© 2022 IJRTI | Volume 7, Issue 9 | ISSN: 2456-3315

OBJECT DETECTION FOR VISUALLY IMPAIRED PEOPLE USING YOLO V4ALGORITHM

AUTHORS:
ASHIKA.K [1]
MONIKA.N [2] NIVETHA.M [3] ROHINI.A [4]
GUIDED BY:
Mrs.UMA

ABSTRACT

Blind navigation has become a challenging task nowadays. Blind people cannot detect and avoid obstacles like sighted people and need guidance to avoid such obstacles. But the limited potential of white canes makes it impossible for a navigator to detect all possible threats. Therefore, there is not enough aid to navigate safely on the white cane. To protect blind people's safe and independent navigation, more insight into their current environment must be provided. This study proposes a novel approach for obstacle detection based on deep learning to assist in blind navigation.In this study, a prototype was developed using single-shot detector (SSD) for obstacle detection and distance estimation due to real-time performance and high accuracy of SSDs. To train the SSD for obstacle detection data was gathered using a simulation environment. The result of the obstacle detection model was used to estimate the distance of the obstacles. The final result is communicated to the user through audio sequences by combining the feedback from obstacle detection and distance estimation.. The prototype system is deployed on a smartphone and a real-time video stream captured by the smartphone camera is processed to detect obstacles. To train the SSD for obstacle detection SSD MobileNet Architecture was used and the data to train the SSD was generated using a simulation environment. To estimate the distance of the detected obstacles, SSD based MonoDepth algorithm was used.The mean average precision (mAP) value of all the classes of the SSD for obstacle detection reached more than 70%. High accuracy and high speed of obstacle detection can be achieved by computer simulation but there is delay when estimating distance. Usability and efficiency of the prototype system exceeded 65% according to the usability evaluation feedback.

INTRODUCTION

Overview

According to World Health Organization (WHO) statistics, the number of blind people is estimated at 1.3 billion, of which 36 million were blind in 2018.By 2019 a total of 2.2 billion people suffers from some form of visual impairment [1]. According to these statistics, the blind and visually impaired community is increasing yearly. Engaging in day-to-day activities without hazel is an extremely difficult task for a visually impaired/ blind person. It becomes more difficult when it requires traveling through unfamiliar locations without a close companion to assist them along the way. Guide dogs are used in assisting visually impaired persons, but it is not easy to get a trained animal due to the high cost. Furthermore, traveling in familiar environments without help could also be challenging since the dynamic situations along the way cannot be predicted earlier, and responding to those situations in real-time is not possible for a blind per Blind people navigate without a clear visual map about the obstacles in their path. Therefore, it is not possible to take precautions to avoid such obstacles similar to a normal person with good vision. This study focuses on improving the independent navigation of the blind and ensuring their safety while navigating.

Problem Statement

The white cane is the most commonly used assistive device by the blind community for their navigation. It is being used to detect and avoid obstacles that could collide with them along the way. But when considering the rapidly changing surroundings in the present, putting all the trust in the white cane does not ensure the safe independent navigation of the blind people. Furthermore, the limited reachability of the white cane reduces the possibility of detecting potential threats to the blind navigator. Considering the aforementioned limitations of white cane, a solution that improves the safety and efficiency of independent navigation by providing additional insights into the current environment is more convenient and useful for the blind community. Replacing a white cane with another assistive device may not be a practical solution for many blind users, as it is also an indicator to inform others of a blind navigator.Although white canes are available with advanced technologies, they are very expensive. Therefore, introducing a solution for blind navigation that can be used in combination with white cane and is inexpensive will be of great help to the visually impaired. Furthermore, to be used in blind navigation, the assistive methods and solutions must be accurate to guide and be convenient to use while navigating.

