captured image to identify the type/coordinates of object.
The voice instructions from the VI person (e.g. to specify a desired destination) are recognized by speech recognition technology through the USB microphone while the guiding commands and recognition system output are received as voice message through earphone.

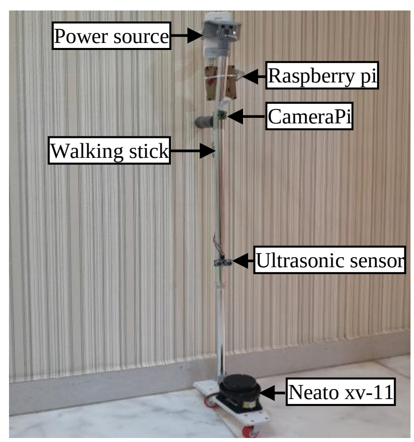


Fig. 1. The proposed system initial prototype.

III. THE SYSTEM IMPLEMENTATION

The proposed system consists of different modules which have a specific output and works together to enable the VI person navigate safely by himself/herself as shown in Fig. 2. These modules are: mapping and localization, global path planning, object avoidance and object recognition. The system in this work has two main facilities: recognition and navigation.

Recognition: firstly, the VI person tells the Raspberry pi what he/she wants to know (e.g. what is this). The microphone on Raspberry Pi records the speech and the speech recognition module converts his/her speech to text. This text is compared with other commands defined inside the configuration file of commands in order to perform the matched command which is in this case executed object recognition process. In object recognition process, CameraPi captures the image in front of the VI person and the object recognition method handles the captured image to identify the object type. Finally, the speech synthesis process converts the output result to speech to tell the VI person what he/she is acquired for. This will help VI person to perceive the surrounding environment and/or choose his/her navigation destination.

Navigation: after a VI person recognizes the objects in their surrounding and decides his/her desired goal, he/she tells Raspberry Pi through the microphone where he/she wants to go (i.e. final goal). The VI person's voice is converted to text by speech recognition module to match the goal with any of recognized object types and gets the coordinate of a recognized object as a goal. After that, SLAM algorithm based on 2D LiDAR creates a 2-D map and localizes a VI person into this map then a building map and the coordinates of current/final location are sent to global path planning module to generate an optimal path. The path following associated with object avoidance method using Ultrasonic sensors is used to follow an optimal path with avoiding a transit obstacle and yields guiding commands.

The guiding commands will be converted to audio message through earphone by speech synthesis module. The overall proposed system flowchart is shown in Fig. 3.

This assistive device guides a visually impaired person from his/her current location to the final destination in an indoor environment and works online. The following subsections describe the implementation of each module in details.

A. Mapping and localization module

Mapping and localization is implemented using 2D LiDAR-based Hector SLAM method [9]. Hector SLAM is one of the most widely used 2D SLAM approaches in robots world which is used to construct a final 2D map of the surrounding environment and at the same time the user position is estimated within a resultant map.

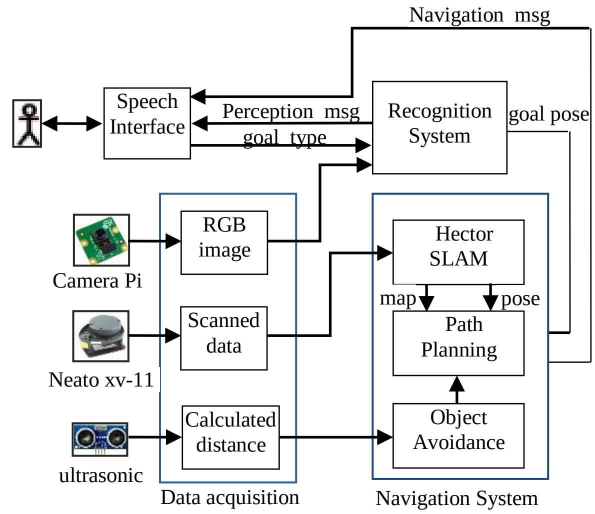


Fig. 2. The overview of the proposed system.

