SMART ASSISTANCE FOR VISUALLY IMPAIRED AND BLIND PEOPLE

Neha N[#1], Dr. Sivakumar V [*2], Ria Sathya[#3], Skandana G[#4], Yashaswini R[#5]

[#]UG student, Department of Computer Science, Dayananda Sagar Academy of Technology and Management
[*]Associate Professor, Department of Computer Science, Dayananda Sagar Academy of Technology and Management
Bangalore, Karnataka, India

[1]neha.nagaraj0123@gmail.com [3]riaaaaa31@gmail.com [4]skandana.gowda@gmail.com
[5]yashaswini.ravi20@gmail.com [2]sivakumarvgym@gmail.com

Abstract
The goal of this research is to create a smart app that will assist persons with vision problems. People who are blind or visually handicapped are at a disadvantage in today's society since they are oblivious of their surroundings. The wider populace holds stereotypes about blind persons that they feel are accurate. Many people feel that blind persons are incapable of working or living a normal life. One of the most challenging obstacles is getting from one site to another without the help of others. Other difficulties include identifying people, detecting obstacles, and so on. Some commercially available products can help them overcome some of these challenges. There are constantly a slew of fresh research in the works, all with the goal of inventing solutions to help these visually impaired people. As a result, advanced new technology can be used to increase the amount of support supplied to them. The concept is achieved using a smartphone app that includes functions like as smart speaker, image processing, and more. The device is a voice-activated tool that aids people who are blind in doing daily tasks. The software incorporates a variety of opportunities for the developing into a management framework that is accessible to the visually impaired. The goal is to help blind or visually impaired persons navigate around obstacles.

Keywords:
blind, visually impaired, obstacle detection, voice warning messages, mobile application

1. INTRODUCTION

Blindness is a common and widely accepted disability. The general public holds stereotypes about blind persons that they feel are accurate. Many non-blind individuals feel that blind persons are incapable of working or living a normal life. Blind people live normal lives and follow their unique set of norms. However, they face challenges in terms of inadequate welfare and services problems. People are affected not only by loss of vision, as well as by a decrease in their entire standard of living, which has an impact on their flexibility.

The country's population of legally blind persons has become so high that it has a considerable impact. In up with the fast world, the average citizen does not have time to notice these people with unique needs. Like a result, those who are blind or visually impaired necessitate support in their everyday activities.

One of the most difficult activities is moving from one location to the other without the support of others. Becoming reliant on another person can lead to a lack of drive and, in certain situations, identity. Additional difficulties are identifying individuals and finding impediments to their progress.

Certain commercially accessible items may be able to assist them in overcoming a few of these obstacles. There are constantly a slew of fresh research in the works, all with the goal of inventing solutions to help these visually impaired people.

As a result, an application that can assist visually impaired people in recognizing items while on the go is required. Artificial Intelligence is used in the proposed system, which is completely reliant on voice instructions to assist visually handicapped people. It can also distinguish between still photos and video data.

The following categories are used to organize the remaining sections: Section II goes over the Literature Review, Section III goes over the Current System, Section IV goes over the Proposed System, which includes Problem Description and Methodology, Section V goes over the Tools and Libraries used, Section VI goes over System Architecture, Section VII goes over Conclusion, and Section VIII goes over References.

2. LITERATURE SURVEY

Joe Louis Paul I et al.[1] presented a device consisting of a Raspberry Pi computer coupled with a camera, sensors, headphones, and other accessories. All of the components are connected via GPIO pins (General Purpose Input Output). An LRD circuit is also connected to a Light Emitting Diode on the gadget (LED). The device uses GPS to track the user's current location in real time. The device makes use of two subsystems: route navigation and facial detection and identification. It does not rely on the internet to function. From a given distance away, the device snaps images of everything.

Divya.S, et al.[2] proposed a method for determining the obstruction that relies on ultrasonic and infrared waves generated by the required sensors. During an emergency, the device sends alarm messages and tracks its whereabouts using GSM and a GPS module. A blind individual can send an emergency message if he or she is in danger. A battery with a 12-hour capacity powers the device. The difficulty with the walking stick is that it gets stuck in cracks and uneven surfaces.

Kasthuri R et al.[3] suggested a technology that enables blind and visually impaired people to simply connect to

all web servers via their smartphones. It has a Speech Recognition Engine (SRE) that transforms speech into text for quick operations, as well as an interface that allows visually impaired people to obtain the most up-to-date information from their smartphones. This makes Google Maps and Music systems more accessible to them. When the smartphone's battery falls below 10%, it stops providing optimal results to its consumers.

