EasyChair Preprint

№ 7813

# Blind Assistance System Using Machine Learning

Naveen Kumar, Sanjeevani Sharma, Ilin Mariam Abraham and Sathya Priya



April 19, 2022

# Blind Assistance System Using Machine Learning

* Naveen Kumar[1], Sanjeevani Sharma[2],Ilin Mariam Abraham[3],and Sathya Priya S[4]

[1,2,3,4] Department of Computer Science and engineering
Hindustan Institute of Technology and Science
Chennai, India

[1*]18113100@student.hindustanuniv.ac.in
[2]18113101@student.hindustanuniv.ac.in
[3]18113117@student.hindustanuniv.ac.in
[4] sathyap@hindustanuniv.ac.in

* Sathya Priya S
Department of Computer Science and engineering
Hindustan Institute of Technology and Science
Chennai, India
sathyap@hindustanuniv.ac.in

**Abstract.** Blindness is one of the most frequent and debilitating of the various disabilities. There are million visually impaired people in the globe, according to the World Health Organization (WHO).

The proposed system is designed to aid visually impaired persons with real-time obstacle detection, avoidance, indoors and out navigation, and actual position tracking. The gadget proposed is a camera-visual detection hybrid that performs well in low light as part of the recommended technique, this method is utilized to detect and avoid impediments, as well as to aid visually impaired persons in identifying the environment around them. A simple and effective method for people with visual impairments to identify things in their environment and convert them into speech for improved comprehension and navigation. Along with these, we have depth estimation, which calculates the safe distance between the object and the person, allowing them to be more self-sufficient and less reliant on others. We were able to achieve this model with the help of TensorFlow and pre-trained models. The approach we suggest is dependable, inexpensive, practical, and practicable.

**Keyword :**Depth Estimation, Object Detection, Single Shot Detection, TensorFlow.

# 1 Introduction

The goal of restoring vision to those who have been blinded from the outside is the focus of a comprehensive evaluation in both design and medication. Exploring their environment without colliding with any impediments is one of the most difficult challenges for vision-impaired people. The vision handicapped have used long sticks and guide hounds to complete this test for a long time. These long sticks and guide dogs, on the other hand, will only provide information about nearby obstacles within a limited range, while excluding weather information. This project was created by merging the ssd algorithm into the field of machine learning using the TensorFlow api packages. Having the ability to provide a completely aiding guide by offering vocal feedbacks is one significant factor in all prior systems where object detection is regarded a key step for the visually challenged in their navigation.

Millions of people suffer from a variety of issues, including visual impairment. Even the most basic tasks are challenging for them. Even in the privacy of their own house or the struggle to travel from one point to another in the business without becoming dependent on others. In truth, according to a new study, blindness can be restored to a degree, but in the long run, the majority of people would require assistance to carry out their daily duties.

Even though there are proven ways to help these people, such as a cane to help them travel or a guide dog, they are insufficient and prone to errors. Despite the fact that they can help them avoid any obstacles in their way, they cannot help them see what is right in front of them.

As a result, our contribution to solving this problem statement is to develop a blind aid system that enables object detection and real-time speech conversion so that they can hear it as and when needed, as well as depth estimate for safe navigation. This is encapsulated in a single system that is utilized to make it easier for the blind to function in society.

The goal of this exercise is to find a solution. This would not only assist these exceptional individuals in avoiding obstacles and managing their daily tasks, but it would also allow them to use their brains to see their surroundings. Giving kids real-time voice feedback is one method to help them become more self-reliant and safer. This device will assist blind individuals by tolerating their voice directions and applying picture handling algorithms to perceive items, as well as providing the

client with a hearable result so they can explore the bar. This framework will also recognize a few major events, such as a "Washroom" sign, which will provide information. The visually challenged person can recognize the image as soon as it is discerned. Objects, signs, and boards, among other things, are likely to be detected by the framework. This will assist the vision-impaired person in completing duties and organizing daily errands.

Our project's main goal and motive is to achieve this. The rest of the paper is formatted as follows. Section II describes the literature review. Section III explains the proposed system. Section IV contains the prerequisites. The suggested system is defined in Section V. Section VI shows how to design a system. The suggested system and its implementation are detailed in Section VII. Sections VIII and IX, respectively, provide the discussion and conclusion.