In comparison to the most 2D SLAM methods, the odometry information or IMU (Inertial Measurement Unit) data is not necessary in Hector SLAM method. On the other hand, the estimation of 2D position in Hector SLAM is dependent on the robust scan matching algorithm which is obtained using Gaussian-Newton method. By the scan matching algorithm, the end points of a laser scan are optimally aligned with the constructed map. The laser scan end points are projected within a map and the probabilities of occupancy are estimated. The algorithm seeks to determine the best laser scan alignment within the map by finding the rigid transformation $\xi = (p_x, p_y, \psi)^T$ which can be expressed in (1):

$$\zeta^* = argmin_{\xi} \sum_{i=1}^{n} [1 - M(S_i(\xi))]^2 \quad (1)$$

where the map value is returned by the function $M(S_i(\xi))$ at $S_i(\xi)$ which are the world coordinates of scan end points $s_i = (s_{i,x}, s_{i,y})^T$ that can be expressed in (2)

$$S_i(\xi) = \begin{pmatrix} cos(\psi) & -sin(\psi) \\ sin(\psi) & cos(\psi) \end{pmatrix} \begin{pmatrix} s_{i,x} \\ s_{i,y} \end{pmatrix} + \begin{pmatrix} p_x \\ p_y \end{pmatrix} \quad (2)$$

where, p_x and p_y are map coordinates.

According to (3), $\Delta\xi$, which optimizes the error measure, is estimated:

$$\sum_{i=1}^{n} [1 - M(S_i(\xi + \Delta\xi))]^2 \to 0 \quad (3)$$

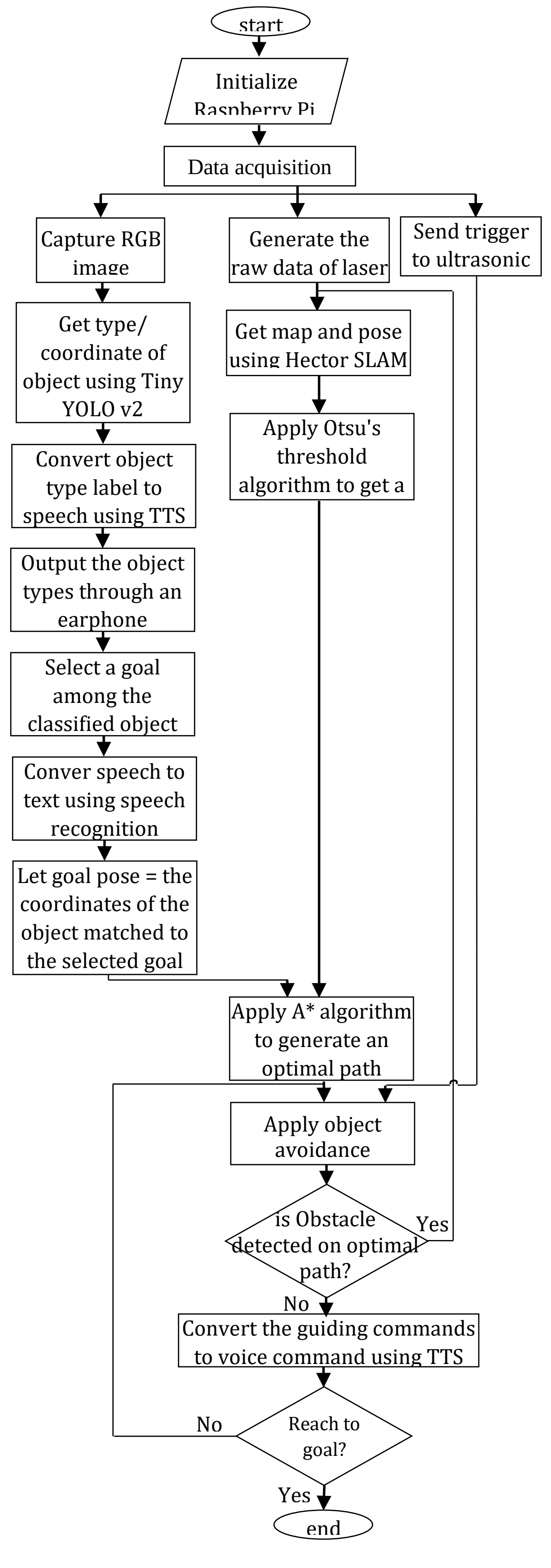


Fig. 3. The flowchart of the proposed system

The 1st order Taylor expansion is applied to $M(S_i(\xi + \Delta\xi))$ and the partial derivative with respect to $\Delta\xi$ is set to zero yields the Gaussian-Newton equation for the minimization problem:

$$\Delta\xi = H^{-1}\sum_{i=1}^{n}[\nabla M(S_i(\xi))\frac{\partial S_i(\xi)}{\partial \xi}]^T[1-M(S_i(\xi))] \quad (4)$$

where,

$$H = [\nabla M(S_i(\xi))\frac{\partial S_i(\xi)}{\partial \xi}]^T[\nabla M(S_i(\xi))\frac{\partial S_i(\xi)}{\partial \xi}] \quad (5)$$

In this work, IMU sensor does not be used and LiDAR (Neato XV-11) is utilized to provide data for Hector SLAM. The Neato XV-11 sensor is a 360° 2-D laser scanner that reads data with distance range 0.15m-6m and the scanning frequency of 5Hz. The Hector SLAM method is implemented using ROS (Robot Operating System), which includes a set of libraries and tools to develop the applications of robot.

Initially in mapping and localization module of the proposed system, the captured data from laser scanner are sent to ROS system to create a 2-D occupancy grid map of the surrounding environment by Hector SLAM. The white color of occupancy grid map indicates that obstacle non-existence, while the black one indicates obstacle existence. The dark grey part indicates unknown cell. The Otsu's threshold segmentation method [10] is used to convert Hector SLAM map into binary image in order to classify a map cell as free (obstacle non-existence) or occupied cell (obstacle existence/ unknown cell) as shown in Fig. 4. Finally, the resultant map is resized to match the distance between the two successive pixels in map with one walking step. This will lead to reduce the computation time for path planning approach. Since ROS is a robotic framework, there is no function in it to help the VI person to reach his/her destination. Therefore, a function in python programming language is implemented to get the current position and the resultant map provided by Hector SLAM in ROS. The path planning module uses the current position and Hector map as inputs.

B. Path planning module

In global path planning approach, the current location is collected from the mapping and localization module and the desired goal from the speech recognition algorithm. The building map will be used to plan a global path from the current location to the goal by A* algorithm [11]. The output of A* path planning algorithm is the shortest path which is followed by VI person.

When following a planned path, the visually impaired person encounters the static obstacles and temporary obstacles. The static obstacles (such as walls) which are part of a building map and A* algorithm has considered them into account. On other hand, the temporary obstacles that appear during path following need to be detected in real time. Therefore, obstacle avoidance based on ultrasonic sensor is adopted in this paper. The ultrasonic sensor placed on white cane will be utilized to ensure there is no object detected at each step of VI's movement. Ultrasonic sensor detects object within a 30° and in detection range 0.02m-4m. Based on the shortest global path provided by A* algorithm and ultrasonic based obstacle avoidance, the guiding command is generated to guide VI person at each step to his/her final destination. These guiding commands will be converted to voice message by Text-To-Speech method and output on an earphone.

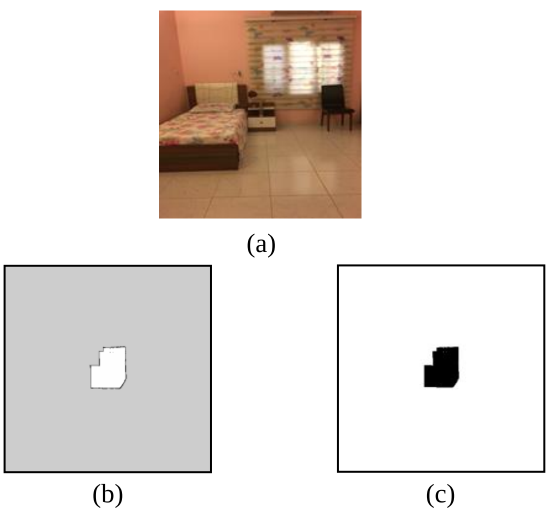


Fig. 4 Example of the navigation system process on the real-time scenario (a) RGB-image (b)Hector map (c) Otsu image

C. Recognition module

In spite of the navigation system improves the visually impaired person mobility in his/her daily lives, he/she still faces a challenge in understanding the unknown environment (perception of objects in the environment) while navigates through it and how to identify the final destination in the navigation system.