Arjun Pardasani et al.[4] suggested an object-classifying and-detection glass. It recognises the object's name and provides an audio output to the user. It will also convert text to braille, which will be saved for later processing, such as printing the output in braille for blind people. The second type of device is a pair of shoes that are connected by sensors and are used to notify visually impaired people to objects that are close by. This proposed system is extremely complex, and the ultimate result is extremely slow.

Shubham Melvin Felix et al.[5] describe a technique to assisting persons who are blind or visually challenged that incorporates Artificial Intelligence, Machine Learning, Image and Text Recognition. The concept is realised using an Android smartphone software that incorporates features such as voice recognition, image identification, money recognition, an e-book, a chat bot, and more. Written analysis can be used to detect words inside a text document, and voice commands can be used to detect items in the immediate area. For someone who is unable to communicate, the technology will be tough to utilise. It also helps with full-intensity detection and major colour directions.

S.Durgadevi et al.[6] propose a paradigm in which photos are collected using the camera on their smartphones and then interfaced with a Raspberry Pi device to perform the next stages. The images are then transmitted to the processing board after passing through the Raspberry Pi. They modify and recognise the text from the photos using ROI (region of interest). The Python programming language is utilised for image processing because it allows the device to adapt to changing trends.

3. EXISTING SYSTEMS

Dealing with vision loss or impaired eyesight may be difficult. For visually challenged persons, moving and traveling may be challenging and frustrating. There are several walking poles and canes available on the navigation market. It all began with a simple foot cane, which evolved into a smart walking stick that detects obstacles using infrared (IR) or ultraviolet (UV) sensors. The current procedures are designed to help blind people become more aware of their surroundings and live a more independent life.

The prior method had several faults, notably sporadic Internet access and people's dislike of canes. Numerous people are also afraid to use the stick since it irritates many around them. Almost every available device only addresses one aspect of the blind's concerns. This segment examines the current methods for assisting blind people in detecting barriers and leading separate lives.

3.1 SMART GLASSES

The eyeglasses can acquire resistance and inform the user using earbuds and a noise producer. The circuit board for this project is designed in a specs pattern that the sight impaired may wear. An Arduino linked to the eyewear will determine what is obstructing the visually impaired person's view.

The smart glasses constructed using Arduino are shown in Fig 1.

Fig 1. Smart Arduino Glasses

3.2 WEARABLE BLIND ASSISTANCE GLOVE

A clever technological gadget for blind people that can aid them in moving. By hearing varied durations of buzzer, a blind person may evaluate the distance of adjacent objects with this gadget. Wearable gadgets such as watches and belts are continually being developed.

Fig 2 shows a glove worn by a person that warns the user about obstructions through the use of sensors.

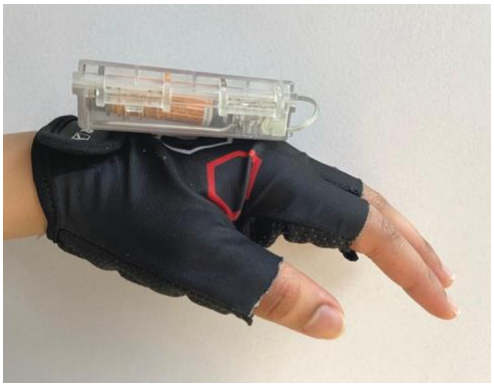

Fig 2. Smart Wearable Glove

3.3 SHOE-BASED BLIND ASSISTANCE SYSTEM

For a visually challenged person, getting from one place to another is the most challenging chore. A suitable equipment is required to assist them and the elderly in walking confidently and independently. The visually impaired might benefit from obstacle

detection and alerts to improve their movement and security, especially in unfamiliar environments. An application named 'Smart Shoes' is introduced, which uses technologies such as IoT and a variety of other algorithms. This is a method of assisting vision-impaired persons by using technology to address their frequent issues when moving about.

A smart shoe, as shown in Fig 3, will detect the existence of an item in front of the wearer.

Fig 3. Smart Shoes

3.4 ULTRASONIC BASED ASSISTIVE HAT OR CAP

People who are blind can walk better securely on the road without worry of colliding with something. The primary goal of this hat is to make life easier for blind people. The Arduino UNO is used to power this hat. It is equipped with three ultrasonic sensors that function as radar. It comprises two components: a broadcaster and a detector. The buzzer beeps and the vibrator starts vibrating to inform the user when an impediment comes within range of the ultrasonic sensor.

Mrs. Subhasini Shukla, et al.[7] proposed solution is an assistive wearable 'Smart Cap' that assists persons with vision impairment engage and navigate their way to safety. For those who live independently, this approach is simple. The purpose of our project is to develop a headpiece for blind people that will guide them from point A to point B.

