Audio Assist for the Blind

by

Akshay J Rai	(USN : 4JK19EC006)
Aman Sasidharan K P	(USN : 4JK19EC008)
Amrita Sinchana	(USN : 4JK19EC009)
Yashas Shetty K	(USN : 4JK19EC063)

Under the Guidance of

Mrs. Deepashree A P

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
A J INSTITUTE OF ENGINEERING AND TECHNOLOGY
A Unit of Laxmi Memorial Education Trust ®
(Approved by AICTE, New Delhi, Affiliated to VTU, Belagavi, Recognized by Govt. of Karnataka)
Kottara Chowki, Mangaluru-575006, Karnataka

2022-2023

CONTENTS

- Objective
- Introduction
- Block Diagram
- Methodology
- Result
- Application
- Conclusion
- Progress

OBJECTIVE

- Audio Assist for Blind is wearable navigation system based on live image recognition and obstacles perception used as an audio assistance for blind people.
- The prototype is enriched with information obtained in real time sensors.
- A map lists these points and indicates the distance and direction between closer points.
- The blind users wear glasses built with sensors like The Raspberry Pi Camera. It is a high quality 8-megapixel Sony IMX219 image sensor custom designed add-on board for Raspberry Pi.
- The user navigates freely in the prepared environment identifying the free path.