To achieve this perception, the object recognition system is implemented in this work that provides VI person with information about the objects which encountered along his/her way (i.e., object type and location). The object recognition system in this paper depends on deep learning technique called YOLO v2 (You Only Look Once) [8]. Specifically Tiny YOLO v2 using pre-trained weight on Pascal VOC dataset is adopted in our recognition module which is the smallest version of the complex YOLO v2 algorithm. Since it uses fewer layers, this makes it lite and faster than original one and suitable to work on raspberry pi

. Tiny YOLO v2 as shown in Fig. 5 consists of 9 convolution layers with 6 pooling layers.

Once the image in front of VI person is captured by CameraPi, it is resized to (416 × 416) pixels and then is applied to Tiny YOLO v2 method to predict the object class among 20 classes in Pascal VOC dataset and making the bounding boxes. The output of the recognition module is the class of the detected object which is converted to audio message by TTS process to notify VI person with the surrounding information. Since the coordinates of the goal in navigation system is needed, the VI person selects his/her desired destination among the classified objects on the audio message. The bottom right coordinates of bounding box for the selected destination is taken as a goal coordinate to our navigation system. Fig. 6 shows an example of applying a tiny YOLO v2 on a real-time scenario.

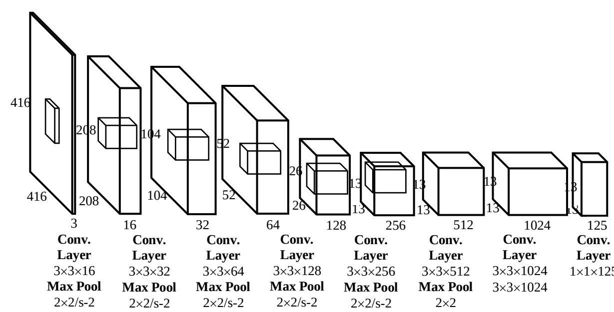


Fig. 5 Tiny YOLO v2 architecture.

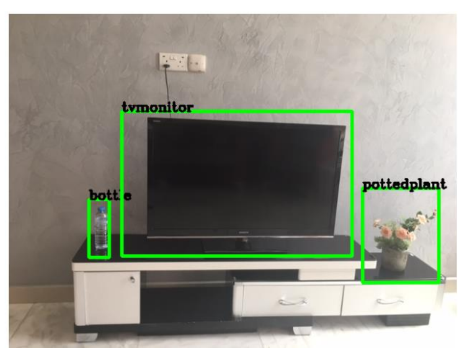


Fig. 6 Tiny YOLO v2 on a real-time scenario.

D. *Speech interface module*

VI person interfaces with this navigation system via speech synthesis and recognition modules. USB microphone is used for audio in since the raspberry pi does not have a sound card and the existing audio jack is used for audio out by earphone.

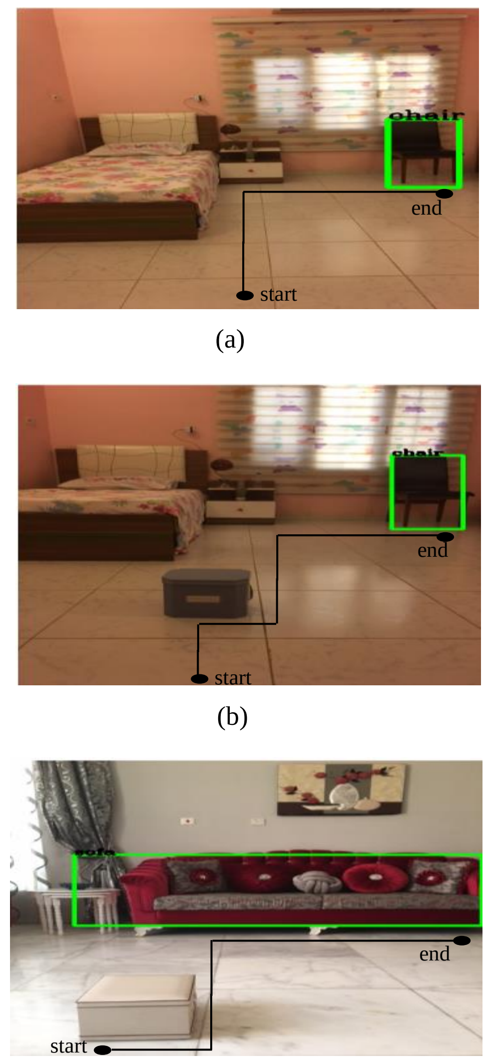
Text-To-Speech process is implemented to render system audio feedback such as object information and guiding commands via earphone. Furthermore, a speech recognition technology is used to allow VI person interacted with the system for query information and desired destination.

