VIVEKANANDA EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER ENGINEERING

A SMART HOME NAVIGATION FOR VISUALLY IMPAIRED FOR EASY AND SAFE NAVIGATION

In partial fulfillment of the Fourth Year (Semester–VII), Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai Academic Year 2019-2020

Group Mentor

Dr.Mrs. Sharmila Sengupta

Submitted By

Anish Adnani (D17B - 01)

Nihal Bhandary (D17B - 10)

Supriya Patil (D17B - 50)

Kaif Siddique (D17B - 63)

(2020-21)

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
Department of Computer Engineering

Certificate

This is to certify that ***Anish Adnani, Nihal Bhandary, Supriya Patil, Kaif Siddique*** of Fourth Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on "***A smart home navigator for visually impaired for easy and safe navigation***" as a part of their coursework of PROJECT-II for Semester-VIII under the guidance of their mentor ***Dr. Mrs. Sharmila Sengupta*** in the year 2020-21 .

This thesis/dissertation/project report entitled ***A SMART HOME NAVIGATOR FOR VISUALLY IMPAIRED FOR EASY AND SAFE NAVIGATION*** by ***Anish Adnani, Nihal Bhandary, Supriya Patil, and Kaif Siddique*** are approved for the degree of ***Computer Engineering***.

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7, PO8, PO9, PO10, PO11, PO12 PSO1, PSO2	

Date: 15/05/2021
Project Guide: Dr. Mrs. Sharmila Sengupta

Project Report Approval

For
B. E (Computer Engineering)

This thesis/dissertation/project report entitled ***A SMART HOME NAVIGATOR FOR VISUALLY IMPAIRED FOR EASY AND SAFE NAVIGATION*** by ***Anish Adnani, Nihal Bhandary, Supriya Patil, and Kaif Siddique*** are approved for the degree of ***Computer Engineering***.

Internal Examiner

External Examiner

Head of the Department

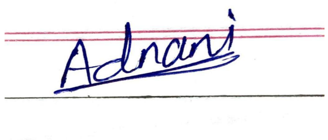
Dr. Mrs. Nupur Giri

Principal
Dr. Mrs.J. M. Nair

Date: 15/05/2021
Place: Mumbai

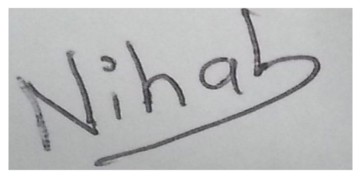
Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.



(Signature)

(Anish Adnani D17B 01)

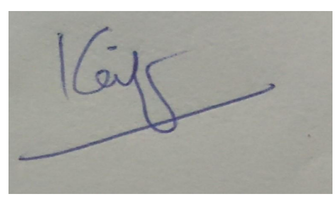


(Signature)

(Nihal Bhandary D17B 10)

(Signature)

(Supriya Patil D17B 50)



(Signature)

(Kaif Siddique D17B 63)

Date:15/05/2021

ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to Assistant Professor **Dr. Mrs. Sharmila Sengupta** (Project Guide) for her kind help and valuable advice during the development of the project synopsis and for her guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr.(Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair ,** for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is a great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

COURSE OUTCOMES FOR B.E PROJECT

Learners will be able to,

Course Outcome	Description of the Course Outcome
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solution for the problem.
CO 4	Able to interpret the data and datasets to be utilized.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one's competency.
CO 11	Able to develop professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

Abstract

There are about 285 million visually impaired people in the world, though they manage to do their daily chores they do face many difficulties. Keeping that in mind the proposed solution provides indoor and outdoor assistance to the visually impaired. Conventionally the visually impaired generally use a cane for proper and hassle-free navigation. The proposed solution aids the visually impaired with a routine task like walking, aid while writing, going up and down the stairs, searching for common objects like keys, brushes, chairs, etc. Voice-based aid would be provided to the visually impaired through speakers and sensors. The ultrasonic sensor helps in recognizing objects within the vicinity that might not be captured by the cameras and produce an alerting sound for the user to avoid any sudden bump.

Certain hardware units using Jetson Nano and Arduino along with ultrasonic sensors are used for navigation and locating certain objects thus helping them manage their daily schedules and improve their activities. The proposed system works offline and can be extended to work outside home premises by making minor changes in hardware and software.

<u>Index</u>

Title **Page no.**

References (papers + books + Patent)

Appendix (Paper I & II Details)

1. Paper I & II Details

a. Paper published
b. Certificate of publication
c. Plagiarism report
d. Project review sheet

2. Competition certificates(if any)

3. Video Presentation Link

Chapter 1: Introduction

1.1 Introduction

Blindness is the inability to see anything, including light. So due to lack of perception of light, the visually impaired cannot identify objects and their placement in the surroundings to have a safe navigation route. Traditionally, the visually impaired use a cane to have safe navigation and to avoid bumping into objects. The aim to develop a navigation system must also include its mobility, cost-efficient nature, user-friendly adaptability. Substantial research is being conducted on a particular topic and every research seems to have setbacks in either of the above-mentioned aspects.

1.2 Motivation

There are about 285 million visually impaired people in the world, though they manage to do their daily chores they do face many difficulties. Keeping that in mind the proposed solution provides indoor and outdoor assistance to the visually impaired. Conventionally the visually impaired generally use a cane for proper and hassle-free navigation. The proposed solution aids the visually impaired with a routine task like walking, aid while writing, going up and down the stairs, searching for common objects like keys, brushes, chairs, etc. Voice-based aid would be provided to the visually impaired through speakers and sensors. The ultrasonic sensor helps in recognizing objects within the vicinity that might not be captured by the cameras and produce an alerting sound for the user to avoid any sudden bump.

1.3 Problem Definition

An indoor AI-based navigator for hassle-free navigation of the visually impaired. An indoor AI-based navigator for hassle-free navigation of the visually impaired. The problem can be broadly classified mainly into two tasks: object detection and measuring distance of objects. So, according to the requirement, we have developed two separate modules for the above-mentioned which make the system to be an all-inclusive, portable, and user-friendly system. The development cost of the system is also far less than the modules currently available in the market. Voice-based aid would be provided to the visually impaired through speakers and sensors. The ultrasonic sensor helps in recognizing objects within the vicinity that might not be captured by the cameras and produce an alerting sound for the user to avoid any sudden bump.

Certain hardware units using Jetson Nano and Arduino along with ultrasonic sensors are used for navigation and locating certain objects thus helping them manage their daily schedules and improve their activities. The proposed system works offline and can be extended to work outside home premises by making minor changes in hardware and software

1.4 Existing Systems

Traditional applications to provide aid to the visually impaired have now existed for many decades without providing any significant support to the concerned bracket.

According to a survey conducted among a group of people, it is estimated that on average the traditionally used cane can only detect up to 67.8% of ground obstacles, which further proves that traditional applications have not been able to completely solve the concerning situation.

There are also various other technological solutions available in the market that cover the specified domain but are either too expensive to be bought by a common man or are not easily available in the market.

One such device, the OrCam device is an artificial vision device that allows the visually impaired to understand the text and identify objects through feedback, describing what they are unable to see. OrCam is priced around 3500$ (2,60,000 INR) which would not be a suitable budget for most in India. Hence the development of a cost-effective tool is necessary

1.5 Lacuna of the existing systems

There are various available solutions for the assistance of the visually impaired, one of them as mentioned is OrCam priced around 2,60,000 INR which would not be a suitable solution for most people, moreover, our solution suits the typical budget range of any common man would generally afford for.

Our solutions require an active internet connection which can also hamper the accessibility range of the device, as it is restricted in the area of the active network. Our system is completely based on offline modules.

1.6 Relevance of the project

The target audience of the project, visually impaired, by this system, will aid the visually impaired to walk freely in the nearby surroundings by giving them information about various objects, placement of the objects, alerting them in case of sudden bumping, all of which will be informed to the end user via headphones.

Finding a particular article can be easily done by the device, the system also provides directions to reach the required object without bumping into any obstacle.