Fig 4 below displays a vector depiction of a smart cap that uses ultrasonic sensors to detect the presence of an item.

Fig 4. Smart Wearable Cap

3.5 ULTRASONIC BASED SMART CANE

The Ultrasonic Based Smart Cane for Blindness is a revolutionary gadget that detects nearby obstructions using ultrasonic waves and alerts the user through vibrations and buzzer tones.

Agarwal A, et al.[8] introduced a smart walking stick with a mechanism that determines the obstacle using ultrasonic waves emitted by the required sensors.

Fig 5 below displays the sensor-equipped walking cane. This is useful for detecting roadblocks.

Fig 5. Ultra-Sonic Cane

3.6 LASER-BASED SMART WALKER

The image below shows a smart walker for blind people who have trouble walking. The laser are used to look for any obstructions in the surroundings. It uses two lasers, one stationary and the other movable with the help of a servo motor, as well as a belt of five vibrator motors to assist users when an impediment is detected in a laser-based walker aid. Lasers may damage eyes and skin, thus they should be used with caution.

In this field, ongoing research implies that major advancements in human recognition, object tracking, and other areas for guiding blind people are still possible.

Andreas Wachaja, et al.[9] displays a navigation gadget that uses sophisticated vibrotactile signals to safely assist blind people who have problems walking. It has a new vibro-tactile user interface as well as a controller that is based on user research and considers human characteristics. Humans can only perceive and comprehend a limited number of orders and respond slowly, therefore the walker was built to remedy this problem.

4. PROPOSED SYSTEM

4.1 PROBLEM DESCRIPTION

It is common knowledge that there are around 285 million visually impaired persons in the globe, almost equivalent to 20% of the Indian population. Individuals have regular and persistent navigational challenges, especially when they are alone. They rely heavily on others to provide their fundamental requirements on a consistent basis. As a result, it's a difficult work, and a technology solution is critical and urgently required. The issue statement is to devise a system that will enable them to engage with the world without the assistance of others, reducing their reliance on others. This research looked at the problems that blind people encounter on a daily basis and sought to provide a satisfactory solution.

4.2 SCOPE

The proposed Obstacle Detector provides a reliable model which is capable of identifying and recognizing the objects in the captured real-time images.

This programme provided voice help to guide visually impaired people through their daily tasks.

It will be a practical approach for blind users to relate with the surroundings while utilising sophisticated equipment and no longer feeling reliant on others for assistance.

4.3 OBJECTIVE

- The system is structured so that the application may take genuine photos and send them to a computer that acts as a host for all operations.
- The fundamental goal of this project is to create an Integrated Machine Learning System that can detect obstacles.
- The fundamental purpose of object detection is to discover and identify one or more effective targets in still photography or video data.
- The primary purpose is to recognize interior and outdoor items and alert users.
- It enables visually impaired people to recognize and classify common everyday objects in real time, receive verbal feedback, and compute estimated distances, resulting in warnings.

4.4 METHODOLOGY

- The system is set up in such a way that the app may take genuine photos and transfer them to a computer that serves as a server for all of the computations.
- The laptop that serves as a server will employ a detection model that was trained utilising coco datasets. The model will test the output class and reliably detect it.
- After testing with speech modules, the object's class will be transformed into a collection of standard audio content, which will be provided to the blind victims for aid.
- In addition to object detection, we deployed an alarm system that calculates an estimate. If the Blind Person is extremely close to the frame or in a safer location, it will provide vocal style responses and also proximity units.

