AI Guidance System for Blind Peoples - Implementation

[1]Digvijay A. Sathe, student, Dept. Computer Engineering, Imperial College of Engineering and Research, Pune.
[2]Pankaj Y. Sarvade , student, Dept. Computer Engineering, Imperial College of Engineering and Research, Pune.
[3]Dhananjay A. Warkari, student, Dept. Computer Engineering, Imperial College of Engineering and Research. pune.
[4]Yogesh Y. putage, student, Dept. Computer Engineering, Imperial College of Engineering and Research, Pune.

---***---

Abstract: Blind people faces many difficulties in there day to day life. The artificial intelligence can be used to build a system that could help these people to understand the surrounding. We have created a system that could help them for becoming more independent. As we have discussed solution to build a system that can help blind peoples in the previous paper "AI guidance system for blind people". This paper will focus on the practical implementation of this system. The system works in offline mode hence it can be used anywhere without a network. This system can detect objects like staircases, person, currency, readable papers, vehicle, door, digs-hips, etc. from the surrounding in offline mode and instruct and controlled by using voice. This increases the ease of use for a blind person.

Technical keywords: AI, ASR, deep learning, IoT.

Introduction

More than three million peoples are blind in the world according to the WHO. These peoples suffer a lot of difficulties in their day-to-day life. They became dependent on others our system helps them to identify some daily interacting objects. This system can identify the objects from the surrounding and notify about it to the user using voice instruction also, the entire system's operations can be understood because of voice outputs generated by the system. The user can give voice instructions to the system to do the operations they want to perform. The system uses various sensors like a camera, Ultrasonic sensor, PIR sensor which increases the operational effectiveness of the system. Hence, it is an interdisciplinary approach of IoT and artificial intelligence. With the use of a single-board computer Raspberry pi 4, we execute the program we have created which detects and inform info about an object to the user. Sensors are handled by separate Arduino Uno r3 and computed signals from it provided to the raspberry pi through the serial port interface which helps to minimize the computational work of raspberry pi.

This paper will mainly focus on: technical solution, implementation, project benefits, methodology, and results.

Scope of the Implementation Review

The Post Implementation Review applies to the:

- Project communication
- Technical solution
- Implementation
- Project dates

Scope and Objectives of the Project

We built a system that can help blind peoples to do identify some objects surrounding them. Which will help them to do their work without any help from others. This system operates in offline mode so large availability of the system is possible.

The main objectives of our project are:

I) Making blind people more independent.

II) Providing object detection of person, readable, door, stairs, currency, and necessary things, etc.

III) providing voice assistance to provide information in form of voice.

IV) providing sensor sensed data to improve project efficiency.

V) Providing all these features in a compact wearable glass.

Project Summary

Project Communication

We used the following communication strategies to co-ordinate work along with all project partners and guide:

- Project Status Reports.
- Project presentations.
- Virtual meetings.
- E-mails.