## 2 LITERATURE SURVEY

A gadget called a multi sensor probe is utilized to conduct the duty of people detection anytime the user is walking in a busy area. Infrared radiation is used by the PIR sensor to detect or check the movement of the person. The target distance and velocity provided by the Sonar module are used to calculate the actual distance. Smart glasses and a walking stick were utilized to identify obstacles. The smart glass does this, and the walking stick serves as a reminder to the blind of the impediment. In the event of a collision, information will be transmitted to the caregivers via an internet portal. A visually impaired system was invented by Devashish Pradeep Khairnar, Rushikesh Balasaheb Karad, and Apurva Kapse. A number of publications have been reviewed and categorized according to their application areas. The proposed VC (Vision Challenged) Assistant System contains four components that assist persons who are visually impaired: Obstacle recognition, avoidance, inside and outside navigating, and actual position sharing are all features available. The recommended device is a hybrid of a smart glove and smartphone software that performs well in low-light situations. The suggested solution includes the use of a smart glove to identify and avoid obstacles, as well as to aid visually impaired persons in recognizing their surroundings. The smartphone-based obstacle and object detection system detects a variety of items in the surroundings.

Deepak Gaikwad, Chaitalee Baje, Vaishnavi Kapale, and Tejas Ladage recommended that the system's functioning start with infrared sensors detecting impediments. Infrared sensors detect obstacles in three directions: front, left, and right. It uses sensor circuitry to detect impediments. This is then supplied as an input to the microcontroller. This controller is connected via Bluetooth to an Android phone, which creates speech depending on microcontroller commands. The influence of multipath transmission on multi-server offloading was predicted and investigated. They tested this approach using both real-world and simulation

versions. Ali Khan and Aftab Khan wrote a study with the goal of developing something exclusively for the blind. This study resulted in the development of an obstacle detection system using ultrasonic sensors. The wearable garment has ultrasonic sensors, as well as a vibration device and a buzzer. Sensors monitor the user's surroundings and notify them to any obstructions via vibration and a buzzer sound.

An Arduino-based ultrasonic blind walking stick is described in this paper. This gadget assists blind people in sensing obstacles and doing activities with ease and comfort. Because the blockage is not detected with a standard stick, it is useless for visually impaired persons. An ultrasonic sensor is used to detect the length between objects and the Smart Walking Stick. The user may hear the Buzzer and learn about the impediment when items or barriers come within range of the ultrasonic sensor. It identifies impediments in the user's path and allows him to move freely. Ashish Mishra and Changzhi Li invented a gadget that detects barriers using RADAR architecture. It is based on the RADAR's transmitter and receiver. Miniaturization and portability are the device's advantages. Shubham Suman, Sushruta Mishra, Kshira Sagar Sahoo, and Anand Nayyar worked on echolation and image processing using an image detecting sensor. The static or dynamic items put in the region are identified using these collected photos. They identify the barrier and distance using an ultrasonic sensor. It has a GPS module that can assist blind people in navigating.

## 3 PROPOSED SYSTEM

The design aims to replace existing technologies of eyeless navigation systems based on detectors and buzzers with a simpler yet effective way of creating an eyeless backing system based on machine literacy, where we can descry an object while receiving real-time voice feedback and depth estimation with the required delicacy. The proposed system is more efficient and reliable. The system is set up to record real-time frames and execute all calculations.
The object's class will be turned into dereliction voice notes after speech module testing, and will also be sent to the eyeless victims for assistance. We've employed an alert mechanism in addition to the item finding, where an approximate will be calculated. If that Eyeless Person is truly close to the frame or is far down at a safer area, nevertheless, it will induce voice- grounded labor’s as well as distance units. The discipline of computer vision is concerned with detecting meaningful things in photos and movies (by creating rectangle boxes around them in our case). We can also offer accurate distances and convert labeled texts into voice answers. Our strategy is trustworthy, cost-effective, feasible, and practical. This allows the blind to be self-sufficient in society and to avoid the societal barriers that still exist. An Integrated Machine Learning System that allows eyeless sufferers to recognize and classify common everyday items in real time, generate verbal feedback, and calculate distance, producing alerts whether they're veritably close or veritably far down from the object was one such endeavor on our side. Handicap Research The same approach can be used to enforce medium. Now that the world is changing

dramatically and new discoveries in medical science are occurring, it is necessary to improve the status of the visually bloodied as well. To make them more independent in all aspects of their daily lives, as well as from a business one, a change like this was more vital to bring about and implement.