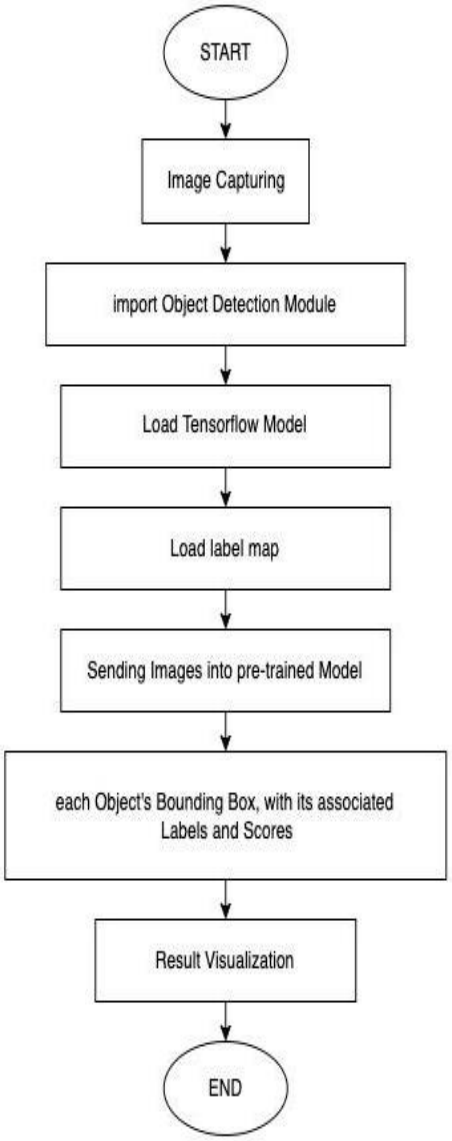

Fig 6. Workflow of Object Detection Algorithm

4.5 DESIGN AND IMPLEMENTATION

A. Overview of the Application

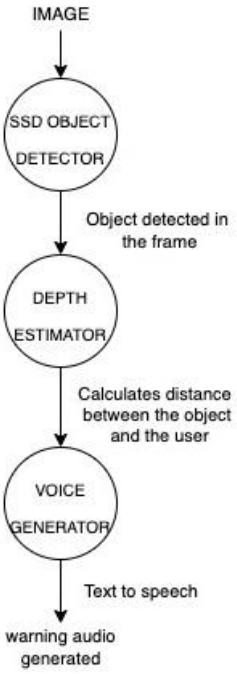

Fig 7. Diagram

Our project consists of 3 major modules:

1) Object detection module using SSD object detector
2) Depth estimator
3) Voice generation

B. DATA FLOW DIAGRAM

i) Object Detection

The smart phone captures the images which acts as the dataset for the already trained object detection module.
Object detection mainly consists of the following:

- Object detection which detects the object within the image.
- Object tracking which detects the position of the object within the image.

After detection, it is compared with identified matched record for identifying the kind of object detected in the image. Once the object is recognized it labels the object identified enclosed within a bounding box. The output of object recognition is further sent to the other modules.

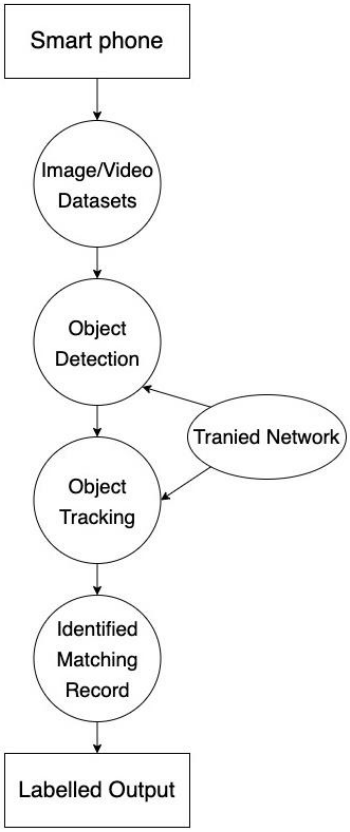

Fig 8. Object Detection Diagram

ii) Depth Estimation

Depth is crucial when it comes to understanding the world around us such as navigation and mapping. To interact with the physical world around, depth is the deciding factor. The RGB image undergoes texture extraction feature and segmentation. It is very necessary to extract meaningful features from raw images that describes texture properties.

It can be defined as function of spatial variation of brightness intensity of pixels. A method of separating a digital image into several image segments is known as segmentation.

Depth map creates a distance representation of an image from a reference point, it provides details of the depth based on how far or near the object is in terms of perspective.

Texture segments and texture features are added together to give segments with texture which provides integration of 2D location and depth information. Using this information, the relative distance between the item and the human is computed and determined, and an output is produced.