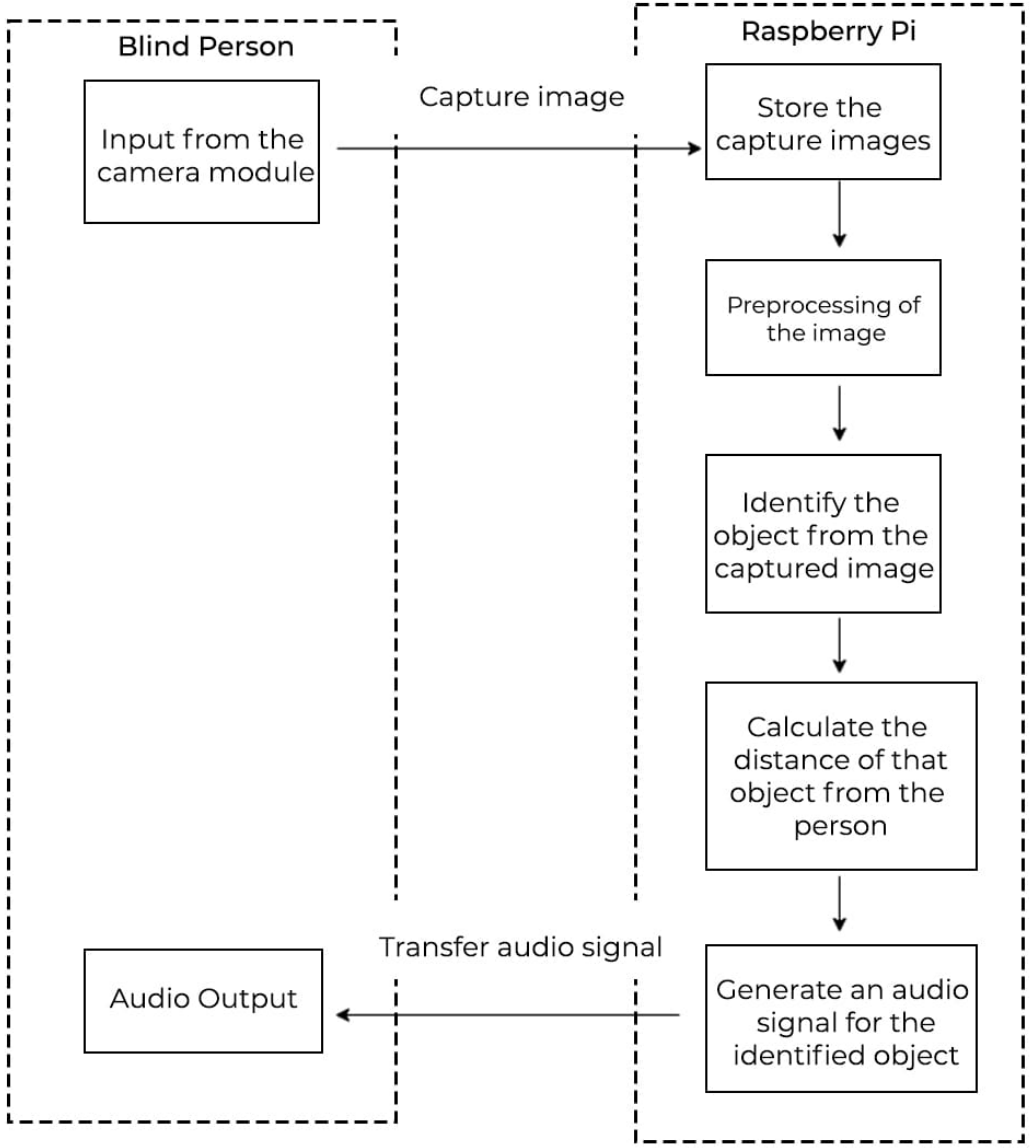
INTRODUCTION

- Blind people may lose perception and have a higher risk of falling or colliding but people need to move whether at home, work or addressing meeting. Most of blind people depend on other human for movement and environmental sensitivity. Blind people are suffering from lot of hardships in their daily life.

- The affected ones have been using the traditional white cane for many years which although being effective still has a lot of disadvantages.

- One such try from our side is that we came up with an Integrated Machine Learning System which allows the Blind Victims to identify and classify Real Time Based Common day-to-day Objects and generate voice feedbacks and calculates distance which produces warnings whether he/she is very close or far away from the object.

BLOCK DIAGRAM



METHODOLOGY

Object Detection & Recognition

- Object detection and perception is used to identify or detect the objects in a frame of a video sequence

- OpenCV is the huge open-source library for computer vision, machine learning, and image processing and now it plays a major role in real-time operation for providing real-time output

- So to perform distance calculation between sensor and object in frame we make use of basic mathematics that is shown in the figure below

METHODOLOGY

The basic principle of working:

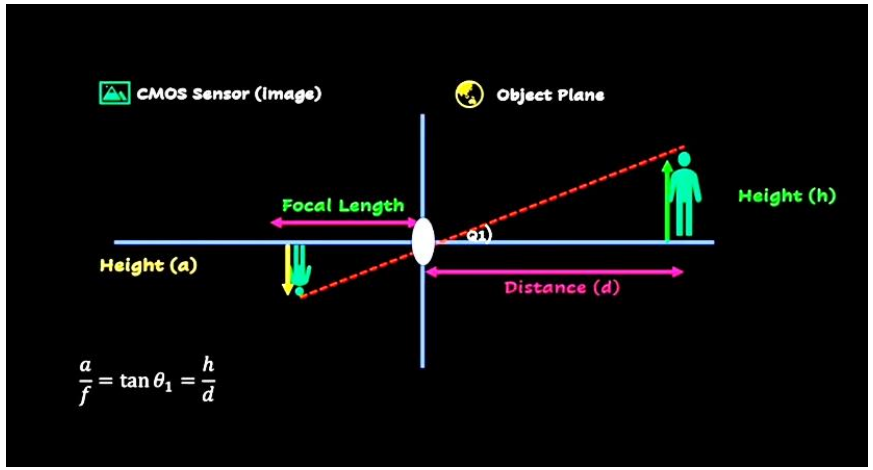
$$\frac{a}{f} = \tan\theta_1 = \frac{h}{d} \quad \longrightarrow \text{Equation No: 1}$$