Chapter 2: Literature Survey

A. Brief Overview of Literature Survey

There have been numerous studies conducted using different sensors,algorithms and models for assisting the blind person in navigation. We found that either some of them required the internet, a good processor, or their surroundings needed to be mapped beforehand. The existing systems also lacked in fulfilling all the three features we found necessary in the mobility of a blind or visually impaired person i.e distance estimation, Orientation and obstacles and hazards.

B. Related Works

2.1 Research Papers referred.

- ### A Machine Vision Based Navigation System for the Blind:

  Shang Wenqin, Jiang Wei and Chu Jian, "A machine vision based navigation system for the blind," 2011 IEEE International Conference on Computer Science and Automation Engineering, 2011, pp. 81-85, doi: 10.1109/CSAE.2011.5952638

  Abstract: Blindness aid system has been part of the blind's life. In this study, we present a new computer vision based blindness navigation system, using robust and stable computer vision technologies to strengthen the weakness of existing electronic intellectual aid devices. Our system can offer state information of the road scene, and assist the blind to find out the obstacles and the walkable planar regions, through which the blind can walk independently. The real scene experiments show that our method can provide a robust cognitive result of the road scene.

- ### Drishti: an integrated indoor/outdoor blind navigation system and service:

  L. Ran, S. Helal and S. Moore, "Drishti: an integrated indoor/outdoor blind navigation system and service," Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. Proceedings of the, 2004, pp. 23-30, doi: 10.1109/PERCOM.2004.1276842.

Abstract: There are many navigation systems for visually impaired people but few can provide dynamic interactions and adaptability to changes. None of these systems work seamlessly both indoors and outdoors. Drishti uses a precise position measurement system, a wireless connection, a wearable computer, and a vocal communication interface to guide blind users and help them travel in familiar and unfamiliar environments independently and safely. Outdoors, it uses DGPS as its location system to keep the user as close as possible to the central line of sidewalks of campus and downtown areas; it provides the user with an optimal route by means of its dynamic routing and rerouting ability. The user can switch the system from an outdoor to an indoor environment with a simple vocal command. An OEM ultrasound positioning system is used to provide precise indoor location measurements. Experiments show an in-door accuracy of 22 cm. The user can get vocal prompts to avoid possible obstacles and step-by-step walking guidance to move about in an indoor environment. This paper describes the Drishti system and focuses on the indoor navigation design and lessons learned in integrating the indoor with the outdoor system.

- **Design and Development of an Indoor Navigation and Object Identification System for the Blind:**

Andreas Hub, Joachim Diepstraten, and Thomas Ertl. 2003. Design and development of an indoor navigation and object identification system for the blind. *SIGACCESS Access. Comput.*, 77–78 (Sept. 2003 - Jan. 2004), 147–152.

Abstract: In this paper we present a new system that assists blind users in orienting themselves in indoor environments.We developed a sensor module that can be handled like a flashlight by a blind user and can be used for searching tasks within the three-dimensional environment. By pressing keys, inquiries concerning object characteristics, position, orientation and navigation can be sent to a connected portable computer, or toa federation of data servers Providing models of the environment.Finally these inquiries are acoustically answered over a text-to-speech engine.

- ## ISAB: Integrated Indoor Navigation System for the Blind:

Abu Doush, S. Alshatnawi, A. Al-Tamimi, B. Alhasan and S. Hamasha, "ISAB: Integrated Indoor Navigation System for the Blind," in Interacting with Computers, vol. 29, no. 2, pp. 181-202, March 2017, doi: 10.1093/iwc/iww016.

Abstract: Abstract Generally, indoor navigation is considered as a challenging task, especially when people navigate through an unfamiliar place (e.g. a university or a mall). It is even a more challenging endeavor for the visually impaired and blind community. This paper presents an innovative approach to the precise indoor navigation challenge for the blind individuals using a multi-tier solution with the help of an intuitive smartphone interface. We utilize a set of different communication technologies (WiFi, Bluetooth and radio-frequency identification) to help users reach an object with high accuracy. As a proof of concept, we deploy a fully functional testbed and we evaluate our entire solution inside our university library by helping blind users find a specific book. Our results demonstrate the high accuracy of the proposed system to reach an object with accuracy up to 10 cm. The intuitive smartphone interface provides step-by-step navigation voice instructions of the least hazardous path for the blind users while minimizing the cognitive load on their short-term memory. In addition, we show that our iterative improvements on our smartphone's interface has improved the system's efficiency and its accuracy in reaching specific objects successfully.

- ## Voice assisted navigation system for the blind:

A. Noorithaya, M. K. Kumar and A. Sreedevi, "Voice assisted navigation system for the blind," *International Conference on Circuits, Communication, Control and Computing*, 2014, pp. 177-181, doi: 10.1109/CIMCA.2014.7057785.

Abstract: The visually impaired are at a considerable disadvantage because they often lack the information for avoiding obstacles and hazards in their path. They have very little information on self-velocity, objects, direction - which is essential for travel. Previously developed navigation systems use costly equipment which is often not affordable by the common blind community. The navigation systems available are heavy and very complicated to operate. This research has been aimed at design and development of a smart and intelligent cane which helps in navigation for the visually impaired people. The navigator system will detect an object or obstacle using ultrasonic sensors and gives audio instructions for guidance. The signals from the ultrasound sensor are processed by a microcontroller in order to identify sudden changes in the ground gradient and/or an obstacle in front. The algorithm developed gives a suitable audio instruction depending on the duration of ultrasound travel which in turn is made available by an mp3 module associated with the system. This work presents a new prototype of a navigation system on a cane which can be used as a travel aid for blind people. The product developed is light in weight, hence, does not cause fatigue to the user. This project is developed by keeping in view the affordability and reliability. An obstacle as close as 4cm can be detected by this module. A resolution of 15cm of obstacle distance has been designed and achieved. This system can also detect potholes on the path.

- **Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants:**

dos Santos, A.D.P., Medola, F.O., Cinelli, M.J. *et al.* Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants. *Univ Access Inf Soc* 20, 93–103 (2021). https://doi.org/10.1007/s10209-020-00712-z

Abstract: Visually impaired individuals often rely on assistive technologies such as white canes for independent navigation. Many electronic enhancements to the traditional white cane have been proposed. However, only a few of these proof-of-concept technologies have been tested with authentic users, as most studies rely on blindfolded non-visually impaired participants or no testing with participants at all. Experiments involving blind users are usually not contrasted with the traditional white cane. This study set out to compare an ultrasound-based electronic cane with a traditional white cane. Moreover, we also compared the performance of a group of visually impaired participants ($N = 10$) with a group of blindfolded participants without visual impairments ($N = 31$). The results show that walking speed with the electronic cane is significantly slower compared to the traditional white cane. Moreover, the results show that the performance of the participants without visual impairments is significantly slower than for the visually impaired participants. No significant differences in obstacle detection rates were observed across participant groups and device types for obstacles on the ground, while 79% of the

hanging obstacles were detected by the electronic cane. The results of this study thus suggest that electronic canes present only one advantage over the traditional cane, namely in its ability to detect hanging obstacles, at least without prolonged practice. Next, blindfolded participants are insufficient substitutes for blind participants who are expert cane users. The implication of this study is that research into digital white cane enhancements should include blind participants. These participants should be followed over time in longitudinal experiments to document if practice will lead to improvements that surpass the performance achieved with traditional canes.

- **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications:**

Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile VisionApplications", 2017,eprint:1704.04861,arXiv.

Abstract: We present a class of efficient models called MobileNets for mobile and embedded vision applications. MobileNets are based on a streamlined architecture that uses depth-wise separable convolutions to build light weight deep neural networks. We introduce two simple global hyper-parameters that efficiently trade off between latency and accuracy. These hyper-parameters allow the model builder to choose the right sized model for their application based on the constraints of the problem. We present extensive experiments on resource and accuracy tradeoffs and show strong performance compared to other popular models on ImageNet classification. We then demonstrate the effectiveness of MobileNets across a wide range of applications and use cases including object detection, finegrain classification, face attributes and large scale geo-localization.