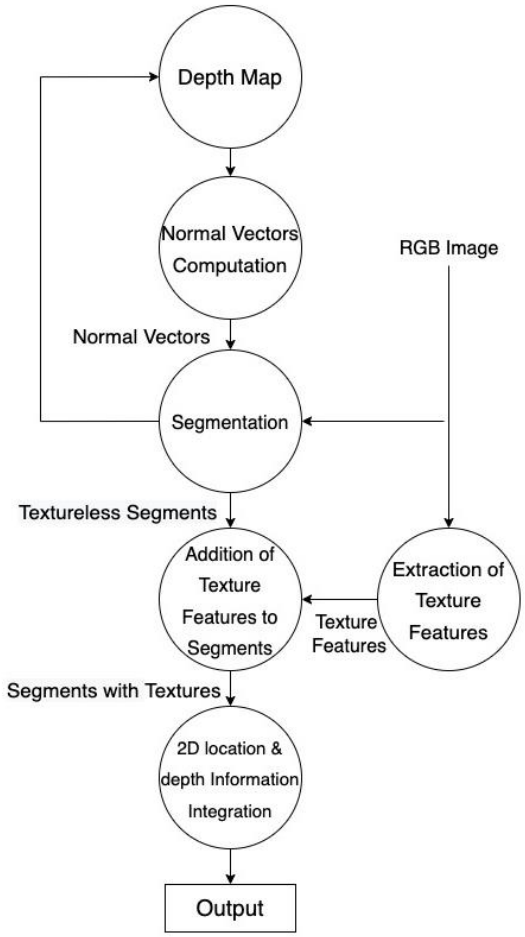

Fig 9. Depth Estimation Diagram

iii) Text to Speech

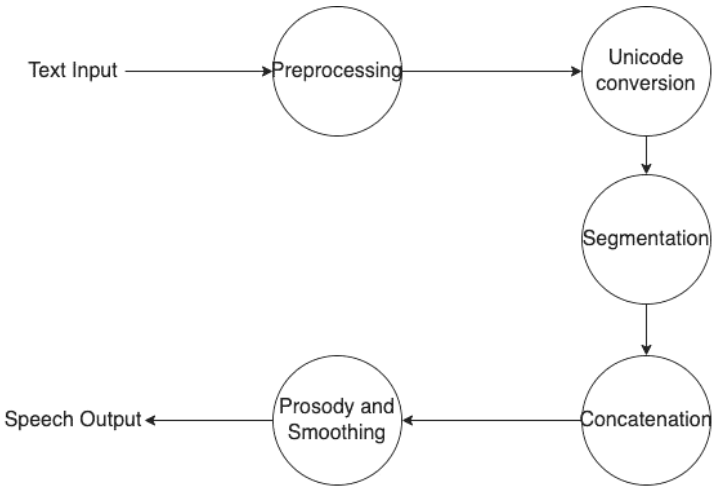

Fig 10. Text to Speech Diagram

PYTTSX3 is a text to speech conversion library in python. It is performed by beginning with a pre-processing procedure for a text corpus. Expanded text, abbreviators, acronyms, and numerals After that, the pre-processed text will be transformed to Unicode.

Unicode was created with the express purpose of overcoming the constraints of conventional character compression algorithms. Now a text is segmented into syllables and duplicates are removed.

The syllables are then concatenated and it further undergoes prosody and smoothening process. This finally gives a smooth human speech output. Variations can be applied to resulting output.

5. TOOLS AND LIBRARIES

TensorFlow is a free and open-source machine learning platform. It includes a number of technologies, packages, and civic engagement that enable academics to learn about machine learning while also enabling designers to easily construct and launch ML applications.

5.1 SSD ARCHITECTURE

Tensorflow comes with a number of models that have already been trained. SSD detection is utilized to deliver quicker and more precise results, depending on the system specs.

A Single Shot Detector (SSD) is made up of two parts: one with an SSD head and another with a backbone. As a feature extractor, the backbone model is basically a learnt image categorization network. In the spatial position of the last layer activations in the SSD head, one or more convolutional layers are added to the backbone and read as bounding boxes and object classifications. Each grid cell in a picture is responsible for recognizing items in its own area.
Object detection includes anticipating an object's class and placement confidential a given area.

5.2 ANCHOR BOX

Many anchor boxes can be linked to a single pixel in SSD. These specified anchor boxes are individually accountable for a certain size and form within a grid cell. SSD performs the matching step during training to guarantee that each real data object within a picture's anchor frame and surrounding enclosures is a good match. The anchor box that bears the strongest similarity to a certain item may be used to anticipate the object's class and position. After the system has been trained, this characteristic is utilized to both train the system and forecast the locations of identified objects. Each anchor box has an aspect ratio and a zoom level assigned to it.

5.3 ZOOM LEVEL

The shapes of all items are not the similar. Some are shorter, longer, or both, while others are narrower, wider, and shorter, longer, or both. This is made possible by the SSD architecture, which allows the anchor boxes to have pre-determined aspect ratios. At each zoom level, the ratios parameter of the anchor boxes coupled with each grid can specify several aspect ratios. The anchor boxes may or may not have to be the same size as the grid cell. The user may be looking for both smaller and larger items within a grid cell. The zooms option sets the amount by which each pixel's anchor boxes should be changed up or down.

5.4 DEPTH ESTIMATION

The methods and procedures used to obtain a depiction of a scene's spatial structure are referred to as depth estimation, also known as extraction feature. To put it another way, it's utilized to figure out how far two objects are apart. In order to aid blind persons with obstacles on their path, it is important to determine the distance between the obstruction and the person in any given situation. When an object is detected, a rectangular box is created around it. If the object takes up the majority of the frame, the approximate distance between the object and the person is determined, taking into account some constraints. Following the detection of an object, the algorithm is utilized to calculate the object's relative distance from the human.