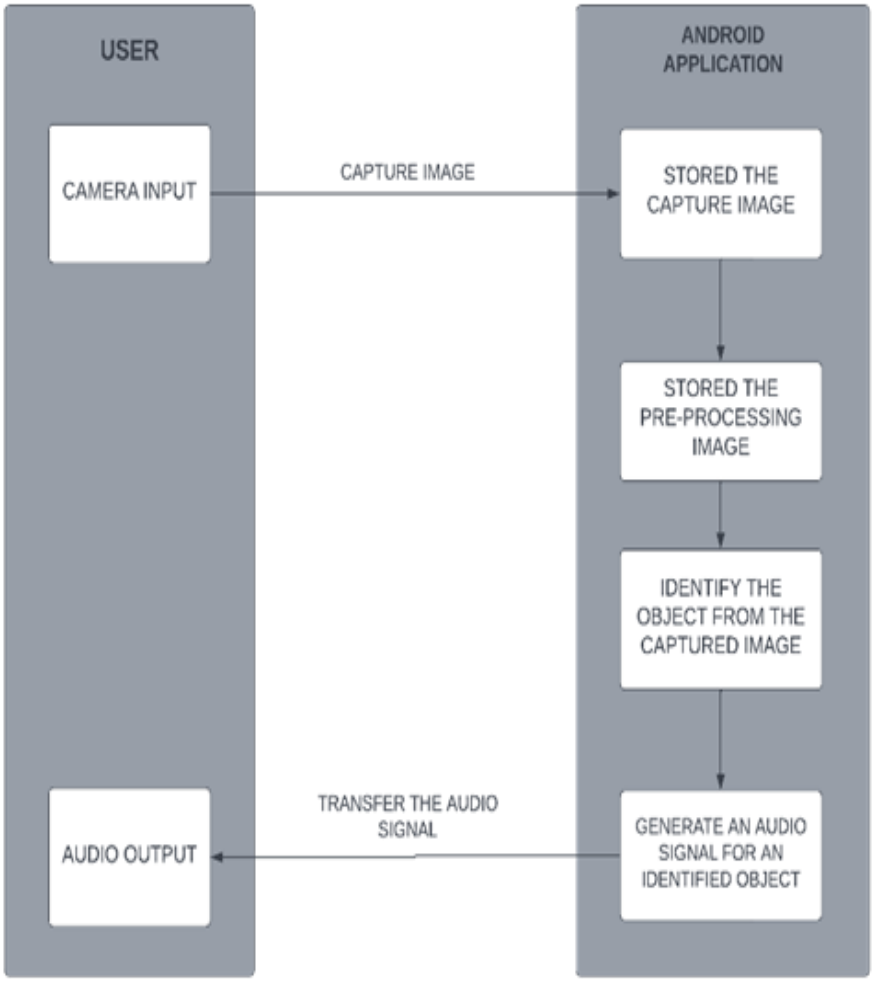


**Fig 1:** Block Diagram of the system

The camera input is shown in this diagram will capture the image and send it to storage as well as it is stored and pre-processed and try to identify the image and after identifying the image, it will generate a voice singel which can be heard by the user . Voice feedback will also be delivered once the image has been captured.

# 4 METHODOLOGY AND SYSTEM DESIGN

Our current application has the following design flow. It basically consists of three modules. They are:
1.Object Detection
2.Converting the detected object into Speech
3.Depth Estimation

### 4.1 Object Detection

Object detection is the first module in our project, and it serves as its foundation. It essentially entails detecting near and far objects using datasets for which the model has been trained. Because our application's target users are visually challenged persons, detection is crucial.

### 4.2 Converting the detected object into text

This module consists of turning text to speech, which is critical since it aids blind victims in identifying and analyzing who and what is close by and around them. It assists kids in navigating and understanding what is happening around them.

### 4.3 Depth Estimation

The processes and calculations used to create a depiction of the spatial construction of a scene are known as profundity evaluation or extraction highlight. To put it another way, it's used to figure out how far two objects are apart. Our model is used to assist visually impaired people, and it anticipates giving them ahead notification of any problems that may arise. To accomplish so, we'll look at the distance between the deterrent and the individual in the given situation. A rectangular box is created around the item after it has been identified.