2.2 Patent Search

1. The patent "Portable blind aid device" provides a blind aid device consisting of:

I. An image capturing module for capturing the images of surrounding environment of the user

II. An image analysis module attached with the above one for

(i) detect moving objects from the images obtained,

(ii) identify spatial relationships of the moving objects, and

(iii) analyze the one or more images obtained from the blind aid device to classify the identified relationships of the moving objects to predefined moving object data.

III. A conversion module for converting the information gathered from the device to audio commands for the end user;

IV. A bidirectional communication module for audio commands to the blind person to notify the blind person of one or more occurrences determined by the blind person as actionable .

The system also provides a computer-readable medium which stores programmable instructions configured for being executed by at least one processor for performing the above processes.[8]

2. “APPARATUS FOR ASSISTING VISUALLY IMPARED PERSONS TO IDENTIFY PERSONS AND OBJECTS" describes a system made of a headset with a camera mounted on top.

The system apparatus for assisting visually impaired persons comprises : a headset; a camera mounted on the headset; a microprocessor communicating with the camera for receiving the images by the camera and converting the images to an audio signal ; and a speaker communicating with the processor for outputting the audio signal. The camera is in communication with a processor that processes either video signals or still photographs output by camera. The processor itself is described as being made of two modules : an image to text module, which recognises and identifies people or objects; a text to speech module which decides what and which audio commands to be output to the user.[9]

3. “Assistance device for blind and partially sighted people” is a wireless system[16] to assist visually impaired people comprising voice-activated portable devices comprising communicating means for transmitting and receiving data to and/or from a network; scanning means for scanning an object selected by a user; memory means for storing information regarding a scanned object; sensor and identification means for locating an object previously scanned; journey means for planning a route to a destination selected by the user and for identifying the correct transport for travel to a selected destination, wherein the sensor identification and journey means are able to communicate with the wireless system through the communicating means.

The patent describes a device comprising a laser scanning technology, mobile service facility, voice recognition facilities, memory, data acquisition and storage. The device can be used in “smart homes” or emergency response services. Moreover, the device could be used in shopping and supermarkets for automated check-out for blind people.

it is configured as a voice-activated, portable device to assist visually impaired people comprising: communicating means for transmitting and receiving data to and/or from a network; scanning means for scanning an object selected by a user; memory means for storing information regarding a scanned object; sensor and identification means for locating an object previously scanned.[10]

2.3 Inference Drawn

The systems described in the papers and patents studied in literature survey provide a variety of ways to guide a user. The proposed system makes use of cameras and ultrasonic sensors to achieve the same goal. The object detection module in existing systems are slower and take time to process the real time data.

Further, apart from the object detection and alerting the user about the object’s presence, the system does not expand on assisting the user with other scenarios where a person might need guidance in interacting with the surrounding environment. That could include finding desired objects from the surrounding environment; or guiding the user in performing activities like turning on lights, TV, etc. With the above deficiencies in the existing system, these particular areas are where the future system could develop more on providing much needed help to the users. Another important factor inferred from the study is ease of use for the user. The existing system still needs assistance in terms of setting up and running it for the user. True self-reliance could be achieved if the future systems are made easily accessible to the end user and are fairly non-complicated in terms of use for a visually impaired person.

2.4 Comparison with existing system

	Paper[1]	**Paper[2]**	**Paper[3]**	**Paper[4]**	**Paper[5]**	**Proposed System**
Internet required	No	Yes	Yes	No	No	No
Processor Used	An embedded computer vision processor	Xybernaut MA IV with a Pentium 200MHz processor	JVC MP-XP7250	Smartphone	Microcontroller	Jetson Nano.
Sensors	Stereo cameras.	DGPS, camera, and ultrasonic sensors.	Camera and ultrasonic sensor.	Bluetooth, Wifi and RFID.	Ultrasonic sensor.	Ultrasonic sensor and a camera.
Working environments	Outdoor	Indoor and outdoor	Indoor and outdoor	Indoor	Indoor and outdoor	Indoor and outdoor

Chapter 3: Requirement Gathering for the proposed system

3.1 Introduction to Requirement Gathering

Requirements gathering is one of the most essential parts of any project and adds value to a project on multiple levels. When it comes to smaller budgets, tighter timelines, and limited scopes, exact documentation of all the project requirements become crucial. Requirements gathering is easier said than done, it is generally an area that is given far less attention than it needs. Many projects start with basic lists of requirements only to find out down the line that many of the customers' needs may not have been fully understood and implemented.

3.2 Functional Requirements

- Object Detection
- Generating commands to move towards a specific object
- Object distance calculation
- Alerting user to avoid sudden bumping (i.e when obstacle_distance < safe_distance)
- Providing speech output about the environment

3.3 Non-Functional Requirement

- Performance: The performance of the system should reach to provide desirable results to the user. The accuracy of the detection model must reach par 80%
- Maintainability: The system requires no to low maintainability, it only requires the battery to be charged at regular time intervals
- Scalability: Currently the system is designed to work best in indoor environments but the model can be extended to provide support in outdoor environments with less chaos

3.4 Hardware, Software, Technology and tools utilized

Hardware Specifications:

- Arduino UNO R3 Microcontroller
- Ultrasonic Sensor
- Jetson nano
- Web cameras and Pi cameras
- Speakers
- 32GB SD card
- Battery
- Buzzer

Software Specifications:

- Torchvision
- Matplotlib, NumPy and Keras modules
- Google collab
- Arduino IDE
- GoogLeNet, SSD MobileNet, AlexNet

Tools:

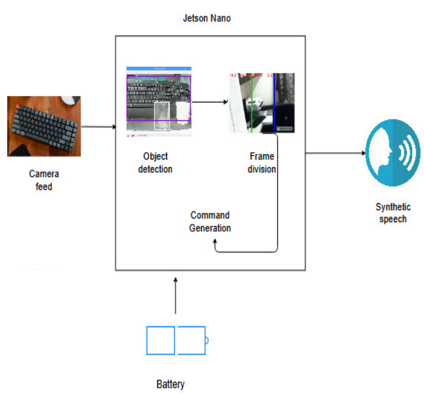
- Python
- PyTorch
- OpenCV
- Tensorflow

3.5 Constraints:

- Insufficient memory for training on a larger dataset or employing other algorithms. Processing power is another limiting factor.
- Objects encroaching or present outside the camera vision range, i.e backside, cannot be worked upon.
- The battery must be sufficiently charged before using the system to avoid unexpected shutting down