5.5 VOICE GENERATION MODULE

After detecting an object is done, it is necessary to notify the individual about the existence of an object. pyttsx3 is an essential element of the voice generation module. The Python conversion module pyttsx3 transforms text to speech. Python 2 and 3 are both supported by this library. pyttsx3 is a straightforward text-to-speech converter. To find hidden text in an image, we use Python-tesseract for character recognition. Optical Character Recognition (OCR) identifies textual data on images and converts it to a computer-readable format. The content is detected by scanning and analyzing the image. As a result, Python-tesseract recognizes and "reads" text encoded in images. These texts are referred to a pyttsx.

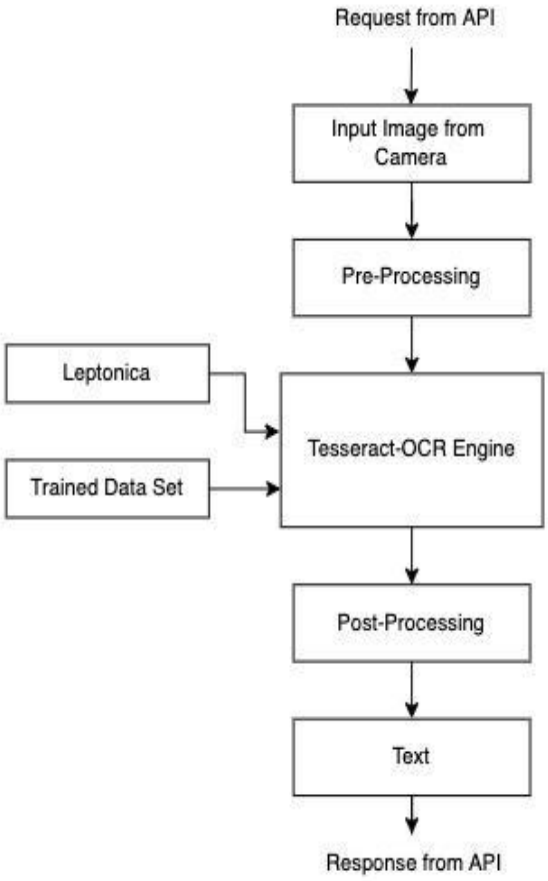

Fig 11. Flowchart for OCR Process

6. SYSTEM ARCHITECTURE

Real-time photos obtained using a smartphone are used in the software. These photos are sent to a server, which subsequently saves them. After that, the photos are examined and divided down into frames.

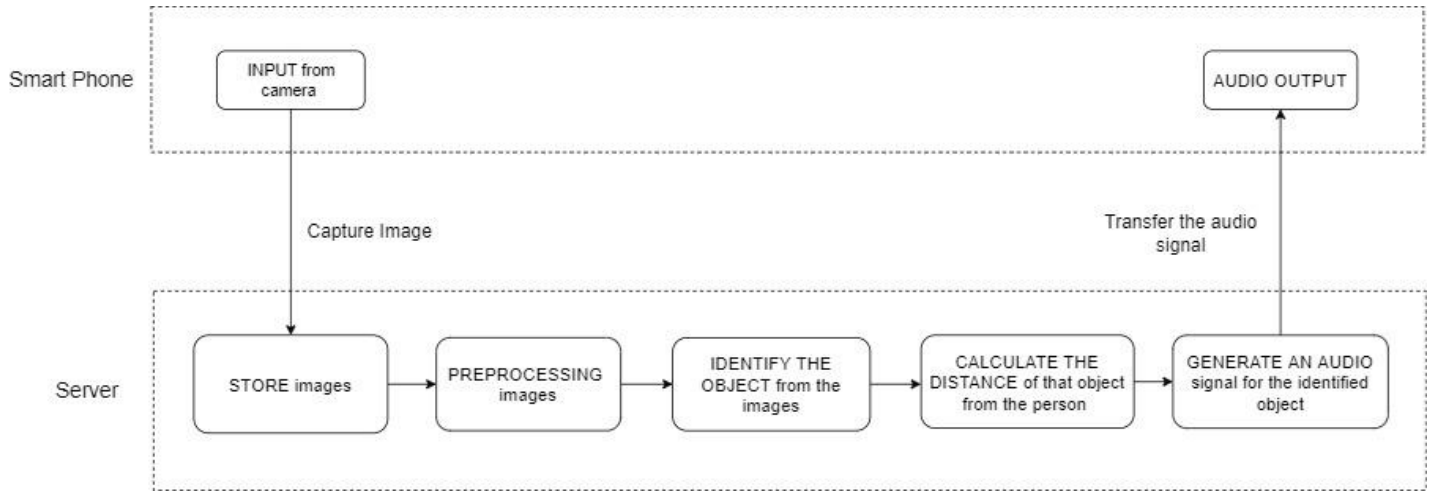

Fig 12. System Architecture Flowchart

These frames are supplied to the pre-trained model, which then identifies the item inside them. The depth estimation technique calculates the distance between the camera that took the picture and the item in the frame. After the item and distance are established, an audio message is created and sent to the user. On the user's smartphone, this audio is played.