Background

In recent years, many researchers have focused on the topic of blind navigation and proposed several approaches to accomplish safe navigation, including obstacle detection/avoidance, distance estimation, and feedback [2]-[5]. Obstacle detection is an important part of blind navigation and calculating the distance to obstacles and providing the necessary feedback to avoid them are important aspects of a blind navigation system. Most of the existing outdoor navigation systems use GPS (Global Positioning System) technology for localization [6]-[8]. But their low accuracy and signal loss issues have made such systems unreliable for the blind. Major drawbacks in GPS systems are the inability of giving any feedback on moving obstacles and the inability of giving information on the obstacles near the user [9]. So such systems are not useful to assist blind people. In recent years computer vision-based approaches, named ETAs (Electronic Travel Aids) are proposed.[10]-[13]. ETAs should be reliable, affordable, and light. An electronic travel aid (ETA) named "EyeCane" was designed in [14] that translates point-distance information into auditory and tactile cues. According to [15], wearable distraction avoidance ETAs that provide survey, audio feedback and tactile feedback are identified as Echolocation, Navbelt, vOICe and tactile handle. Dynamic and static obstacle detection plays an important role when guiding the blind and research has been carried out to discover efficient obstacle detection methodologies [2]-[4]. These existing systems have adapted vision-based techniques by using Monocular cameras, Stereo-cameras, RGB-D cameras, Time-of-flight (TOF), etc. and other sensors such as Ultrasonic sensors, SONAR, LIDAR, etc. to Capture the environment and associated obstacles. Structure from Motion algorithm, RANSAC algorithm, SIFT, Multicase Lucas-Kanade algorithm and Event-based algorithm are some of the algorithms that were adapted for obstacle detection. Deep learning-based approaches are widely adapting to blind navigation in recent years [5][16][17]. Convolutional neural network (CNN) techniques are commonly used in object recognition tasks due to their high accuracy, although there are concerns about collecting a large set of training data. and overfitting due to noisy data when using CNNs.Smartphone usages of the blind persons are increasing due to their features such as screen reader, haptic feedback, audio feedback, adjustable contrast, audible battery indicator, and vibration. Light weight of a smartphone is also an advantage and it does not affect the navigation. Therefore, smartphones are widely being used in assisting the blind navigation [2]-[5]. Further laptops and other customized devices are used in designing assistive devices for blind navigation [18]-[21].

Aims and Objectives

Navigation is a difficult task for blind people. Since many blind people use only the white cane for navigation, there is a higher possibility of facing to harmful situations. Therefore, there is a critical need of providing more information of the environment around them to ensure their safety. The main objective of this research is to propose a novel obstacle detection mechanism to improve the safety and independent mobility of the blind people.The objectives of the study are described below.

The main aim of the proposed approach is real-time obstacle detection with higher accuracy. To achieve this, a deep learning-based approach is used to obstacle detection due to the higher accuracy and real-time prediction of results of single-shot detector (SSD).It is expected to fill the gap of detection when using white cane due to its low accessibility. Mainly when considering obstacles that are above ground-level are not reachable using the white cane and thus not detected by the white cane. With this proposed mechanism it is aimed to cover such obstacles.

• Furthermore, communicating to the user about potential threats in their surroundings in an effective manner to avoid such threats successfully is also an aim of the study.

Proposed Approach

With the improvements of the technology, it is possible to create deep learning models that provide real-time results with high accuracy [22]. Therefore, deep learning is used for object detection tasks in various applications. Similarly, this proposed system uses a deep learning-based obstacle detection mechanism.Training a deep learning model demands a large amount of data. To train this object detection model, the data was collected using a simulation environment instead of using real-world data.There are different architectures and frameworks which are available for deep learning. Faster R- CNN [23], Masked R-CNN [24], YOLO [25], SSD [26] are such popular architectures for object detection. Among these, some architectures are fast enough to be running on mobile devices such as YOLO (YOLO family) and SSD (Single Shot MultiBox Detector). It is a challenging task to select a suitable architecture and considering requirements to perform a task and its deployment requirements an architecture must be selected. Here, considering the requirement of mobile deployment, the SSD MobileNet architecture is used to train a deep learning model. The system prototype consists of three main modules namely object detection model, depth estimation module and feedback module. After detecting an obstacle through the object detection model, its depth estimation will be carried out using the depth estimation module. The depth estimation module is deployed in an external server. To perform depth estimation MonoDepth [27], a deep learning-based approach was used. This proposed approach is a deep learning-based obstacle detection mechanism that combines simulation, deep learning, and computer vision techniques.

Contribution

The main contribution of this research is to propose a novel obstacle detection mechanism to assist blind navigation that targets,High accuracy and real-time obstacle detection.A deep neural network is used for obstacle detection. In the proposed system, a novel approach was adapted that uses data generated from a simulation environment instead of real-world data to train the deep neural network. Simulation platforms are used in various domains as a mechanism of data generation. Similarly, a simulation platform was used here and to generate data a 3D realistic environment was created.A common mechanism that can be used to estimate the distance of both ground-level, above ground-level obstacles.A deep learning-based monocular depth estimation methodology is combined with binary thresholding to identify the obstacle with possible threats to the navigator.