Chapter 4: Proposed Design

4.1 Block diagram of the system:



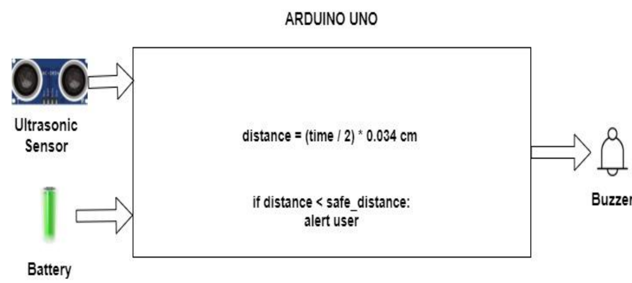


Fig1. Block diagram of module 1

Fig2. Block diagram of module

4.2 Modular design of the system:

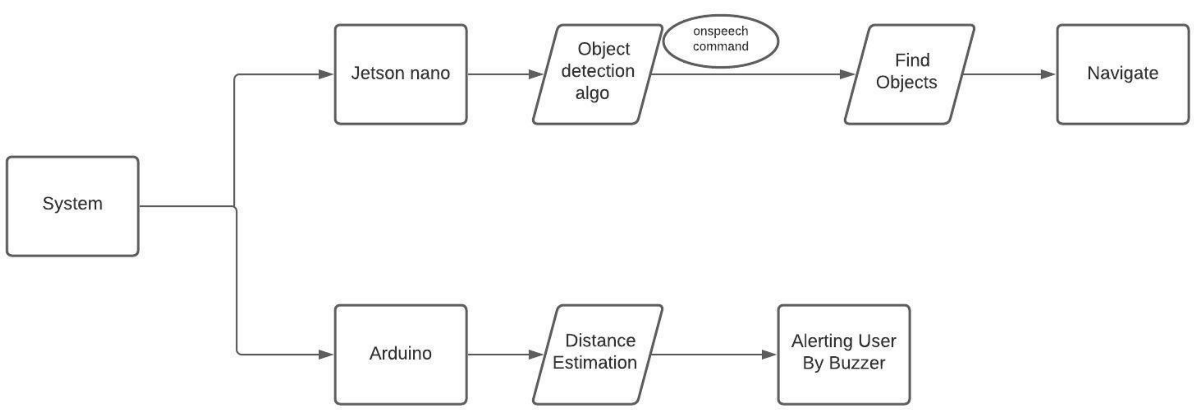


Fig 3. Modular Design

4.3 Detailed Design

DFD 0:

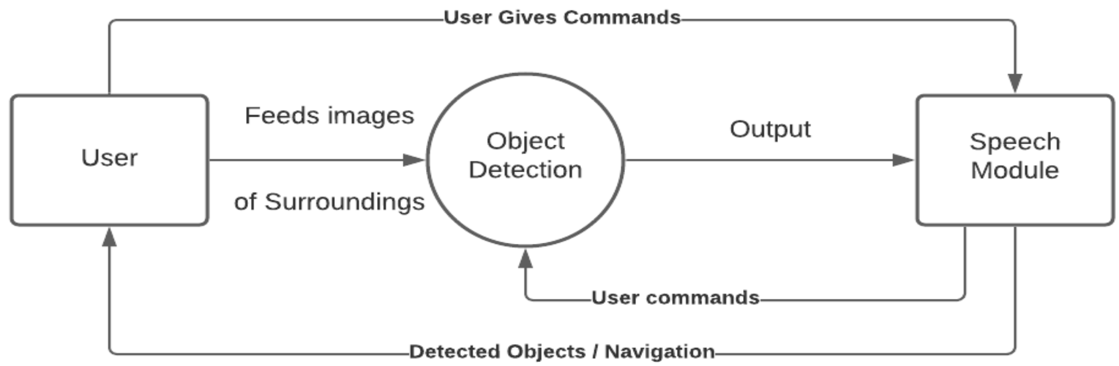


Fig 4. DFD 0

DFD 1:

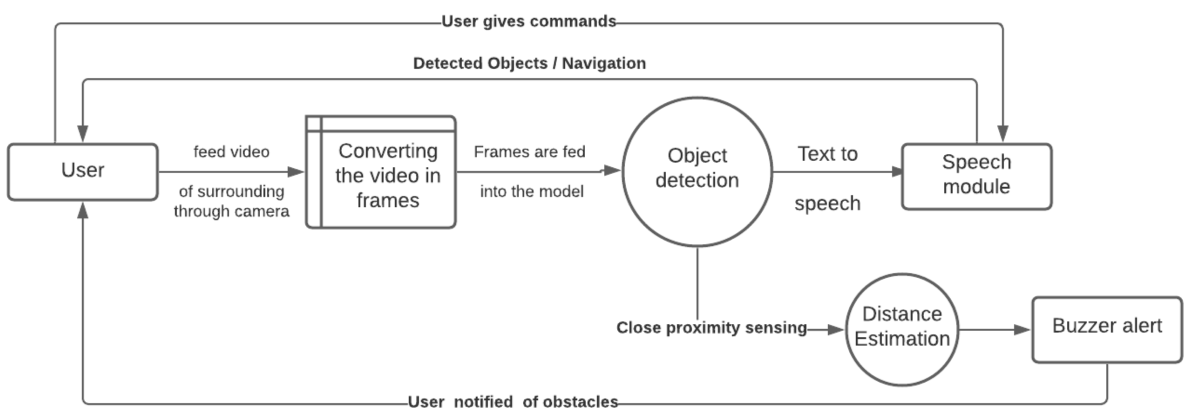


Fig 5. DFD 1

DFD 2:

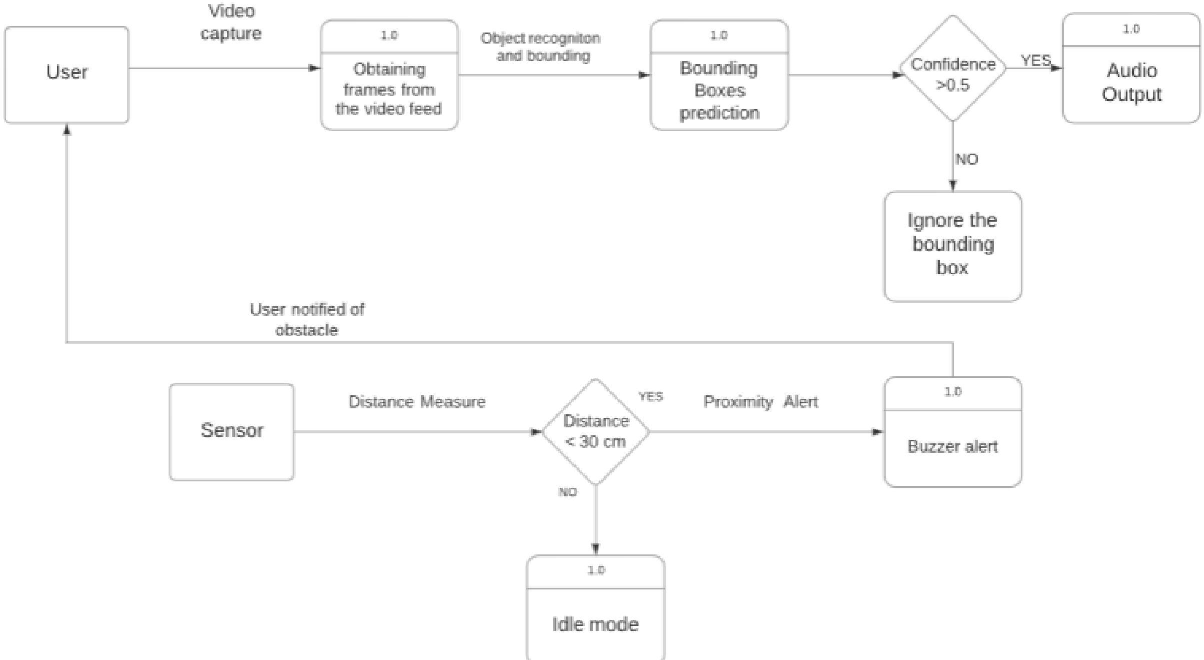


Fig 6. DFD 6

4.4 Project Scheduling & Tracking using Timeline / Gantt Chart

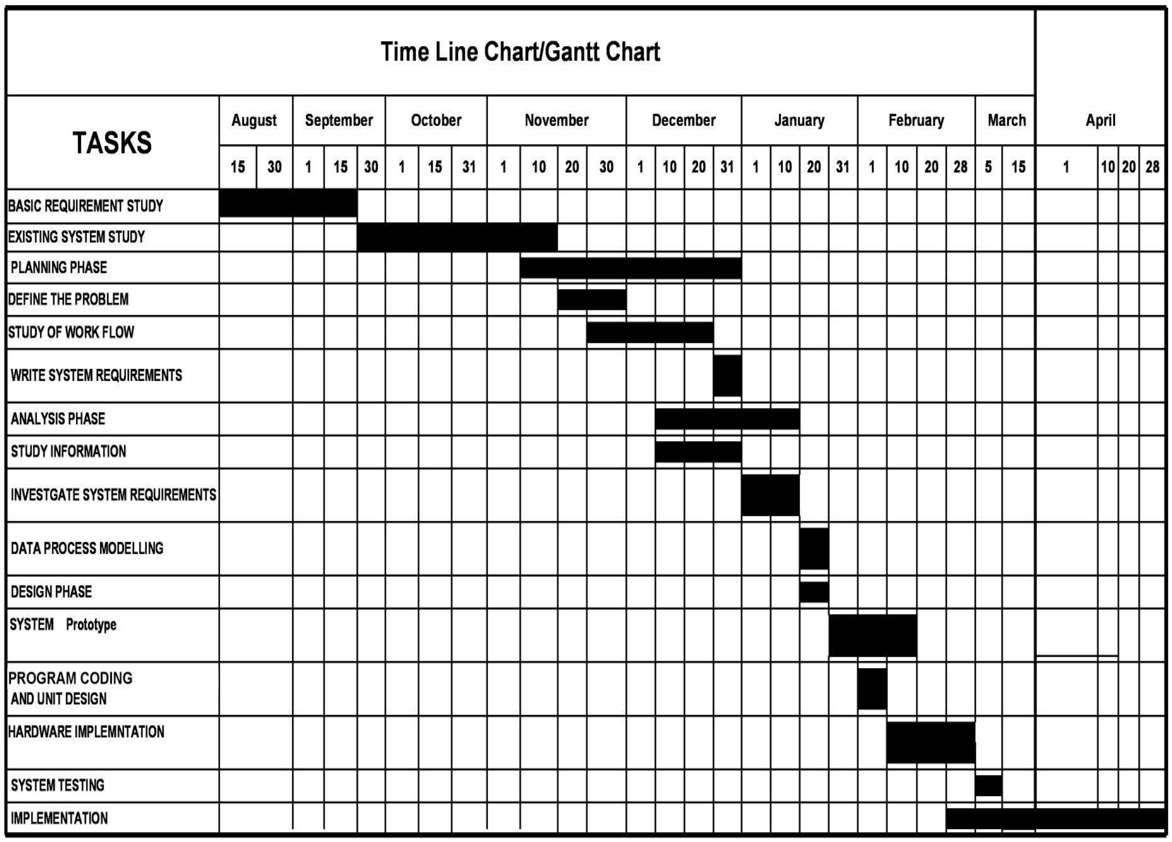


Fig 7: Gantt chart

Chapter 5: Implementation of the proposed system

5.1 Methodology employed for development

The system consists of two independent modules namely; Jetson based module and an Arduino-based system.

Input to the Jetson Nano computing board is provided through cameras which are attached at the eye level of the user, to provide an exact First Person view of the environment. Output computed by the processor is conveyed to the user via headphones.

Moreover, Arduino is mounted with an ultrasonic sensor to calculate the object distances, if any object is in nearby proximity and may obstruct the path of the user, the user is alerted via a buzzer.

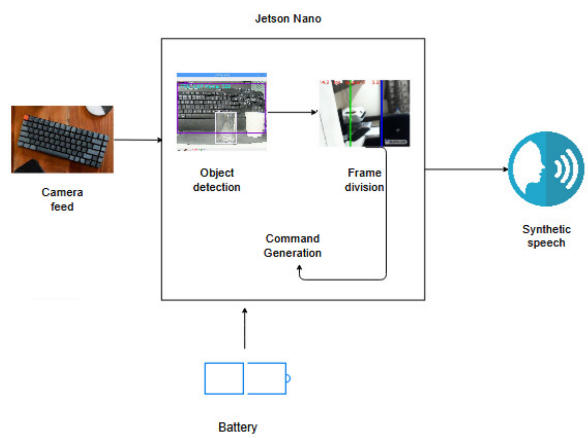


Fig 8: Block diagram of Module1 [Jetson supported]

The above Fig. shows the block diagram of the proposed Jetson-based system. The cameras are attached to spectacles to provide a first-person view of the surrounding. The cameras capture the live surrounding and provide the real-time input to the processor [Jetson Nano], the incoming input is divided into frames which undergo frame division horizontally (into 3 parts), which will be used to identify the coordinates of the object and in turn generate commands to reach that specific object (coordinate). These frames are then available for the object detection model to make predictions.

Following are the list of classes to the object is predicted into:

["person","hat","backpack","umbrella","shoe","eyeglasses","handbag","tie","suitcase","bottle","cup","fork","knife","spoon","bowl","apple","banana","orange","carrot","chair","couch","bed","mirror","desk","door","laptop","mouse","remote","keyboard","cellphone","microwave","oven","refrigerator","clock","vase","scissors","toothbrush"]

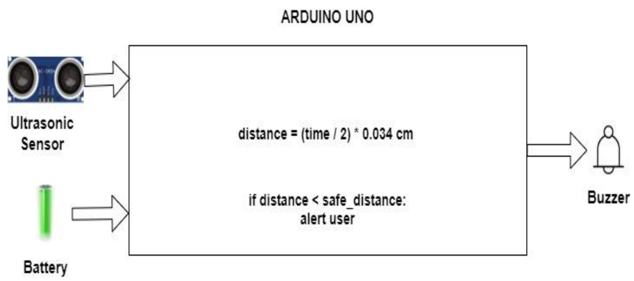


Fig 9 : Block diagram of Module2 [Arduino supported]

The Arduino board mounted with the ultrasonic sensor and buzzer is then fixed in a closed container to provide a complete mobile system that can be directly used by the user.

According to the results obtained, the system works best and provides optimal results when attached to the feet of the user.

The system implemented can be divided into various parts performing separate tasks as mentioned follows

Object Detection Model

For the purpose of object detection, we tried out various models namely, Yolo V5, Yolo V4, GoogleNet, SSD MobileNet, etc. Following are the results obtained.

Model Name	Accuracy	FPS
YOLO V5	~84%	1-2
YOLO V4	82%-84%	2-3
SSD MOBILE NET V2	80%-83%	16-20

GOOGLE NET	75%-79%	10-12
ALEXNET	~80%	10-12

As deduced from the above-mentioned table, the "**SSD MOBILE NET V2**" model gives the best combination of accuracy and frames per second.

Object Detection and Recognition

Before providing the data for object detection it is necessary to prepare the data in a specific format

Data preparation can be divided into a series of steps mentioned below:

- The incoming video stream is broken down into frames.
- Frames are then put into a buffer
- Input images of the buffer are resized to 300 x 300 pixels

The image frame thus obtained by resizing from the previous stage is then inserted into the model. The frame goes through several layers of the MobileNet V2 model. The model has different types of layers such as depthwise separable layer, batch normalization layer, ReLu, and pointwise layer. In the diagram below, the filter shape and input size of the layers of the model are listed. Once the frame reaches the softmax layer, classification takes place and all the objects detected are classified and then released to the program.

Obstacle Avoidance

The other main component of the system is the obstacle avoidance part. It consists of a microcontroller (the Arduino UNO board) attached to an ultrasonic sensor. The sensor works by transmitting ultrasonic waves and receiving those waves. The information about the time of transmission and reception of ultrasonic waves are recorded and notified in real-time to the Arduino UNO board. The Arduino UNO through the program calculates the distance of objects which made the waves pingback. This distance obtained is then used to determine if the user is in close proximity to any obstacle at foot level. The user is then immediately notified by the buzzer if the obstacle is deemed within the threshold distance.

Generating Commands

Once the user has expressed his/her object of interest, the speech-to-text module transcribes the object and then alerts the program about the object. If the aforementioned object is in the list of objects, the Jetson Nano then proceeds to detect, recognize and guide the user to the object.

For navigation, the goal of the program will be to guide the user towards the object in such a way that the object is directly in front of the user. In order to do so, the video frame is divided into three parts: the left, the center, and the right. Once the object is detected and recognized, it has to determine which part the object lies in. Depending on the section of the frame, the system guides the user by generating appropriate commands.