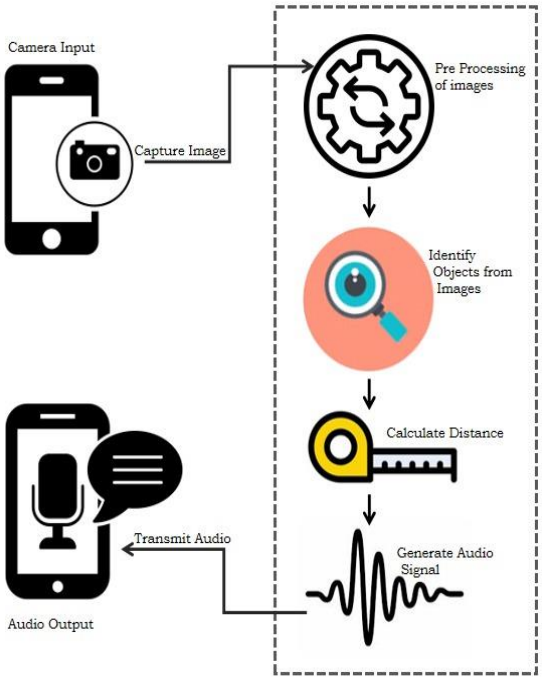

Fig 13. System Architecture Diagram

7. RESULT

7.1 EXPECTED OUTCOME

- The system is set up in such a way that the application may take real photos and transmit them to a pc that acts as a host for all functions.
- The suggested paradigm is anticipated to function similarly to the libraries and framework already in place.
- Develop a deeper grasp of the user's surroundings and characterize the items that are present.
- Among the algorithms, methods, and approaches employed, we will be able to recognise and identify things across different places utilising various optimization strategies in order to reduce waste and maximise model efficiency.
- In comparison to bulkier IoT-based gadgets, assess the efficacy of object detection.
- The frame is used to determine the final distance. If the item is too close to the user, a distance warning is provided, along with a voice output. If the item is at a safe distance, a class recognition voice is created and the name of the object recognized is heard instead of a distance-based warning.
- With the aid of the matplotlib packages, the findings may be further studied.
- Object detection, threat identification, voice messaging, detecting things and individuals, and generating estimated distance are all characteristics of the proposed model.

7.2 OUTPUT

Object detection module identifies items using a pre-defined model, as shown in the figures below.

Fig 14. The system is detecting a scissor with 100% accuracy.

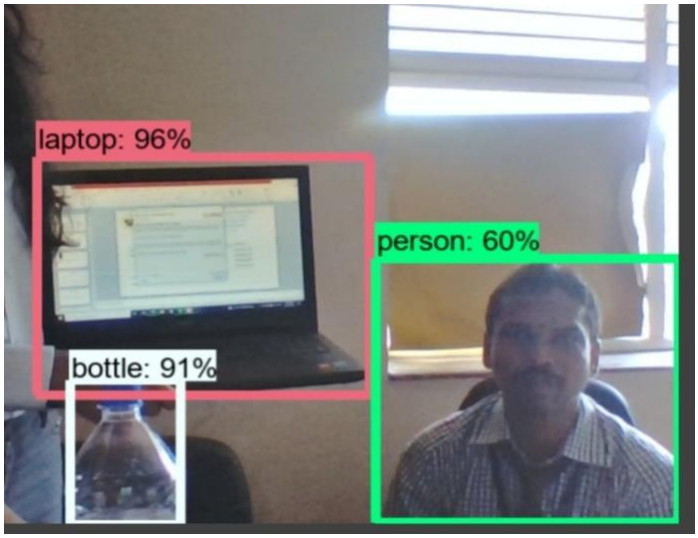

Fig 15. The system is detecting a person, a bottle and a laptop with an accuracy of 60%, 91% and 96% respectively.