Technical Solution

2.2.1. Proposed System

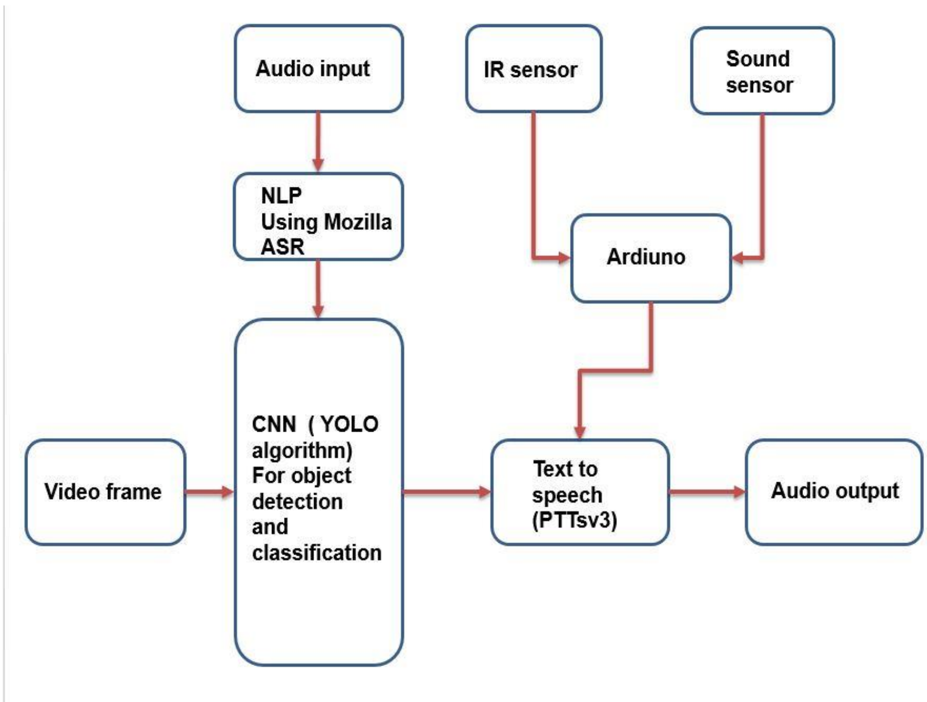


fig. proposed system architecture

The AI guidance system for blind peoples system is mainly based on three main modules. There working as follows:

i) YOLO v3 tiny Detection:

As we are using raspberry pi 4 we have limitations over computational and processing power due to limited RAM, ROM, and CPU power. So we used UOLO v3 tiny that is compact and fast comparing to other algorithms. The detection rate and frame rates increase due to the use of the YOLO algorithm. We trained the algorithm with a custom image dataset which we have created using Google collab, and used its weights and config files to detect objects by loading them into OpenCV.

Working with raspberry pi the cv2.CaputureVideo() does not work with a raspberry pi camera. So we used picamera module to capture video frames. To increase the speed we reduced the size of frames to 320*320.

Working of YOLO v3:

YOLO is single-stage architecture, (You Only Look Once) results in a very fast inference time. The frame rate for 448x448 pixel images was 45 fps i.e. 0.022 s per image on a Titan X GPU while achieving state-of-the-art ***mAP*** (mean average precision). Smaller and slightly less accurate versions of the network also reached 150 fps.

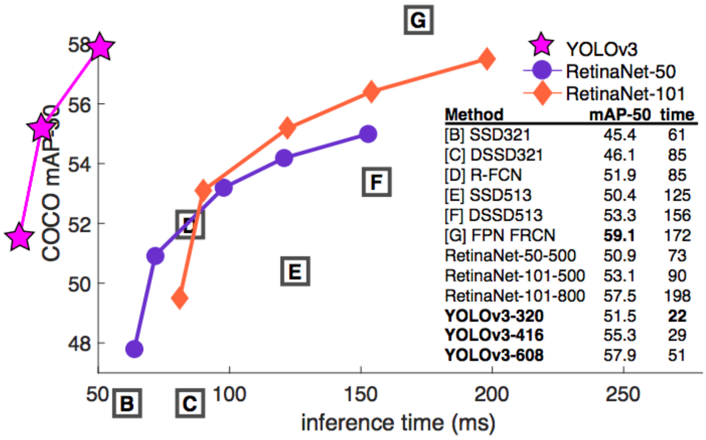


Fig. YOLO v3 timing graph

Inspired by ***ResNet*** and ***FPN*** (Feature-Pyramid Network) architectures, the ***YOLO-V3 feature extractor***, called ***Darknet-53*** (it has 52 convolutions) contains skip connections (like ResNet) and 3 prediction heads (like FPN) — each processing the image at a different spatial compression.