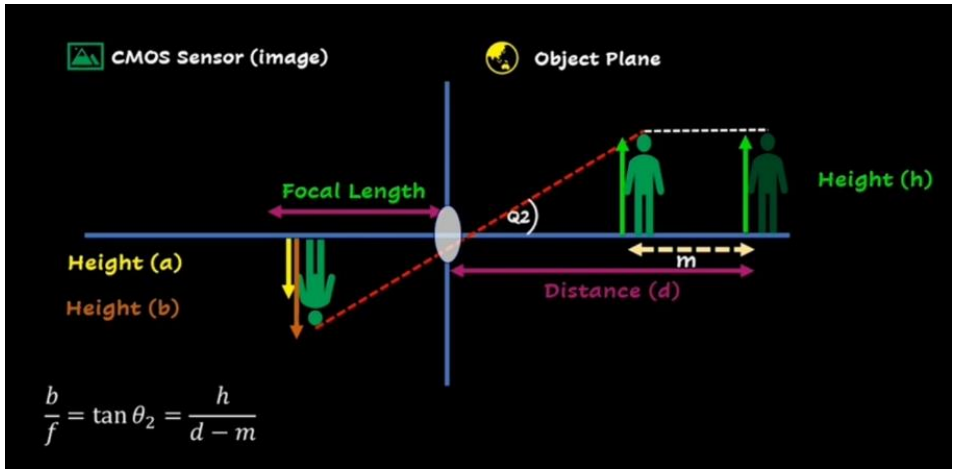
$$\frac{b}{f} = \tan\theta_2 = \frac{h}{d-m} \quad \longrightarrow \text{Equation No: 2}$$

Divide Equation 1 with Equation 2 We get:

$$\frac{a}{b} = \frac{h}{d} \times \frac{d-m}{h}$$

$$d = \frac{m}{1-\frac{a}{b}}$$





METHODOLOGY

Object detection & classification

- Classification identifies objects by classifying them into one of the finite sets of classes, which involves comparing the measured features of a new object with those of a known object or other known criteria and determining whether the new object belongs to a particular category of objects.

- which can be used to form the training set. Once the training set has been obtained, the classification algorithm extracts the knowledge base necessary to make decisions on unknown cases. Based on the knowledge, intelligent decisions are made as outputs and fed back to the knowledge base at the same time

METHODOLOGY

- We will be using a pre-trained SSD detection model trained on COCO DATASETS. It will then test and the output class will get detected with an accuracy metrics.

- After testing with the help of voice modules the class of the object will be converted into a default voice notes which will then be sent to the blind victims for their assistance

- Along with the object detection , we have used an alert system where approximate will get calculated. If that Blind Person is very close to the frame or is far away at a safer place , it will generate voice-based outputs along with distance units.

METHODOLOGY

- **TENSORFLOW APIs**
- We have implemented it by using TensorFlow APIs. The advantage one have by using APIs is it provides us with a set of common operations. Because of which we don't have to write the code for program from scratch.