IV. Demonstration

The series of experiments have been conducted on the real-time scenarios in an indoor environment with/without obstacles to verify the effectiveness of the proposed system. The obstacle includes chair, box and table was used in the visually impaired person path with various start points and destinations. The results have shown that the proposed system is appropriate enough to help a VI person navigate in the buildings from one place to another and can classify the objects in the scene accurately. In navigation system, the length of user's stride was assumed about 35cm in path planning algorithm. The Hector map has a resolution of 0.05 meter/pixel. The recognition process runs online on Raspberry Pi and achieves approximately 0.6 fps. The path from a start point to a goal on the different real-time scenarios is shown in Fig. 7.

V. Conclusion

In this paper, an indoor assistive device that provides both real-time navigation and object recognition capabilities for visually impaired people is presented. The proposed system utilizes low cost sensors to aid VI people who have low income and can't afford the cost of the expensive assistive systems to navigate safely in an unknown environment. Navigation system in this work based on hector SLAM method to construct a 2D map of the surrounding environment and on A* algorithm to plan a shortest path for VI person. Hector SLAM is selected among all other 2D SLAM methods because it is produced accurate map, the pose estimation is calculated by the scan matching without the odometry information and computational complexity is lower due to the use of Gauss-Newton algorithm. Furthermore, Tiny YOLO v2 model which has the best performance between speed and accuracy in real time is implemented in the recognition system to increase the perception ability of the surrounding environment. An audio feedback message is utilized to deliver the guidance information. In future, the obstacle avoidance algorithm will be improved. In addition, object recognition capability will be enhanced, especially the detection of staircase and trained Tiny YOLO v2 on the customized data set. Moreover, the proposed system will be connected with cloud.



(a)

(b)

(c)

Fig. 7. The path from start to end point in the real-time scenarios

REFERENCES

[1] S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das and S. Gupta, " Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 1148–1159, 2014

[2] M. Islam, M. Sadi, K. Zamli and M. Ahmed, "Developing walking assistants for visually impaired people: A review", *IEEE Sensors Journal*, vol. 19, pp. 2814–2828, 2019.

[3] Z. Fei, E. Yang, H. Hu and H. Zhou, "Review of machine vision-based electronic travel aids", *IEEE 23rd International Conference on Automation and Computing (ICAC), Huddersfield, UK*, pp. 1–7, 2017.

[4] A. Krizhevsky, I. Sutskever and G.E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information Processing Systems (NIPS)*, pp. 1097–1105, 2012.

[5] B. Lin, C. Lee and P. Chiang, "Simple smartphone-based guiding system for visually impaired people", *Sensors*, vol.17, 2017.

[6] F. Bashiri, E. LaRose, J. Badger, R. D’Souza, Z. Yu and P. Peissig, "Object detection to assist visually impaired people: A deep neural network adventure", *International Symposium on Visual Computing, Springer*, pp. 500–510, 2018.

[7] R. Trabelsi, I. Jabri, F. Melgani, F. Smach, N. Conci and A. Bouallegue, "Indoor object recognition in rgbd images with complex-valued neural networks for visually-impaired people", *Neurocomputing, Elsevier*, vol. 330, pp. 94-103, 2019.

[8] J. Redmon and A. Farhadi, “YOLO9000: better, faster, stronger”, *arXiv:1612.08242v1 [cs.CV]*, 2016.

[9] S. Kohlbrecher, O. Stryk, J. Meyer and U. Klingauf, "A flexible and scalable SLAM system with full 3D motion estimation", *IEEE International Symposium on Safety, Security, and Rescue Robotics*, pp. 155-160, 2011.

[10] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Transaction on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.

[11] P. Hart, N. Nilsson and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths", IEEE Transactions on Systems Science and Cybernetics (SSC), vol. 4, Issue 2, pp. 100–107, 1968.