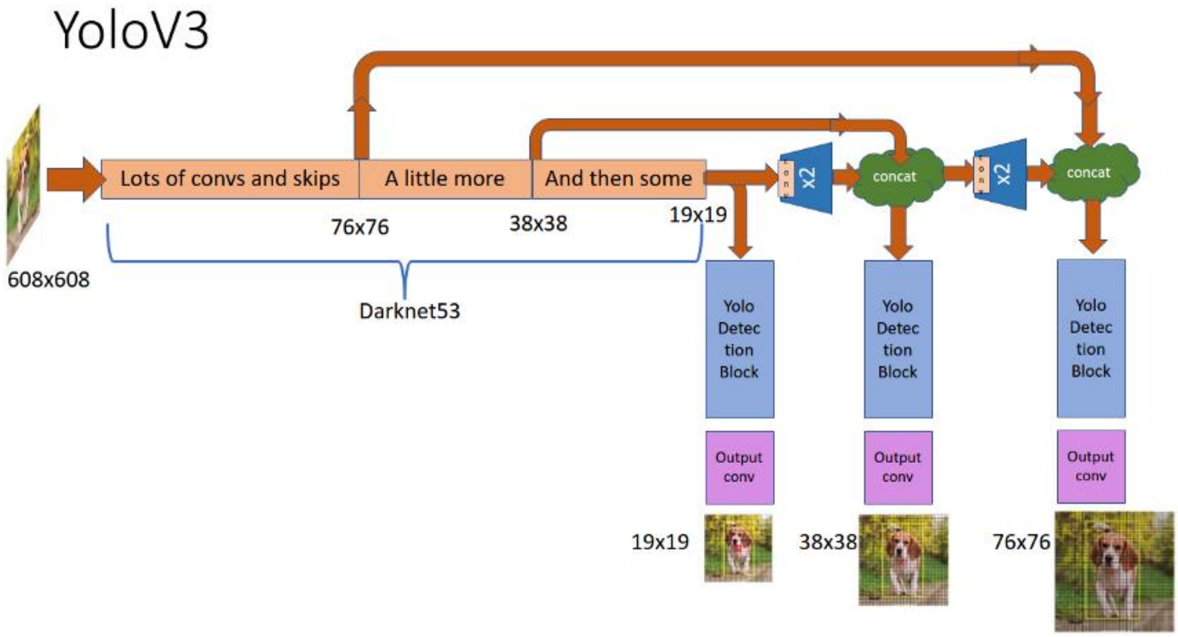


Fig. YOLO-V3 architecture

the FPN topology allows the YOLO-V3 to learn objects at different sizes: The 19x19 detection block has a broader context and a poorer resolution compared with the other detection blocks, so it specializes in detecting large objects, whereas the 76x76 block specializes in detecting small objects. Each of the detection heads has a separate set of anchor scales.

II) Mozilla ASR module:

We created a voice assistant for blind peoples using Mozilla's ASR in offline mode. As we have the aim to implement ASR in offline mode for the high availability of a system. We used pre-trained weights of Mozilla ASR trained on American English and deepspeech module version 0.6.0. and here also raspberry pi requires more time to compute voice captured by pyaudio but pyaudio makes the system slow and not operate accurately so we used SpeechRecognition modules SpeechRecognizer() method to capture voice and Mozilla ASR to process it. Also, we used the pyttsx3 python module to convert text output generated by the system into speech output.

Working of Mozilla ASR:

DeepSpeech v0.6 includes a host of performance optimizations, designed to make it easier to use the engine without having to fine-tune their systems. Our new streaming decoder offers the largest improvement, which means DeepSpeech consistents of low latency and memory utilization, regardless of the length of the given audio being transcribed. Application developers can obtain partial transcripts without worrying about big latency spikes.

DeepSpeech is made up of two main subsystems: an acoustic model and a decoder. The acoustic model od Mozilla's ASR is a deep neural network that receives audio as inputs, and outputs character probabilities. The decoder uses a beam search algorithm to transform the character probabilities into textual transcripts that are then returned by the system to the user.