## 5 IMPLEMENTATION

Compatibility with Python and library set up obstacles are necessary for efficient implementation of this Model. To tell you the truth, this was one of the most difficult sections of the project for me to complete. Thank you to stack overflow and Python Unofficial Binary Releases for providing pre-built docs, which you may download from here if your system supports it.

### 5.1 Anchor box

Every grid cell in SSD can have several anchor/earlier containers. These pre-described anchor containers are each responsible for a length and form within a grid cell. SSD use the matching section during training to ensure that the anchor field is properly aligned with the bounding containers of each floor fact item inside an image. The anchor field with the greatest degree of similarity with an object is

responsible for forecasting that item's elegance and proximity. Once a community has been trained, these resources are used to teach the local surroundings as well, in order to anticipate the recognized things and their locations. Smaller Every anchor field has a part proportion and a zoom level specified explicitly. We appreciate, however, that not all of the objects are rectangular. Some are narrower, while others are longer and wider to varying degrees. To account for this, the SSD structure allows pre-defined anchor container component ratios to be used. At each zoom/scale level, the usage of the depth ratio of the anchor connected with each grid cell may be tailored to the outstanding component ratios.

## 5.2 ZOOM LEVEL

The anchor bins do not always have to have the same length as the grid cell. The customer is undoubtedly curious about the location of each smaller or larger object within a lattice cell. The zooms parameter is used to define how a horrendous part of the anchor containers has to be raised or decreased with respect to each network cell.

## 5.3 DEPTH ESTIMATION

Depth estimates or extraction function refers to the procedures and algorithms used to produce a representation of a scene's spatial form. Miles is used to calculate the distance between things in less basic terms. Our prototype is utilised to assist blind people, with the goal of raising awareness of potential obstacles. To do so, we must first determine how far apart the impediment and the individual are in any given moment. Following the detection of an item, a square container is created around it.


**Fig 2:** Distance Approximations

If the object takes up the most space on the body, the approximate distance between the item and the specific individual can be computed by applying a few conditions. The following code is used to decipher objects and return to the gap and location records.

We've built up a TensorFlow consultation based on Crucial Features for Detection right here. As a result, the boxes are used to evaluate the new release as well. Boxes are the innermost element of an array. So, with the new release, we'd want to lay down the terms and conditions.

I am the index of a container in the containers array. The analysis of the container's rating is accomplished using an index. It's also utilized to gain access to classes. The width of the detected item is now determined. This is accomplished by inquiring about an item's width in terms of pixels. We were given the middle of by subtracting and dividing the equal axis start coordinates. The center of our observed rectangle is computed in this way. Finally, A point is drawn in the last row's middle. We expect the object to be recognised if scores[0][i] >= 0.5, which is the default choice for drawing rectangles (i.e., identical or greater than 50 percent ). if score[0][i] is greater than or equal to 0.5.

The mid x in the preceding formula represents the centre of the X axis, while the mid y represents the centre of the Y axis. If the gap between the item and the precise character is apx distance 0> 0.3 and mid x < 0.7, the item is too close to the precise character. This code can be used to calculate the item's relative distance from a specified character. Following item detection, the code is utilised to determine the item's relative distance from the character. If the item is too close, the character receives a sign or a warning via the voice technology module.

### 5.4 VOICE ESTIMATION

It's vital to alert the person about the presence of an object on their route after it's been detected. PYTTSX3 is an important module for the voice era. Pyttsx3 is a Python conversion module that converts text to speech. This module is compatible with other modules. Python versions 2 and 3 are both supported. An application uses the factory feature pyttsx.init to connect to a pyttsx. Engine instance (). Pyttsx3 is a gadget that easily converts text to voice..

## 6 RESULT AND DISCUSSION

### 6.1 TEST

This project improves performance while also allowing for quick output delivery. This server-based project allows you to break down your paintings into components and recognise the central component of any system. This method allows for the development of reliable and enjoyable software from the inside out.

- Initially, we are taking real-time images from the webcam of a blind person's mobile phone, and a connection is established between the phone and the computer, after which the photos are sent to the user.

- The device in the computer will examine it with the aid of its APIs and SSD ALGORITHM, and it will determine the picture's self-assurance accuracy. Positive lessons like books, drinks, and the remote had a 98 percent accuracy rate.