After the object is detected, its center coordinates are captured. These coordinates are then used to determine which section of the frame the object lies on. If it lies on the left part, the command generated will be to “move left” and vice versa. Once the object is in the center zone, the user is notified about the presence of the object in front of him/her.

In case the object is detected and the user moves away, thereby object exiting the camera frame. The coordinates that are captured when the object was foremost detected are used to determine which part of the screen it lies in before its exit. If the object was last found in the left frame before it went out of frame, the user is then guided to turn left in order to bring the object in the frame.

As seen from the figure a camera is attached at the eye level of the user to provide the most accurate first-person view of the surroundings. The camera is connected to the processor, which recognizes the objects present in the surroundings and the user can select a specific item to get navigation commands to the respective coordinate.

5.2 Algorithms and flowcharts of the respective modules developed

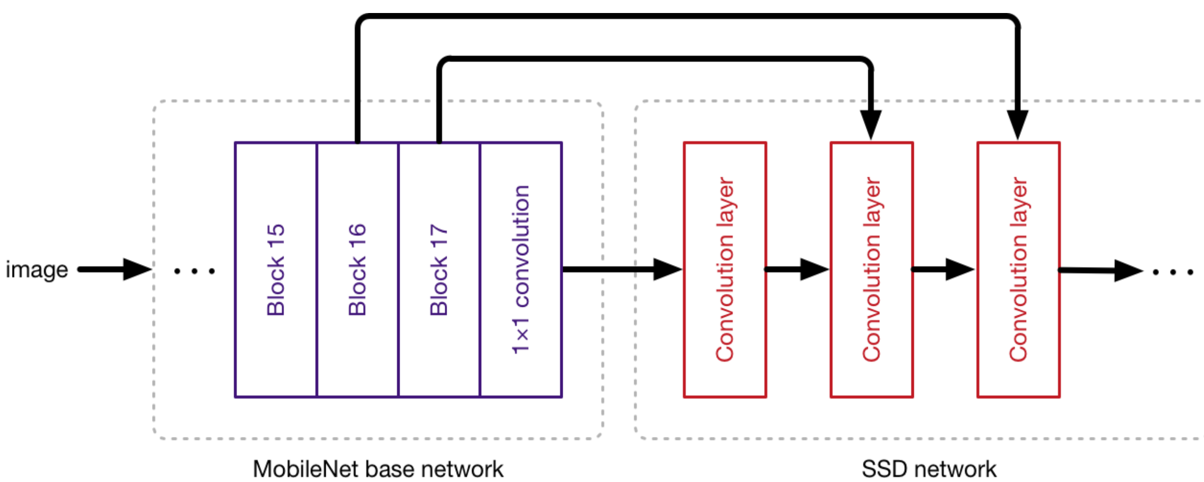


Fig. 10 SSD MobileNet v2

5.3 Dataset Source and utilization

SSD MobileNet V2 is a type of single staged object detection model which is trained on an MS-COCO dataset consisting of over 300,000 images and over 80 classes.

Non-linearities are removed from the intermediate layers, also lightweight depthwise convolution is used.

Chapter 6: Testing of the proposed system

6.1 Introduction to testing

Software testing is a process of executing a program or application with the intent of finding the software bugs.

Softwares is tested to ensure that quality is maintained with required results being obtained in an iterative manner, in which there are developmental phases and testing phases.The contents of a software system test plan may vary from organization to organization or project to project. It depends how we have created the software test strategy, project plan and master test plan of the project.
It can also be stated as the process of validating and verifying that a software program or application or product:

- Meets the business and technical requirements that guided it's design and development
- Works as expected
- Can be implemented with the same characteristic.

The definition of Software testing can be broken into the following parts:

1) **Process:** Testing is a process rather than a single activity.
2) **All Life Cycle Activities:** Testing is a process that take place throughout the Software Development Life Cycle (SDLC).
3) **Static Testing:** It can test and find defects without executing code. Static Testing is done during the verification process. This testing includes reviewing of the documents (including source code) and static analysis.
4) **Dynamic Testing:** In dynamic testing the software code is executed to demonstrate the result of running tests. It's done during the validation process. For example: unit testing, integration testing, system testing, etc.
5) **Planning:** We need to plan as what we want to do. We control the test activities, we report on testing progress and the status of software under test.
6) **Preparation:** We need to choose what testing we will do, by selecting test conditions and designing test cases.

7) **Evaluation:** During evaluation we must check the results and evaluate the software under test and the completion criteria, which helps us to decide whether we have finished testing and whether the software product has passed the tests.

8) **Software products and related work products:** Along with the testing of code the testing of requirement and design specifications and the related documents like operation, user and training material is equally important.

Our system was tested to give an optimal combination of accuracy and frames per second. We tried out various models with placing the camera at various angles and heights (eye level, above head, etc) to obtain optimal results and to find which position covers the most first person view.

The proposed system was tested in indoor and outdoor environments with different objects in the view. The obstacle sensor was tested with the user running into walls and pillars. Chair was used as the desired object sought by the user, and the system then guided the user to the desired object.

After the overall effectiveness of the proposed system is tested, we focus on minor subtle aspects of the proposed system. Frame per second (FPS) rate of the system is tested with various models employed to see which provides the most FPS rate. This test was also taken in consideration of the accuracy, object demarcation and real time feedback factors.

6.2 Types of tests considered

The proposed system consists of two main modules namely : object detection module and obstacle avoidance module. We separately tested both the modules in order to fine tune any errors or difficulties faced by the user.

Using OpenCV, all the performance metrics along with object demarcation is shown in real time on the attached monitor. This was done to help with the evaluation of the models we tested. The model evaluation and testing was crucial in order to select the best and most appropriate model for the system.

In order to make the best use of the sensors available, the position of the sensors were tested. The view of the camera from different locations, angles was tested. This was done in order to achieve user’s comfort as well as improving the field view of the system

The distance for alerting the user of impending collision with an obstacle was also determined by testing various distances by including the user's response time, delay in alerting the user,etc.

6.3 Various test case scenarios considered

The proposed system was tested in both indoor and outdoor environments. Since the goals of the system changes according to the type of environment the user is situated. Both the scenarios were tested to obtain the results and efficiency of the system in indoor and outdoor environments.

To achieve maximum performance and user comfort the cameras were placed on top of the headphones as well as on the glasses. The ultrasonic sensors were placed on different angles on foot to get the correct orientation required for regular working of the proposed system.

The models with high accuracy but low FPS were tested in order to see if FPS was to be sacrificed for higher accuracy would it improve the user experience. The models with high FPS rate but bad accuracy were also considered and tested to observe the final outputs of the system.

Other testing scenarios included testing what happens after and before the desired object has been located, if the object unexpectedly exits the frame of the camera how to re-guide the user towards the object.

6.4 Inference drawn from the various test cases

During model evaluation, we learned that it was beneficial to the user if the user is made aware of the object asap even if the object is not recognized. This led us to select models with high FPS rate and moderate accuracy. Though accuracy was further improved by the use of GPU for processing the images.

The final positioning of the camera was determined after testing several positions. The difference in the view of the system was not limited or constrained in any way so the most comfortable position was selected to place the camera.

The distance between the obstacle and sensor which we inferred from our testing cases is that 45 cm is the appropriate distance where the user can be notified of an impending obstacle and enough response time to evade it if the user is maintaining the regular walking speed.

Another important inference we obtained from testing the proposed system in indoor and outdoor settings is that, in indoor settings the obstacle avoidance is used more than in outdoor settings. And that the indoor environment is where the most help was required for the user.