Scope

This research will be conducted to propose an obstacle detection mechanism to assist blind people to navigate safely and independently. As it is clear that using only a white cane is not adequate for a blind person to navigate safely in outdoor environments, it is important to provide more information about the obstacles in their path. Due to the limited reachability of the white cane, especially the obstacles that exist above the ground-level are not detectable. Such obstacles are mainly focused to detect using the obstacle detection model.Using a simulation the required data is generated to train the deep learning model. The Deep Learning Model (SSD) will be trained using a machine learning library that supports mobile conversion since this model is deployed in a smartphone. To estimate the distance of the detected obstacles a pre-trained MonoDepth.algorithm [27] based implementation will be used. The computer's feedback is conveyed to the blind user through pre-recorded audio sequences.

Smartphone-Based Obstacle Detection for the Visually Impaired

Author : Alessandro Caldini, Marco Fanfani and Carlo Colombo

Year : 2015
In Caldini et al. [3], a calibrated smartphone on users' chest which installed a gyroscope is used to detect obstacles. It implements a modified Structure from Motion (SfM) algorithm for scene reconstruction. Visual data obtained from the camera and measurements obtained from the gyroscope are used in the developed algorithm. The proposed vision-based system focuses on obstacle detection to help visually impaired people to move autonomously in unknown indoor andoutdoor environments.

Virtual-Blind-Road Following-Based WearableNavigation Device for Blind people

Author : Jinqiang Bai, Shiguo Lian, Member, IEEE, Zhaoxiang Liu, Kai Wang, Dijun Liu

A color histogram is created based on the images acquired from smartphones and histograms are created for each frame and chooses the simplest RGB color space. Then, a binary RGB histogram is built for the safe region which is derived from the image region.

A Smartphone-Based Obstacle Sensor for the Visually Impaired

Author : En Peng, Patrick Peursum, Ling Li, and Svetha Venkatesh

Year : 2010

A detection system for obstacles on the ground at any height using a handheld smartphone is proposed in Peng et al. [2]. This proposed method uses computer vision techniques such as color histograms and edge detection for obstacle identification and obstacles are captured through the smartphone camera.

A Navigation System forthe Visually Impaired:
Fusion of Vision and Depth Sensor.

Author : Nadia Kanwal,Erkan Bostanci, Keith Currie, and Adrian F. Clark

Year : 2015

A Sensor-based navigation system for visually impaired persons using a fusion of vision and depthsensors that uses Microsoft Kinect Sensor is introduced in Kanwal et al. [31]. In the research, the expected outcome is to utilize the usage of both sensors by detecting obstacles using corner detection and then estimating their distance using a depth map of the corresponding scene. The data acquired from the Microsoft Kinect camera is processed to detect Edges using the Sobel edge detector, Corners using Harris&Stephens detector and Blobs using SIFT detector. But according to the paper, the Kinect sensor is not reliable as expected in outdoors and some sensors are unable to work in different environmental conditions and the infrared sensor has issues with blind spots insunlight.

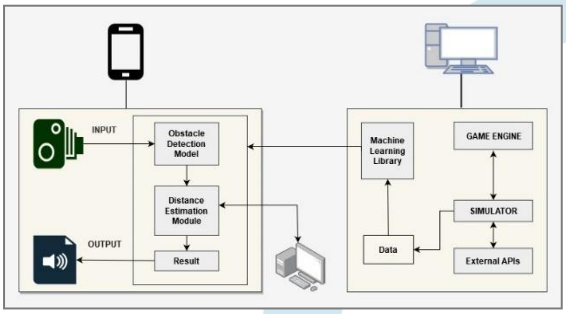
Thesis outline

The thesis is organized in the following format. The first chapter of the thesis identifies the problem of blind navigation and describes the proposed solution to overcome the issue. Chapter two presents the literature review on proposed approaches to assist blind navigation, background on simulation platforms, and usages of simulation in real-world and deep learning algorithms for obstacle detection.The research methodology and the architecture of the prototype are presents in three. The details of the proposed solution's development will be included in chapter four. Evaluation results of the obstacle detection model and the system prototype are presented in chapter five and the chapter concludes with the discussion section. The final chapter of the thesis contains the conclusion and future work.