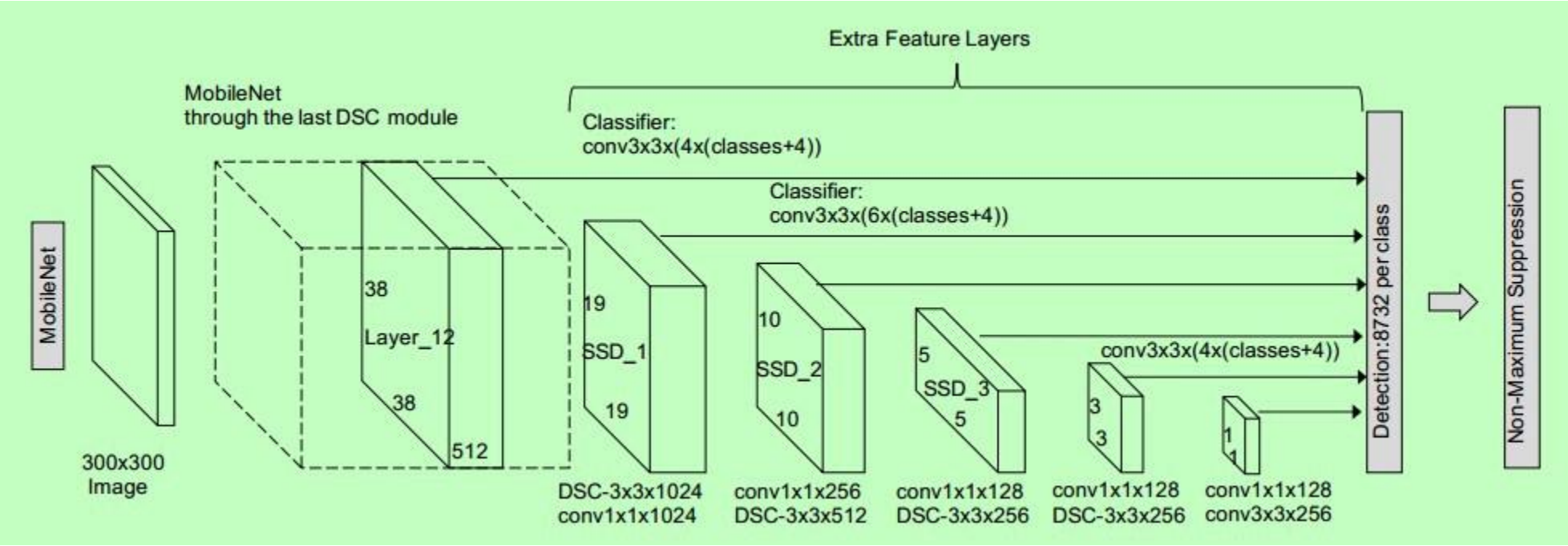
- APIs provides us convenience and hence they are time saver.

- The TensorFlow object detection API is basically a structure build for creating a deep learning network that solves the problems for object detection.

- This includes a collection of COCO dataset, (common object in context)

METHODOLOGY

- For a faster accuracy we will choose SSD DETECTION most of the system shows smooth performance with SSD Mobile_Net DETECTION



Single Shot Detection Architecture

METHODOLOGY

- It’s a architecture and it is trained and used for object detection Single Shot Detector Single Shot detector like YOLO takes only one shot to detect multiple objects present in an image using multibox. It is significantly faster in speed and high-accuracy object detection algorithm.

- YOLO – You Only Look Once

METHODOLOGY

- **MOBILENET**

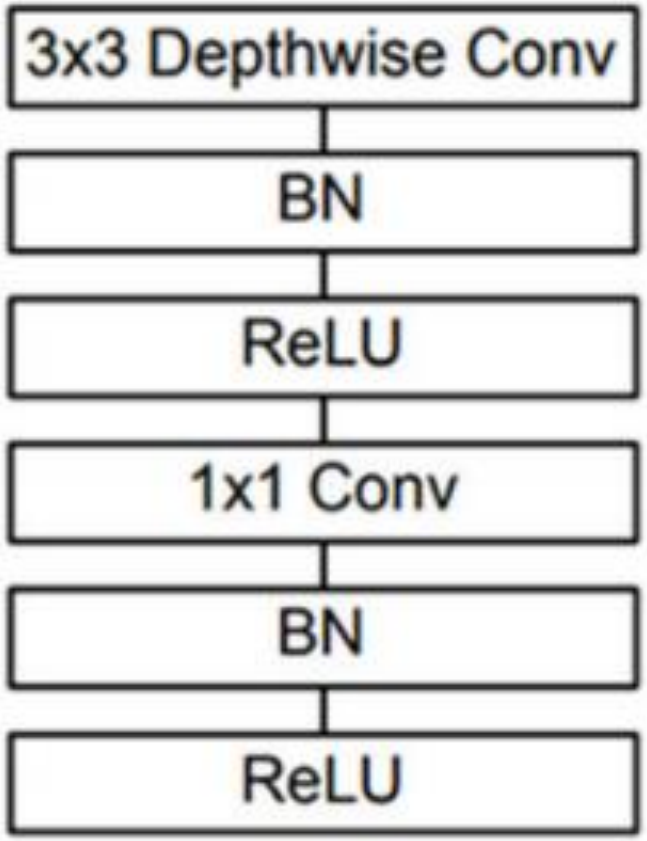
- MobileNet is an efficient and portable CNN architecture that is used in real world applications.