Chapter 7: Results and Discussion

7.1 Screenshots of the UI for respective module

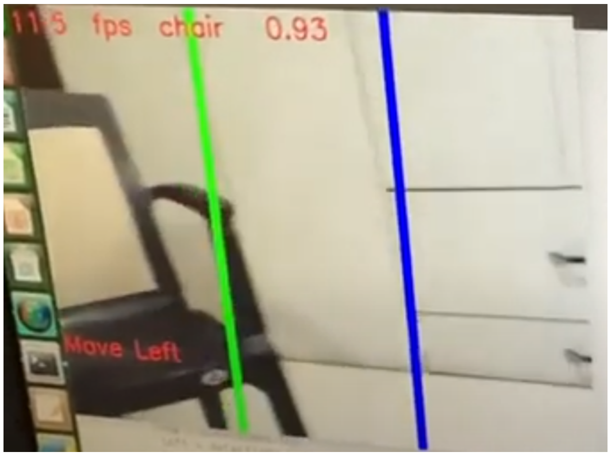


Fig 11 : Command generated by Jetson Nano

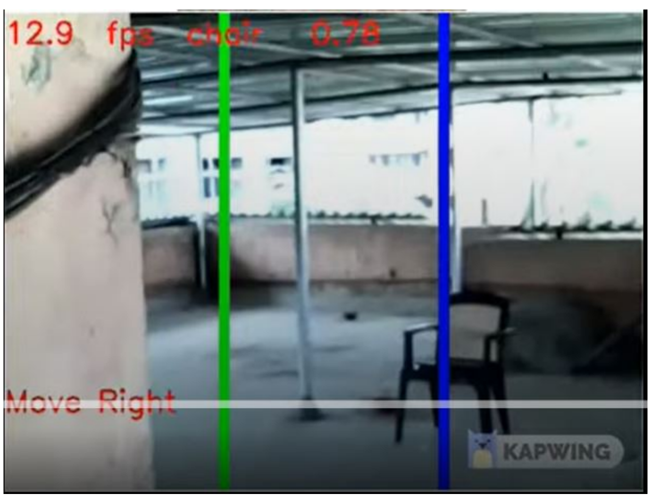


Fig 12 : Command generated by Jetson Nano

7.2 Performance Evaluation measures

Performance of our system is mainly determined by 2 factors :

- Object Detection accuracy
- Frames Per Second (FPS)

To find out the best suited model to give the best combination of accuracy and frames per second (fps), we tried to use various models and maintained a table of accuracy and fps of all the models, and we observed that SSD MOBILENET V2 gave best performance

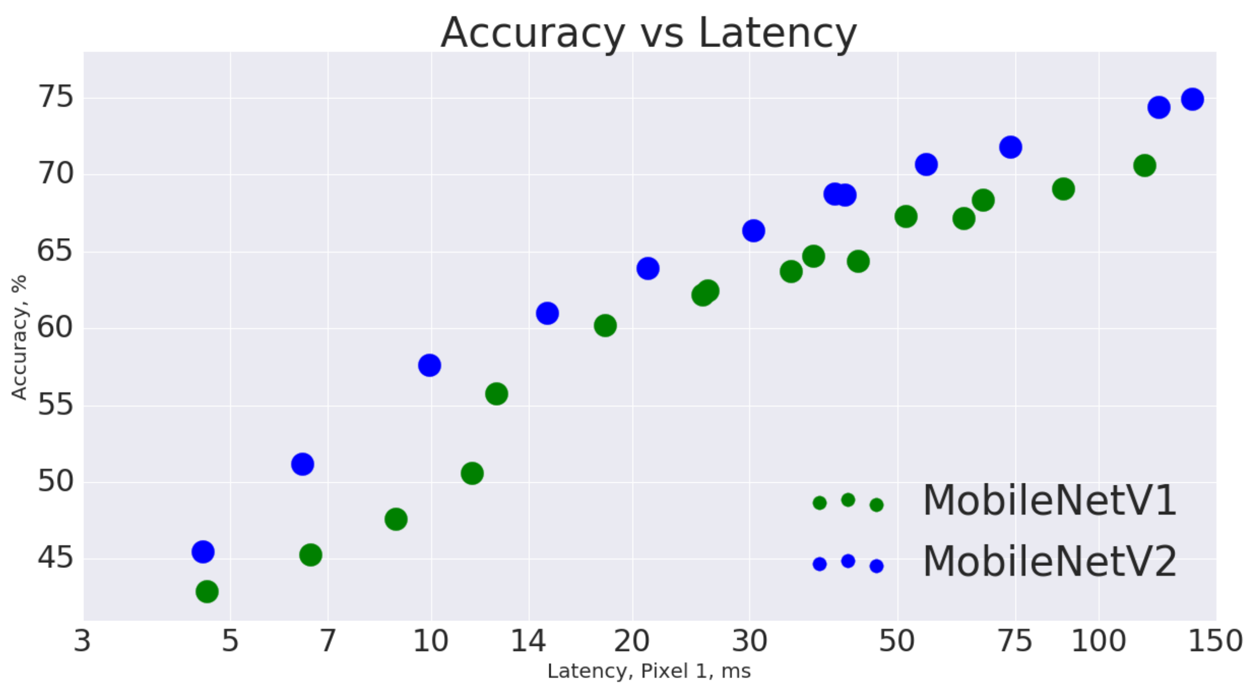
7.3 Input Parameters and Features considered

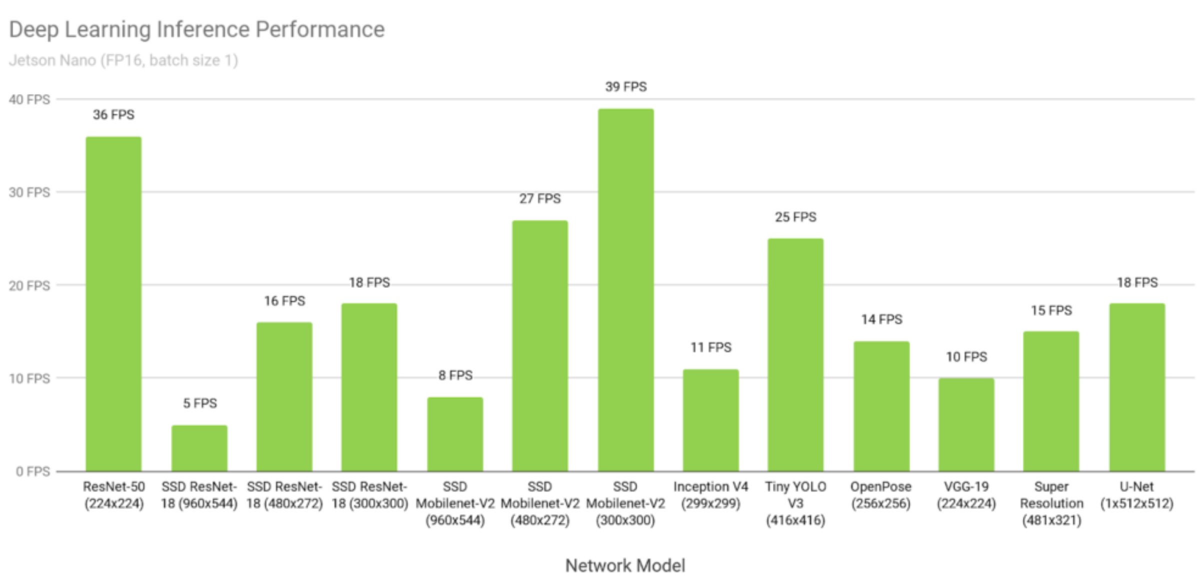
Input parameters to our Module 1(Jetson Nano supported) is input from the web camera. After that the nano process the surrounding. Also the Speech input is provided by the user to find the object he is searching for.