Research methodology
The research methodology is based on constructive research. Constructive research consists of several phases. A relevant and practical problem must be available to apply the constructive research. After analyzing the problem domain using relevant

literature, practical experiences or using other methods a thorough understanding must be gained on the research area. Then considering the identified problems and research gaps, a novel prototype solution must be constructed and the solution could be presented as an algorithm, model or a framework. The constructed solution must be tested and evaluated thoroughly to be considered a theoretical contribution to the field. Considering the threats that the blind people face while navigating, this research is focused on finding a practical solution to assist them. Although the blind people use the white cane to assist them, due to its limited reachability, it is not an adequate solution. Therefore, more insights of thesurroundings must be provided to the navigator. Information of the obstacles with potential threatsmust be informed to the user based on the distance The proposed prototype uses deep neural network-based approaches to detect obstacles and estimate the distance of detected obstacles. This prototype will be deployed in a smartphone and therefore, it is developed as a mobile application. The blind navigator will use the smartphone camera to capture the surroundings in a video stream and after processing the data, the applicationwill provide necessary feedback through audio queues to avoid threatening obstacles in their path.

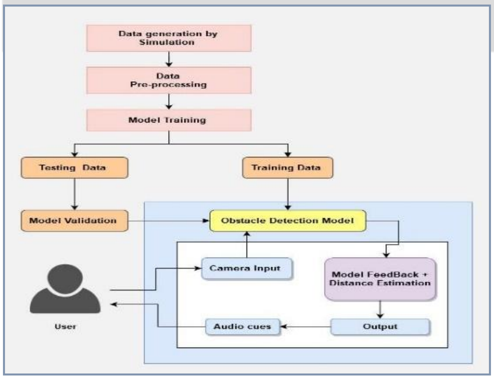
Design of the PoC (Proof of Concept Prototype)



1. Simulation platform
2. Game Engine (used simulation platform is a plugin for game engines)
3. Machine learning Library
4. External sever to perform depth estimation.
5. Mobile device: A suitable android smartphone with a quality camera and an adequate RAM toinstall the application.

A simulation platform is used to generate data to train a deep learning model. The simulation platform works as a plugin for a game engine and there are APIs available in the simulator that canbe extended using external APIs. By extending the available APIs or using other available methodsprovided by the simulation platform, the desired data generation can be performed. The trained deep learning model is deployed in a mobile platform therefore it's important to choose a suitable architecture that supports mobile deployments.

Process flow of the (PoC) System



.

The process flow for building the prototype system is described in the following section. For this step, a simulation platform will be used and such available simulation platforms are mostly built on game engines such as Unreal or Unity, etc. It is possible to create highly realistic 3D- scenes for games using such game engines. Similarly, it is possible to create new and customized environments after adding the plugins of the simulator. To create a suitable simulation environment/s it is to possible use environmental modules already available in the respective asset stores of the game engines. Further, it is possible to use single assets and create customized environments for the required purpose.

Data Generation

The simulation platforms allow modeling different types of physics models such as drones, vehicles and other objects.To dedicate them these physics models have cameras.These camera views can be accessed through API calls. For example, the cameras of the car model in AirSim can be accessed via front_center, front_right, front_left, fpv and back_center API calls. AirSim has car, multirotor and other physics models. Furthermore, some platforms allow to simply move a camera using keyboard keys through the simulation environment and retrieve data.

By using the capabilities of the simulation platforms, it is possible to model different objects with sensor capabilities. For an example a person with sensors attached to the body (wearing a headband with two cameras, etc.) can be modeled within the environment, and while simulating the way that person walks, data can be generated. Similarly, methods without physical engines can also be adapted into the data generation process. The simulation platforms allow us to modify or extend the behavior of physics models and cameras through APIs using different programming languages.To train a deep learning model,data collected through simulation is used. Faster R-CNN algorithm uses a region proposal network to create boundary boxes and utilizes those boxes to classify objects. But this reduces the real-time performance. For real-time obeject detection, a single shot ,multibox detector. SSD speeds up the process by eliminating the region proposal.

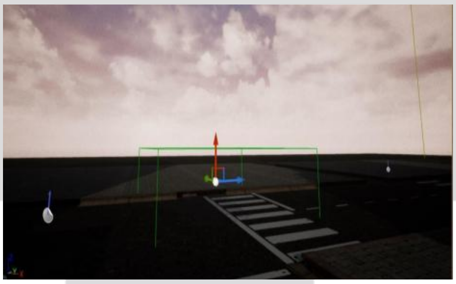
Data generation with AirSim by simulation

AirSim is a simulation platform that enables simulation in realistic 3D environments. It is open- source and has an active repository with good documentation and an active community. It also hasa number of APIs that provide access to cameras and different physics models and such available APIs can be modified using programming languages as Python or C++. AirSim has different physics models such as vehicle, and multirotor models provided for simulation and it is also possible to simulate without any physics model. In AirSim, this approach is called the 'Computer Vision' mode. In this mode the physics engines are disabled and the cameras are moved in desiredpaths and angles by the keyboard. Furthermore, AirSim provides RGB, depth and thermal camera output. It also provides output of the sensors such as accelerometer, gyroscope and barometer. CARLAis another simulation platform and it supports vehicle physics model and sensors such as Lidar. AirSimprovides more features compared to the other simulation platforms such as CARLA and DeepDrive.