- Depth wise Convolution Filters
- Point wise Convolution Filters

- This model is based on the ideology of the MobileNet model based on depth wise separable convolutions and it forms a factorized Convolutions. This converts a basic standard convolutions into a depth wise convolutions.
- This 1×1 convolutions are also called as pointwise convolutions.

METHODOLOGY

Depth wise Convolution filter based MobileNet Architecture



METHODOLOGY

DEPTH ESTIMATION

- Depth estimation or extraction feature is nothing but the techniques and algorithms which aims to obtain a representation of the spatial structure of a scene.

- It is used to calculate the distance between two objects. Our prototype is used to assist the blind people which aims to issue warning to the blind people about the hurdles coming on their way.

- In order to do this, we need to find that at how much distance the obstacle and person are located in any real time situation.
-
- If that object occupies most of the frame then with respect to some constraints the approximate distance of the object from the particular person is calculated.

METHODOLOGY

VOICE GENERATION MODULE

- After the detection of an object, it is important to acknowledge the person about the presence of that object on his/her way. For the voice generation module PYTTSX3 plays an important role. Pyttsx3 is a conversion library in Python which converts text into speech.

- Audio commands are generated as output. If the object is too close then it states

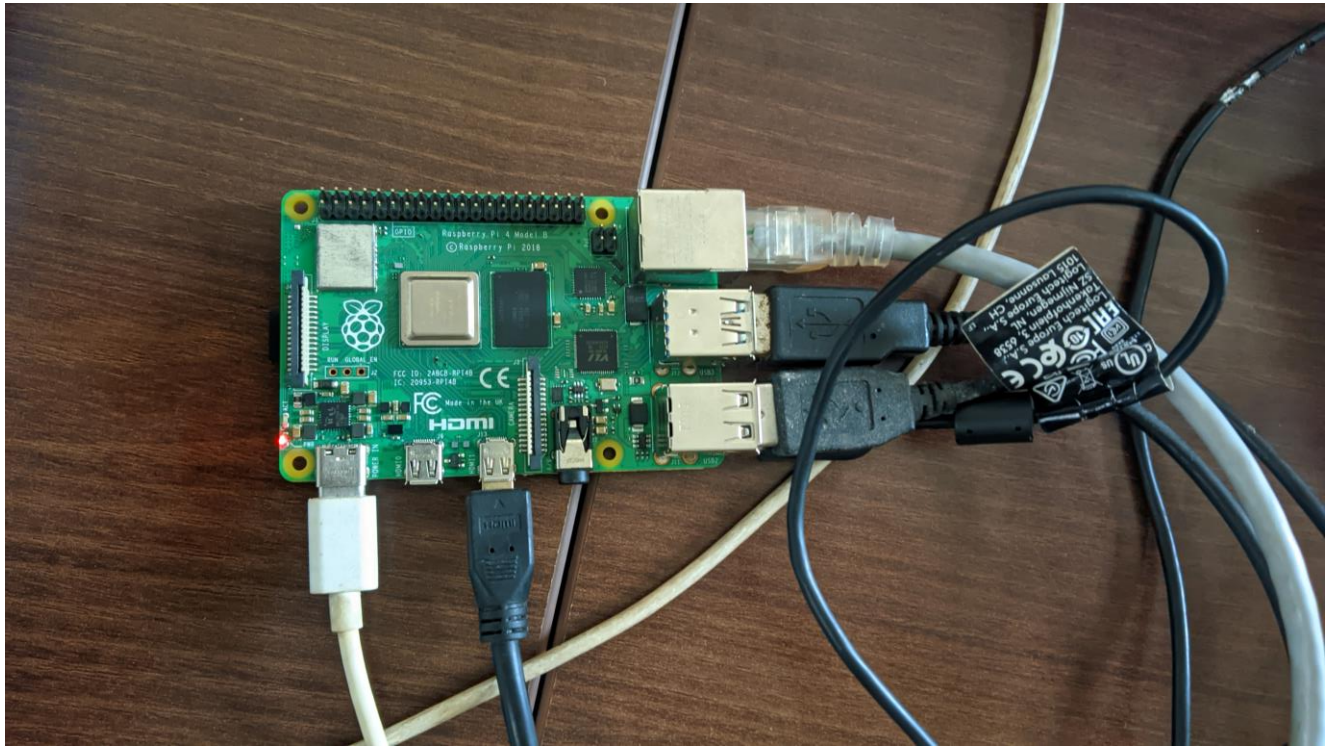
"Warning: The object (class of object) is very close to you. Stay alert!".