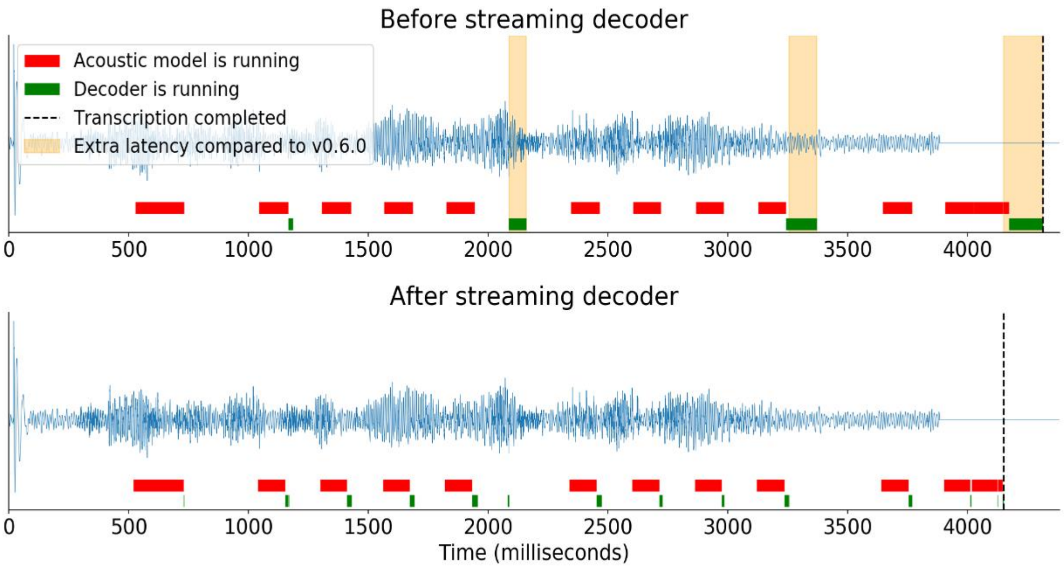


Fig. ASR timing diagram

In the above diagram, you can see the same audio file is being processed by DeepSpeech, before and after the decoder optimizations. The Mozilla ASR program requests an intermediate transcription roughly every second while the audio is being transcribed. The dotted black line marks when the program has received the final transcription. Then, the distance from the end of the audio signal to the dotted line represents how long a user must wait after they've stopped speaking until the final transcript is computed and the application can respond.

III) IoT module:

We created an IoT module using Arduino unor3, PIR, and an ultrasound sensor. This module is integrated with raspberry pi using a serial port connection. It is responsible for the sensing through ultrasound and PIR sensor for the improvement of our system. Arduino is a separate part for IoT operation to reduce the load from raspberry pi. Using an Ultrasound sensor we can calculate the distance of obstacles if any came front. And PIR uses to detect any movement in the range.

Working of IoT module:

Ultrasound sensor:

The Ultra Sonic HC-SR04 emits ultrasound at 40,000Hz. If there is an obstacle in its path, then it collides and bounces back to the Ultrasound sensor.

The formula **distance = speed*time** is used to calculate the distance from Ultrasound sensor.

PIR sensor:

A passive infrared motion sensor also called as a PIR, detects black-body radiation that is temperature, which all objects emit as a function of their temperature relative to absolute zero. The PIR sensor responds to infrared radiation centered around a wavelength of 10μm (10 microns, or 10,000nm). Which is the approximate body temperature of people and animals.

The word **"passive"** in the term **"passive infrared"** refers, which receives infrared radiation passively. Proximity sensors must generate their own infrared radiation actively, which is reflected by nearby objects.

Apart from this module, some hardware-related work had been done. Which includes installation of configuration and OpenCV in raspberry pi, creating a wearable container to hold all things.