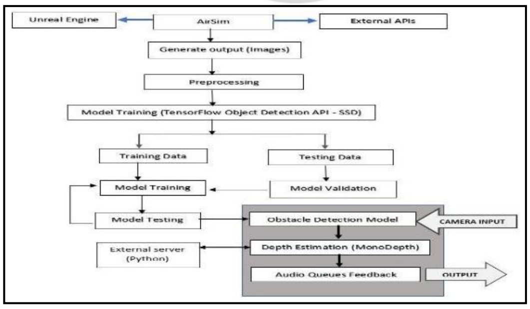
Considering the above facts AirSim was used in developing the proposed system. AirSim is used as a plugin for Unreal Engine which is a suite of integrated tools for game developers to design and build games, simulations, and visualizations.

Setting up AirSim: Creating Unreal projects and integrating AirSim

The AirSim plugin can be downloaded from the Microsoft Git hub repository. Using the Unreal project browser, we can create a new empty project with basic code C++. After creating the project,a suitable simulation environment is downloaded from Unreal Market Place and merge the configuration files and plugins with the project configurations. This process is also possible to do automatically without manually transferring files by directly adding an asset/s into a project. Then the AirSim plugin must be enabled in the Unreal project. This setup was completed using Unreal Engine version 4.23 and Visual Studio 2017 was used as an editor for Unreal projects.



Proposed System



This chapter describes the process carried out to implement the proposed system. Data generationusing a simulation platform is described in section 4.1 and deep learning object detection model training is described in section 4.2. Section 4.3 explains the modification of the prototype to integrate the distance estimation module to the prototype.

REFERENCES

1. "Vision Impairment and Blindness." Accessed November 1, 2019. https://www.who.int/news- room/fact-sheets/detail/blindness-and-visual-impairment.

2. Peng, En, Patrick Peursum, Ling Li, and Svetha Venkatesh. "A Smartphone-Based Obstacle Sensor for the Visually Impaired." In Ubiquitous Intelligence and Computing, edited by Zhiwen Yu, Ramiro Liscano, Guanling Chen, Daqing Zhang, and Xingshe Zhou, 6406:590– 604. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. https://doi.org/10.1007/978-3-642-16355-5_45.

3. Caldini, Alessandro, Marco Fanfani, and Carlo Colombo. "Smartphone-Based Obstacle Detection for the Visually Impaired." In Image Analysis and Processing — ICIAP 2015, editedby Vittorio Murino and Enrico Puppo, 9279:480–88. Cham: Springer International Publishing,2015. https://doi.org/10.1007/978-3-319-23231-7_43.

4. Mocanu, Bogdan, Andrei Bursuc, Titus Zaharia, and tapu. "A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People." In 2013 IEEE International Conference on Computer Vision Workshops, 444–51. Sydney, Australia: IEEE, 2013. https://doi.org/10.1109/ICCVW.2013.65.

5. Lin, Bor-Shing & Lee, Cheng-Che & Chiang, Pei-Ying. (2017). Simple Smartphone-Based Guiding System for Visually Impaired People. Sensors. 17. 1371. 10.3390/s17061371.

6. Velázquez, R., Pissaloux, E., Rodrigo, P., Carrasco, M., Giannoccaro, N., Lay-Ekuakille, A., 2018. An Outdoor Navigation System for Blind Pedestrians Using GPS and Tactile-Foot Feedback. Applied Sciences 8, 578. https://doi.org/10.3390/app8040578

7. Liu, Yongqing & Chen, Qi. (2018). Research on Integration of Indoor and Outdoor Positioning in Professional Athletic Training. Proceedings. 2. 295. 10.3390/proceedings2060295.