Else if the object is at a safer distance then then a voice is generated which says that "The object is at safer distance"

- Pytorch is primarily a machine learning library. Pytorch is mainly applied to the audio domain. Pytorch helps in loading the voice file in standard mp3 format. It also regulates the rate of audio dimension.

contd.

PROGRESS



RESULT

Obstacle detection and Alert Message

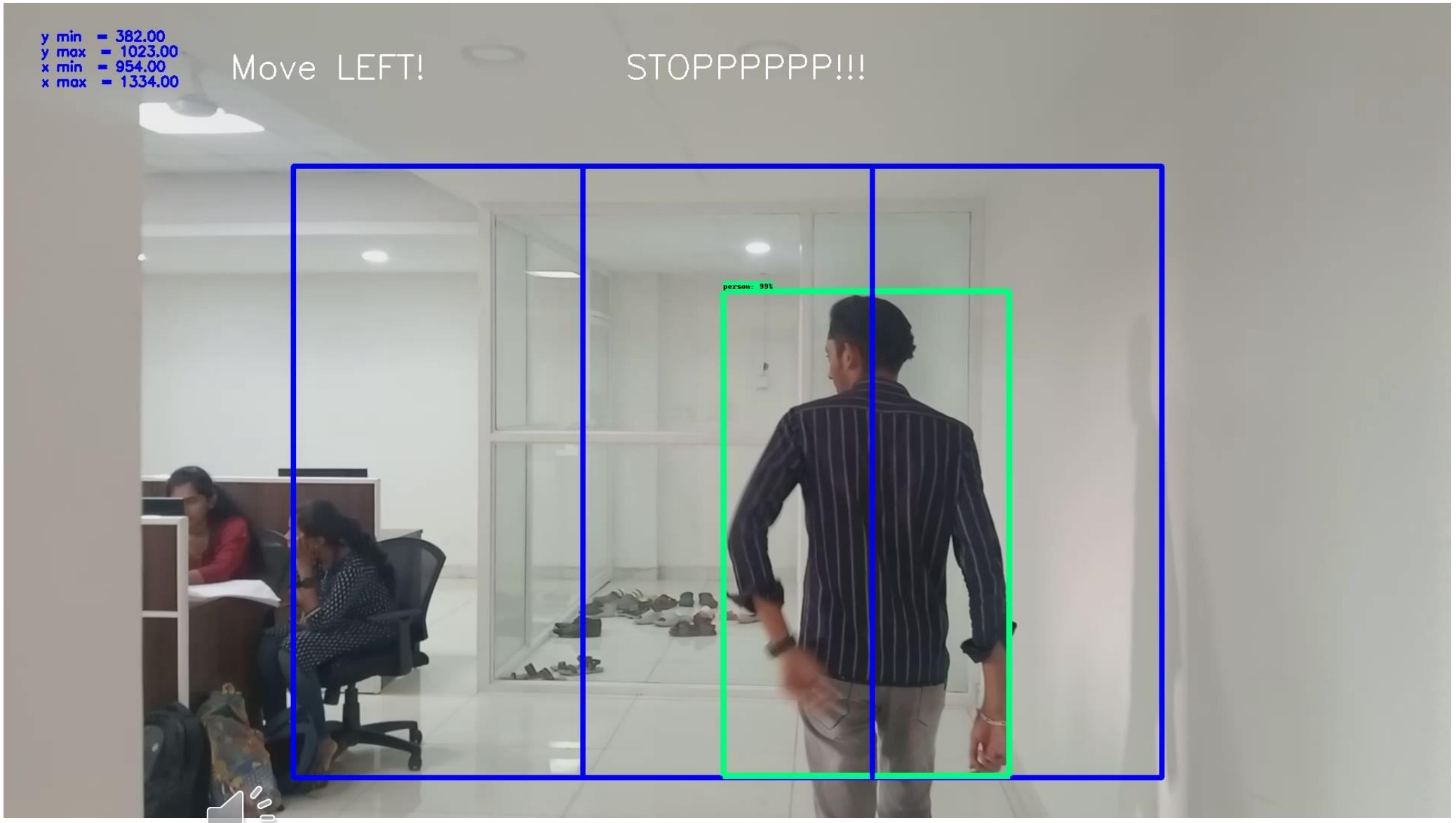
The model can now detect obstacles in front of it and it can specify either to move left right or to stop in a particular direction

```
      cv2.line(frame,tuple(left_boundary),tuple(left_boundary_top), (255,
      cv2.line(frame,tuple(right_boundary),tuple(right_boundary_top), (25
      out.write(frame)
except:
  pass
```

```
1080 1920
0.94111395
left_boundary[0],right_boundary[0] : 768 1152
left_boundary[1],right_boundary[1] : 1026 1026
xmin, xmax : 831 1487
ymin, ymax : 220 1037
1080 1920
0.9638729
left_boundary[0],right_boundary[0] : 768 1152
left_boundary[1],right_boundary[1] : 1026 1026
xmin, xmax : 824 1497
ymin, ymax : 223 1041
1080 1920
0.95539916
left_boundary[0],right_boundary[0] : 768 1152
left_boundary[1],right_boundary[1] : 1026 1026
xmin, xmax : 758 1526
ymin, ymax : 233 1020
1080 1920
0.95556855
left_boundary[0],right_boundary[0] : 768 1152
left_boundary[1],right_boundary[1] : 1026 1026
xmin, xmax : 710 1523
ymin, ymax : 235 1018
```

✓ 1m 57s completed at 12:26 PM

PROGRESS


y min = 382.00
y max = 1023.00
x min = 954.00
x max = 1334.00
Move LEFT!
STOPPPPPPP!!!
person: 99%

PROGRESS

- The operation characteristics of Ultra Sonic Sensor is replaced by Camera module to measure the distance between the user and the object. That defines within the video to move in a right direction

- Used Optimized Algorithms such as SSD Algorithms and Mobile net CNN algorithm

APPLICATION

Assistance through audio to the blind can be used in the following fields

- Blind schools
- Industries
- Home
- Public places

CONCLUSION

This report gives an overall view of the prototypes which have been implemented and yet to implement. Survey of all the assistive devices which helps the visually impaired person has been done. It consists of problems which the visually impaired people are facing in their day-to-day life and solution to these problems has been given. Based on the survey the systems are advanced to help the blind in various fields so that they can be independent and do their work on their own.

THANK YOU