Implementation

We used waterfall model to implement this project, the major tasks then divided to each project member to implement like:

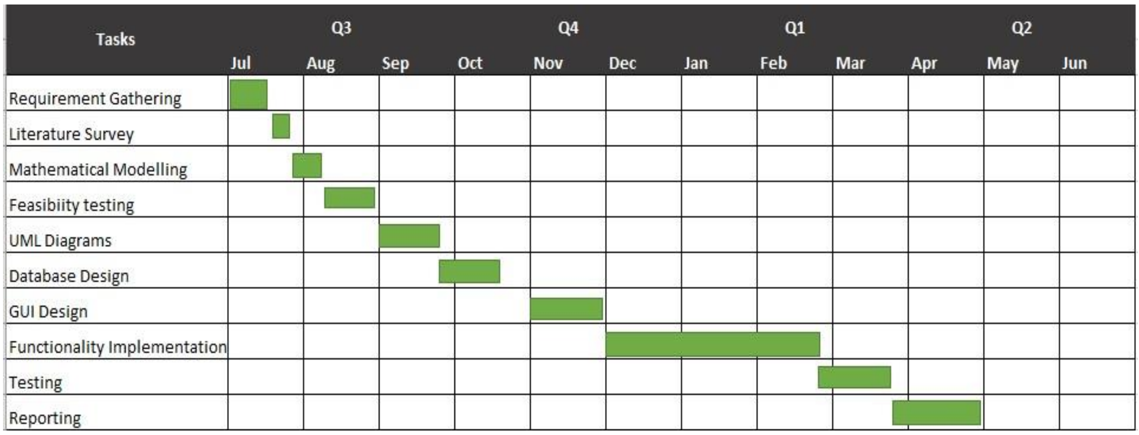
i) YOLO detection.

ii) Voice recognition.

iii) Sensor integration.

Then we integrated all the major tasks to work as one and done testing.

Throughout implementation is as follows:



Project Benefits

List the major benefits that the project is to deliver to your council, library, and community. This may include:

- Replacement of old bulky and expensive system.
- New flexible and sustainable system.
- Offline operations on CPU for higher availability.

- Ease of use due to voice assistance.
- Wearable aid that can detect objects fast and accurately.

Project Methodology

- We used the waterfall model for project implementation.
- Our project tasks were divided across four members of the team.
- the roles and responsibilities of the project team member are provided below:

Project's YOLO model creator (member 1)

The responsibilities were to create training, creating object detection module, and the tasks:

- Creating custom data set for training YOLO Algorithm.
- Training the YOLO algorithm.
- Writing program module for object detection.

Project's Voice Assistance creator (member 2)

The responsibilities were to create a voice recognition module, assistant building, and the tasks:

- Creating offline ASR with help of Mozilla voice recognition.
- building program of voice assistant.

Project's Sensor module Creator (member 3)

The responsibilities were to create sensor IoT module, integration, and the tasks:

- Creating IoT model on Arduino to integrate on raspberry pi.
- Creating program module of Ultrasound and PIR sensor operations.

Project's Tester and Modules Integrator (member 4)

The responsibilities were to integrate, test all modules, and the tasks:

- Integrating all modules
- Testing the working

Guide

- Providing technical advice.
- Assist in the resolution of any technical issues that may arise during the project

Process Improvement

What Went Right... and Why?

- Good Project Management.
- Communication strategies.

- Support of the Project guide.
- Detailed Planning.

What Went Wrong… and Why?

- The incapability of libraries with raspberry pi os.
- Late Delivery of hardware equipments.
- Technical issues with equipment.

Results

We tested our system and come up with the following results:

Fig. YOLO detection

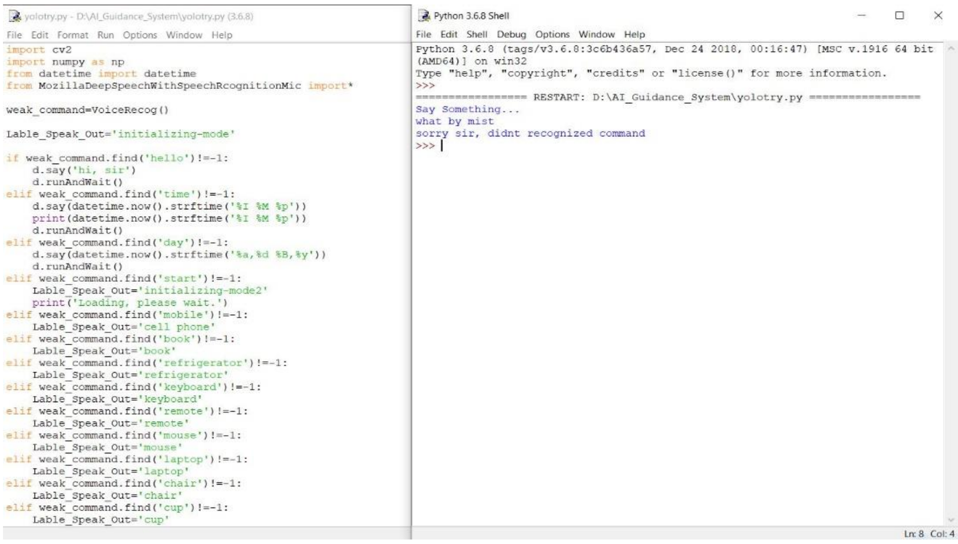


Fig. voice assistant

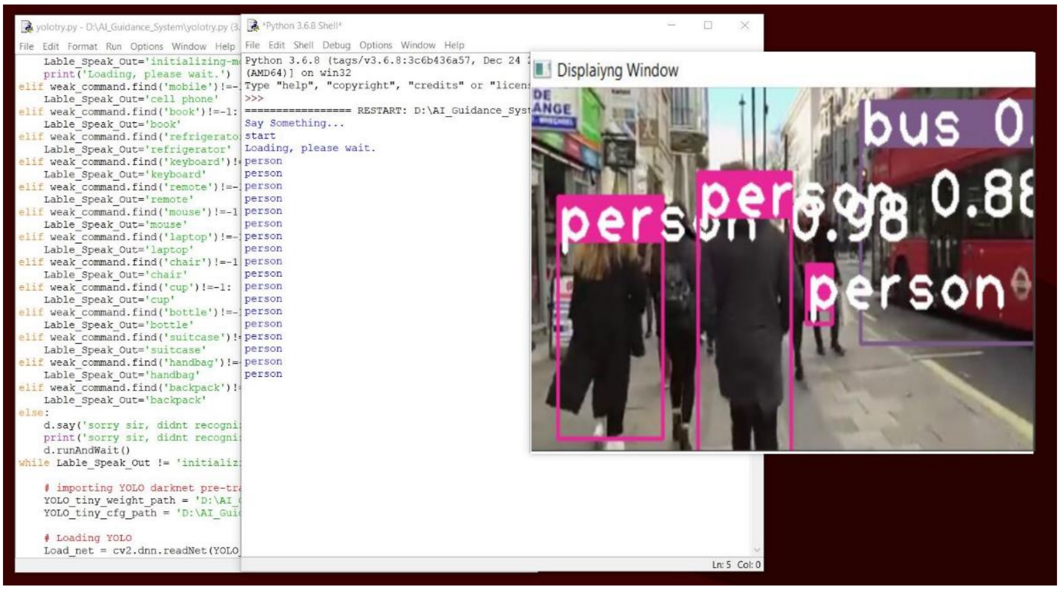


Fig. serial operation of project

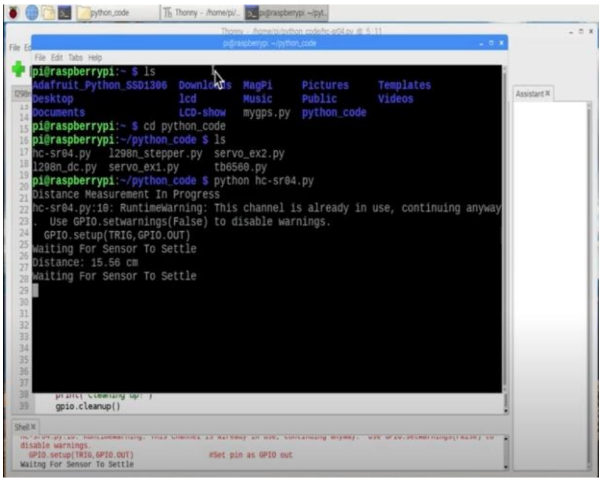


Fig. ultrasound sensor distance

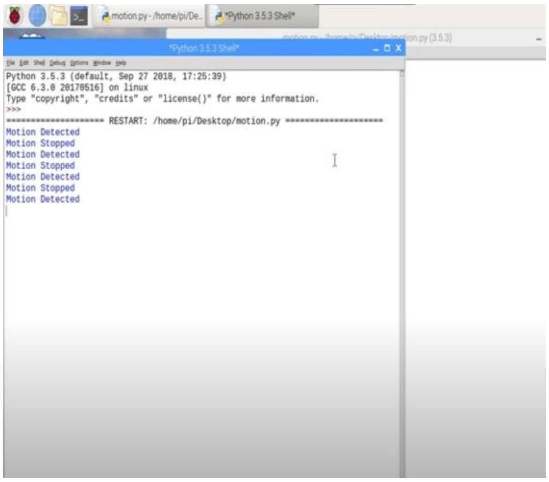


fig. PIR detection

Conclusion

More accurate, low price AI guidance system for blind people is created with higher availability of a system

Future Recommendations

- Building more accurate ASR.
- Reduction in response time (less time complexity).

Acknowledgment

We are very grateful to our Guide Prof. Anup Raut and HOD of Computer Department ICOER, Prof. Nilesh Sabale. for providing us the necessary help and encouragement for successful implementation of our project, Special thanks to friends for supporting us throughout our work.

References

[1] Uncertainty-Aware Visual Perception System for Outdoor Navigation of the Visually Challenged1, Dimitris E. Diamantis, Panagiotis Kalozoumis, and Dimitris K.Iakovidis, 22 April 2020.

[2] Smart Guiding Glasses for Visually Impaired People in Indoor Environment 2. Shiguo Lian, Member, IEEE, Zhaoxiang Liu, Kai Wang, Dijun Liu,2017.