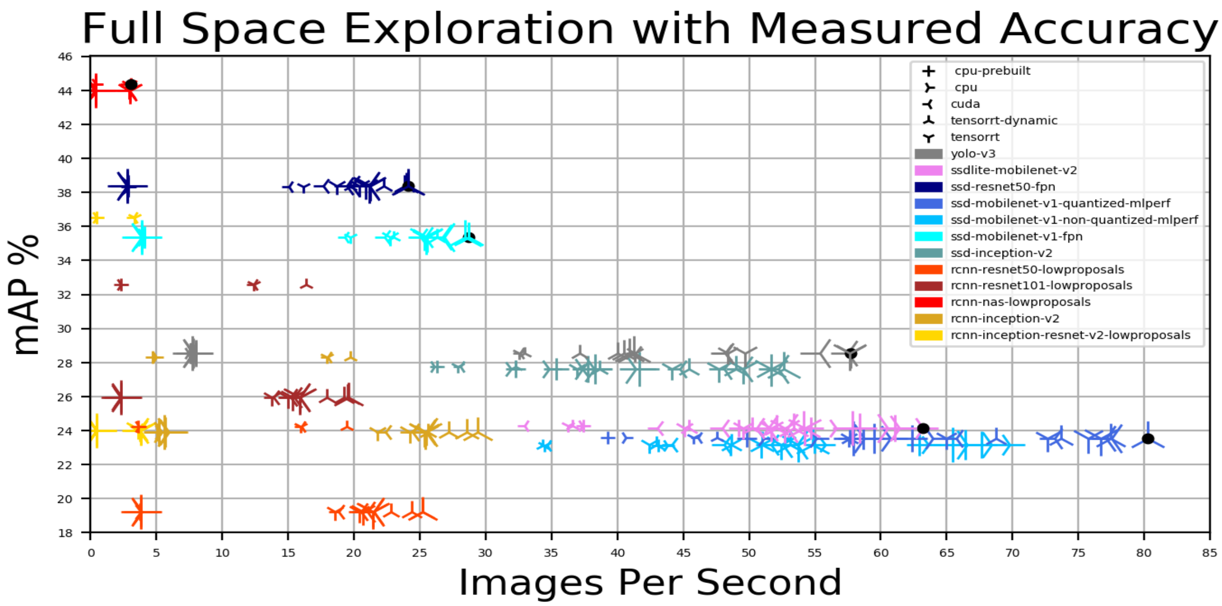
Features that are considered are:

- Object Detection
- Generating commands to move towards a specific object
- Object distance calculation
- Alerting user to avoid sudden bumping (i.e when obstacle_distance < safe_distance)
- Providing speech output about the environment

7.4. Graphical and statistical output







7.5 Comparison of results with existing systems

Accuracy of our system ~80%-85% is comparable to the existing systems currently present in the market

Frames per second (fps) 15-20fps is also one of the optimal conditions that can be obtained for such a complex running system

7.6 Inference Drawn

During model evaluation, we learned that it was beneficial to the user if the user is made aware of the object asap even if the object is not recognized. This led us to select models with high FPS rate and moderate accuracy. Though accuracy was further improved by the use of GPU for processing the images.

Various Models were tested and we first started with the yolov5 model which has better accuracy but it was not working well with jetson nano as it was providing 4-5 fps which was not sufficient and the model consumes more memory. So we came with yolov5 tiny which is a 25mb model which works well when there are single and stable objects it also provides more fps of about 12-15 fps but the accuracy was not good. Then we use SSD mobileNet v2 which works quite well with our model and gives accuracy of 80-85% compared to previous models. Also frames per second was 15 - 20 fps which works quite well.

Chapter 8: Conclusion

8.1 Limitation:

SSD mobilenet V2 is a powerful model which has adequately made the use of all the processing power available to it. The FPS and accuracy could be improved by a huge margin if a powerful processing machine was available, but the cost of such a system would also further increase the cost of the system.

8.2 Conclusion:

Significant improvements in the detection and frame rate are achieved with the use of GPU. The system not only assists the visually challenged in making them aware of the objects in their vicinity but also can now aid in finding and guiding towards any object of interest the user desires. After the specified object is detected, the system helps in orienting the user towards the object and aligns him/her in front of the aforementioned object.

The system makes efficient use of the available GPU processing power and memory. Some of the constraints in the proposed system are overcome with the use of parallel processing and a lightweight object detection module without compromising the frame rate or accuracy. There can be a further increase in the accuracy and FPS rate if a machine with large processing power or more cores is to be used instead of the existing GPU.

In consideration of the overall cost of such a system and the processing power of Jetson Nano, the proposed system is able to achieve the goals satisfactorily.

8.3 Future Scope

- The smart home indoor navigation system can further be implemented to work out in outdoor scenarios as well
- If provide Wifi module support, one can also add road maps navigation commands on the jetson module
- To provide 360 Degree support, one can attach 2 cameras at the back to gain better view of the surroundings
- Further, various other models like VGG-16, VGG-19, RCNN, etc can be used to improve accuracy

References:

1. Shang Wenqin, Jiang Wei and Chu Jian, "A machine vision based navigation system for the blind," 2011 IEEE International Conference on Computer Science and Automation Engineering, 2011, pp. 81-85, doi: 10.1109/CSAE.2011.5952638.

2. L. Ran, S. Helal and S. Moore, "Drishti: an integrated indoor/outdoor blind navigation system and service," Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. Proceedings of the, 2004, pp. 23-30, doi: 10.1109/PERCOM.2004.1276842.

3. Andreas Hub, Joachim Diepstraten, and Thomas Ertl. 2003. Design and development of an indoor navigation and object identification system for the blind. SIGACCESS Access. Comput., 77–78 (Sept. 2003 - Jan. 2004), 147–152. DOI:https://doi.org/10.1145/1029014.1028657

4. I. Abu Doush, S. Alshatnawi, A. Al-Tamimi, B. Alhasan and S. Hamasha, "ISAB: Integrated Indoor Navigation System for the Blind," in Interacting with Computers, vol. 29, no. 2, pp. 181-202, March 2017, doi: 10.1093/iwc/iww016.

5. A. Noorithaya, M. K. Kumar and A. Sreedevi, "Voice assisted navigation system for the blind," International Conference on Circuits, Communication, Control and Computing, 2014, pp. 177-181, doi: 10.1109/CIMCA.2014.7057785.

6. Dos Santos, A.D.P., Medola, F.O., Cinelli, M.J. et al. Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants. Univ Access Inf Soc 20, 93–103 (2021). https://doi.org/10.1007/s10209-020-00712-z

7. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile VisionApplications”, 2017,eprint:1704.04861,arXiv.

[8] Eugene Evanitsky, “Portable blind aid device” ,US 8,606,316 B2

United States Patent and Trademark Office, 10 December 2013.

[9] Ronald L. Klein,James A. Kutsch, JR., “APPARATUS FOR ASSISTING VISUALLY IMPARED PERSONS TO IDENTIFY PERSONS AND OBJECTS AND METHOD FOR OPERATION THEREOF “,US 2013/0046541 A1,
United States Patent and Trademark Office, 21 February 2013.

[10] PEY TOM, “ASSISTANCE DEVICE FOR BLIND AND PARTIALLY SIGHTED PEOPLE”, WO2008015375A1, The International search report, 7 February 2008.

PAPER 1 DETAILS

PAPER 1 PUBLISHED IN 11th INTERNATIONAL CONFERENCE ON COMPUTER COMMUNICATION AND INFORMATICS 2021

PAPER 2 DETAILS

PAPER 2 SUBMITTED FOR REVIEW IN ICICI 2021

ICICV-2021 submission 256 External Inbox ×

ICICV-2021 <icicv2021@easychair.org>
to me

Dear authors,

We received your submission to ICICV-2021 (2nd International Conference on Innovations in Computational Intelligence and Computer Vision):

Authors : Anish Adnani, Supriya Patil, Nihal Bhandary, Kaif Siddique and Sharmila Sengupta
Title : Netra: AI Navigator for the visually impaired
Number : 256

The submission was uploaded by Anish Adnani <2017.anish.adnani@ves.ac.in>. You can access it via the ICICV-2021 EasyChair Web page

https://easychair.org/conferences/?conf=icicv2021

Thank you for submitting to ICICV-2021.

Best regards,
EasyChair for ICICV-2021.

Video Presentation Link:

https://drive.google.com/drive/folders/1LW0PyPcE1ik3bLndo4nGsRaX08C4Cwas?usp=sharing