8. Kumar Yelamarthi, Daniel Haas, Daniel Nielsen, and Shawn Mothersell. 2010. RFID and GPSintegrated navigation system for the visually impaired. In 2010 53rd IEEE International Midwest Symposium on Circuits and Systems. IEEE Press, Piscataway, NJ, USA, 1149–1152.
9. DOI: http://dx.doi.org/10.1109/MWSCAS.2010.5548863

10. "Accuracy of GPS Data - OpenStreetMap Wiki." Accessed November 1, 2019. https://wiki.openstreetmap.org/wiki/Accuracy_of_GPS_data.
11. "Electronic Travel Aids for the Blind." Accessed November1,2019. https://www.tsbvi.edu/orientation-and-mobility-items/1974-electronic-travel-aids-for-the- blind.

12. Amedi, et. al "Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex," Nature Neuroscience, vol.10, no. 6, pp. 687-689, June 2007.

13. T. Schwarze, M. Lauer, M. Schwaab, M. RomMATLABs, S. B¨ohm, and T.J¨urgensohn, "A camera-based mobility aid for visually impaired people",KI-K¨unstliche Intelligenz, pp. 18, 2015.

14. P. Bach-Y-Rita and S. W. Kercel "Sensory substitution and the human machine interface, " Trends Cogn Sci., vol. 7, no. 12, pp.541-546, Dec.2003.

15. Shachar, Maidenbaum, Hanassy Shlomi, Abboud Sami, Buchs Galit, Chebat Daniel-Robert, Levy-Tzedek Shelly, and Amedi Amir. "The “EyeCane”, a New Electronic Travel Aid for the Blind: Technology, Behavior & Swift Learning." Restorative Neurology and Neuroscience, no. 6 (2014): 813–824. https://doi.org/10.3233/RNN-130351.

16. Dakopoulos, D., and N.G. Bourbakis. "Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 40, no. 1 (January 2010): 25–35. https://doi.org/10.1109/TSMCC.2009.2021255.

17. Tapu, Ruxandra & Zaharia, Titus. (2017). Seeing Without Sight — An Automatic Cognition System Dedicated to Blind and Visually Impaired People. 1452-1459. 10.1109/ICCVW.2017.172.

18. Tapu, Ruxandra & Zaharia, Titus. (2017). DEEP-SEE: Joint Object Detection, Tracking and Recognition with

Application to Visually Impaired Navigational Assistance. Sensors. 17. 2473.10.3390/s17112473.

19. Katzschmann, Robert K., Brandon Araki, and Daniela Rus. "Safe Local Navigation for Visually Impaired Users With a Time-of-Flight and Haptic Feedback Device." IEEE Transactions on Neural Systems and Rehabilitation Engineering 26, no. 3 (March 2018): 583–93. https://doi.org/10.1109/TNSRE.2018.2800665.

20. Aladren, A., G. Lopez-Nicolas, Luis Puig, and Josechu J. Guerrero. "Navigation Assistance forthe Visually Impaired Using RGB-D Sensor With Range Expansion." IEEE Systems Journal 10, no. 3 (September 2016): 922–32. https://doi.org/10.1109/JSYST.2014.2320639

21. Kayukawa, Seita & Higuchi, Keita & Guerreiro, João & Morishima, Shigeo & Sato, Yoichi & Kitani, Kris & Asakawa, Chieko. (2019). BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians. 10.1145/3290605.3300282.

Conclusion and Future Work

This study proposed a deep learning-based obstacle detection mechanism to assist blind navigation. The developed prototype consists of three main modules namely obstacle detection, distance estimation, and audio queue feedback. For obstacle detection, a deep neural network is trained using the data generated by simulation instead of real-world data. To estimate the distance, monocular depth estimation was used and the feedback is communicated to the user via audio queues.To improve user-friendliness in smartphone,the prototype is deployed.The distance estimation calculation is performed in an external server and the result is sent back to the smartphone.The main research question focused on this study is,How to find a solution to fill the detection gap of the white cane that arise due to its limited reachability to improve the independent navigation of the blind?Under that, the followinsub-questions were focused.How to generate data using simulation to detect target obstacles?How to develop a learningmodel from collected data that can be deployed in a mobile platform?What object detection algorithms or architecture should be adapted?What is the most suitable mechanism to estimate the distance of the detected obstacles?To improve the safe and independent mobility of the blind people, more insights about their current surroundings must be provided to them. Thus, they can avoid threatening obstacles and situations successfully. Considering the real-time performance and the high accuracy of the single-shot detector, a deep learning-based obstacle detection mechanism was proposed in this study to improve the navigation of the blind people