- Following the evaluation of the photographs, we generate an output on a computer-based device, It is then converted to speech using voice modules and transmitted to the blind individual via wireless audio assistance devices. The information is then converted to speech using voice modules and supplied to the blind individual via wireless audio assistance devices.

## 6.2 RESULT

The devices on which it was tested are listed below, and it produced the following result, which was analyzed along with the help matplotlib libraries.


**Fig 3:** The Cup

The end range is 0.3 units from the frame, and a range warning is provided since it is too near, with the speech output noting that it belongs to the cup class. It emits a warning linke cup when it gets too close to the frame.


**Fig 4:** The Remote

The ultimate distance from the frame is 0.8 units, and no distance-based warning is sent since it is at a safer distance, the class recognition voice can be heard, and the class is distant.


**Fig 5:** The Bed

The ultimate distance from the frame is 0.9 units, and there is no distance-based warning because it is at a safe distance; instead, a class identification voice is created, and the object's name is heard as bed.

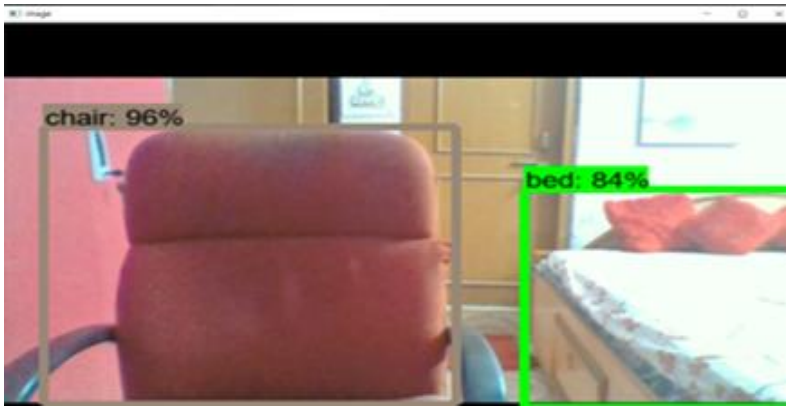


**Fig 6:** The Chaır

Because it is at a safer distance, no distance basd alert is created, and the class recognition voice may be heard as expected. In a single frame, it can identify many objects.


**Fig 7**: The Tv

The ultimate distance from the frame is 0.8 units, and there is no distance-based warning because it is at a safer distance. Instead, a class identification voice is created, and the object's name is heard as tv.

The suggested system successfully recognises 90 items, names them, and indicates their correctness. The version also estimates the distance between the item and the digital digicam and provides audio feedback as the user approaches the item. SSD Mobile net V1 and SSD Inception V2 were used to evaluate the data. The SSD Mobile net V1 version, on the other hand, showed a significant reduction in latency and improved object detection speed

**Table 1**. Objects And Accuracy Level

| S.NO | OBJECTS | ACCURACY LEVEL |
|---|---|---|
| 1 | Cup | 99% |
| 2 | Remote | 98% |
| 3 | Bed | 96% |
| 4 | Chair | 96% |
| 5 | Tv | 96% |
| 6 | Person | 90% |

.

# 7 CONCLUSION

Using machine learning and pre-trained models, we designed a Blind Assistance System that will assist in the detection of objects in the environment. In order to complete this project, we used TensorFlow API, SSD architecture, and COCO Dataset. We used object detection and depth computation to turn the discovered item into speech. There are a variety of uses for this proposed system. It makes it easier for blind people to acquire, analyses, and translate information. Reference or scope in the future the study's purpose is to let visually impaired persons navigate freely so they can move fast while being safe. The device gives the blind person distance and object detection, speech awareness of the object.

We are optimistic that by integrating these services, we will be able to improve the application's effectiveness in the future.

Making this application into a hardware project to expand it. Creating a chatbot that allows the user to converse and engage. Installing GPS to know where you are in real time and making it an all-in-one system. Adding Web Support to it.

# 8 REFERENCES

1.M. E. C. Santos, T. Taketomi, C. Sandor, J. Polvi, G. Yamamoto, and H. Kato, “A usability scale for handheld augmented reality,” in Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology, pp. 167-176, 2014.

2. C. K. Lakde and P. S. Prasad, "Navigation system for visually impaired people," in Proceedings of 2015 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), pp. 0093-0098, doi:

10.1109/ICCPEIC.2015.7259447, 2015.