[3] AI guidance system for blind peoples 1.,Digvijay Sathe, Pankaj Sarvade, Yogesh phutage, dhananjay warkari, 2020.

[4] Developing Walking Assistants for Visually Impaired People4: A Review., Muhammad Sheikh Sadi, Senior Member, IEEE, Kamal Z. Zamli, APRIL 15, 2019.

[5] Assistance System for Visually Impaired using AI5 ., Ms. Kavya. S, Ms. Swathi, MimithaShetty,2019.

[6] Staircase Detection to Guide Visually Impaired People6: A Hybrid Approach., Ahsan Habib1 , Muhammad Nomani Kabir2* , Motasim Billah Mredul, 10 August 2019 .

[7] Blind Navigation System Using Artificial Intelligence7., Ankush Chourasia2, 03| Mar-2018.

[8] Outdoor Navigation for Visually Impaired Based on Deep Learning8., Saleh Hadi,2019.

[9] Obstacle and Fall Detection to Guide the Visually Impaired People with Real9-Time Monitoring., · Md. Milon Islam1 · Shishir Ahmmed1 · Saeed Anwar Khan, 27 June 2020.

[10] Navigation System for Blind10 - Third Eye., Samartha Koharwal, Samer Bani Awwad, Aparna Vyakaranam. 5 March 2019.

[11] Business Intelligence and Analytics: Challenges and Opportunities by Prakash K. Ukhalkar, Dr. Rajesh N. Phursule , Dr Devendra P Gadekar, Dr Nilesh P Sable, in International Journal of Advanced Science and Technology Vol. 29, No. 12s, (2020), pp. 2669-2676