3. Deepak Gaikwad1, Chaitalee Baje2, Vaishnavi Kapale3 , Tejas Ladage4 “Blind Assist System” , International Journal of Advanced Research in Computer and Communication Engineering,Vol.6 Page No- 2278-1021,march 2017.

4. P. Bose, A. Malpthak, U. Bansal and A. Harsola, "Digital assistant for the blind," in Proceedings of 2017 2nd International Conference for Convergence in Technology (I2CT), pp. 1250-1253, doi: 10.1109/I2CT.2017.8226327, 2017.

5. B.Deepthi Jain, S. M. Thakur and K. V. Suresh, "Visual Assistance for Blind Using Image Processing," in Proceedings of 2018 International Conference on Communication and Signal Processing (ICCSP), pp. 0499-0503, doi: 10.1109/ICCSP.2018.8524251, 2018

6. M. Awad, J. E. Haddad, E. Khneisser, T. Mahmoud, E. Yaacoub and M. Malli, "Intelligent eye: A mobile application for assisting blind people," in Proceedings of 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM), pp. 1-6, doi: 10.1109/MENACOMM.2018.8371005, 2018.

7. S. M. Felix, S. Kumar and A. Veeramuthu, "A Smart Personal AI Assistant for Visually Impaired People,” in Proceedings of 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1245-1250, doi: 10.1109/ICOEI.2018.8553750, 2018.

8. A. Khan, A. Khan and M. Waleed, "Wearable Navigation Assistance System for the Blind and Visually Impaired," in Proceedings of 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), pp. 1-6, doi: 10.1109/3ICT.2018.8855778, 2018.

9. A. Khan, A. Khan and M. Waleed, "Wearable Navigation Assistance System for the Blind and Visually Impaired," in Proceedings of 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), pp. 1-6, doi: 10.1109/3ICT.2018.8855778, 2018.

10. Joe Louis Paul, S. Sasirekha, S. Mohanavalli, C. Jayashree, P. Moohana Priya and K. Monika, "Smart Eye for Visually Impaired-An aid to help the blind people," in Proceedings of 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), pp. 1-5, doi: 10.1109/ICCIDS.2019.8862066, 2019.

11. S. Divya, S. Raj, M. Praveen Shai, A. Jawahar Akash and V. Nisha, "Smart Assistance Navigational System for Visually Impaired Individuals," in Proceedings of 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), pp. 1-5, doi:

10.1109/INCOS45849.2019.8951333, 2019.

12. D. P. Khairnar, R. B. Karad, A. Kapse, G. Kale and P. Jadhav, "PARTHA: A Visually Impaired Assistance System," in Proceedings of 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA), pp. 32-37, doi: 10.1109/CSCITA47329.2020.9137791, 2020.

13. F. Al-Muqbali, N. Al-Tourshi, K. Al-Kiyumi and F. Hajmohideen, "Smart Technologies for Visually Impaired: Assisting and conquering infirmity of blind people using AI Technologies," in Proceedings of 2020 12th Annual Undergraduate Research Conference on Applied Computing (URC), pp. 1-4, doi: 10.1109/URC49805.2020.9099184, 2020.

14. S. S. Mohith, S. Vijay, S. V and N. Krupa, "Visual World to an Audible Experience: Visual Assistance for the Blind And Visually Impaired," in Proceedings of 2020 IEEE 17th India Council International Conference (INDICON), pp. 1-6, doi: 10.1109/INDICON49873.2020.9342481, 2020.

15. M. A. Khan, P. Paul, M. Rashid, M. Hossain and M. A. R. Ahad, "An AI-Based Visual Aid with Integrated Reading Assistant for the Completely Blind," in Proceedings of in IEEE Transactions on Human-Machine Systems, vol. 50, no. 6, pp. 507-517, Dec. 2020, doi: 10.1109/THMS.2020.3027534, 2020.

16. K. Patil, A. Kharat, P. Chaudhary, S. Bidgar and R. Gavhane, "Guidance System for Visually Impaired People," in Proceedings of 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp. 988-993, doi: 10.1109/ICAIS50930.2021.9395973, 2021.

17. https://www.researchgate.net/publication/301198969_A_Heuristic_Self-Adaptive_Medium_Access_Control_for_Resource_Constrained_WBAN_Systems URL Accessed on 12 January 2022.