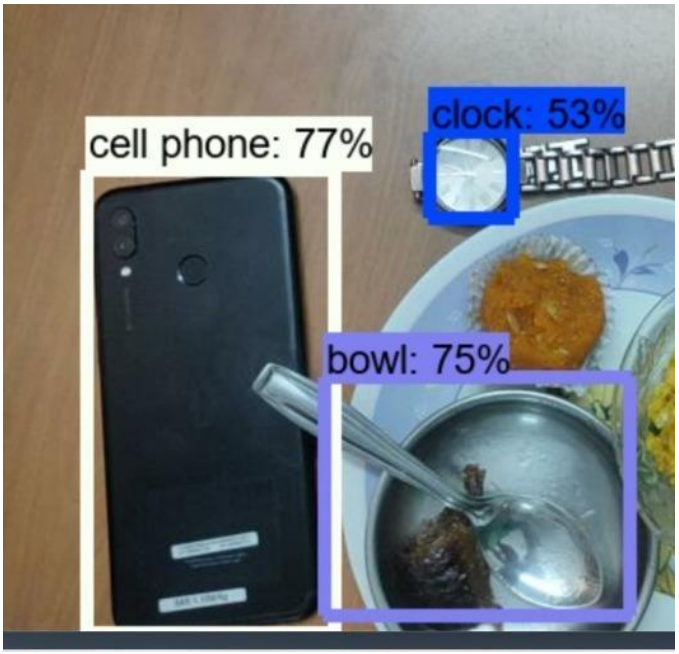

Fig 16. The system is detecting a cell phone, a clock and a bowl with an accuracy of 77%, 53% and 75% respectively.

Fig 17. The system is detecting a cell phone, a spoon and a bowl with an accuracy of 94%, 75% and 54% respectively.

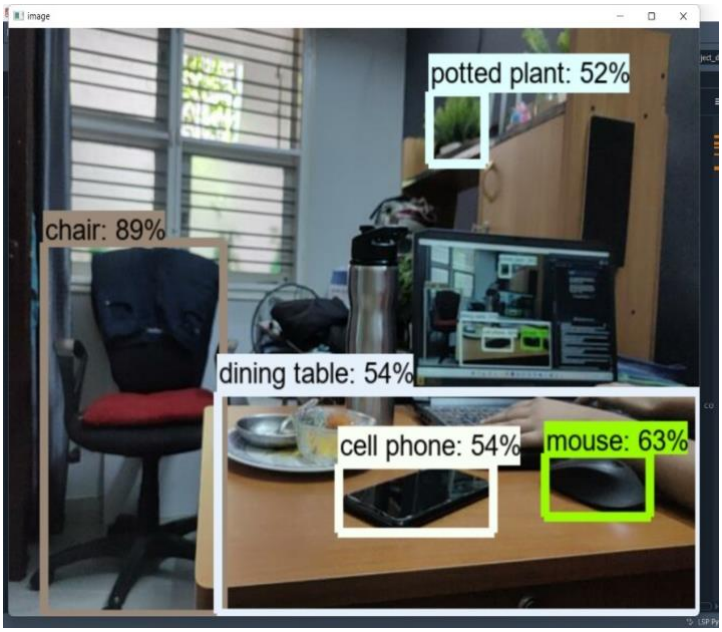

Fig 18. The system is detecting multiple objects with an accuracy above 50%.

8. CONCLUSION

The proposed system blocs the capabilities of the numerous workings to create a versatile smartphone for those who are blind or visually impaired. The application is simple to use because it makes use of cellphones, which are convenient and portable. Object identification and voice guiding are among the functions included in the programme, which is designed for visually impaired people. It contains a user-friendly interface developed exclusively for visually impaired people, as well as detection findings that are read out loud so the user can plainly hear them. As a result, their lives will be made easier. Our proposed solution might be implemented in a multilingual application, allowing users to use it in their local language with ease. Image recognition can be improved with more information about the image obtained. It also works with the Internet of Things.

REFERENCES

[1] Joe Louis Paul I , Sasirekha S, Mohanavalli S, Jayashree C, Moohana Priya P, and Monika K, “Smart Eye for Visually Impaired-An aid to help the blind people”. *Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019).*

[2] Divya .S, Shubham Raj, M. Praveen Shai, A. Jawahar Akash,Ms. V. Nisha “Smart Assistance Navigational System for Visually Impaired Individuals” *2019 IEEE International Conference on Intelligent Techniques in*

Control, Optimization and Signal Processing (INCOS).

[3] Kasthuri R,Nivetha B, Shabana S,Veluchamy M,Sivakumar S "Smart Device for Visually Impaired People". *2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)*,pp- 54-59.

[4] Arjun Pardasani, Prithviraj N Indi, Sashwata Banerjee, Aditya Kamal, Vaibhav "Smart Assistive Navigation Devices for Visually Impaired People ". 2*019 IEEE 4th International Conference on Computer and Communication Systems.*

[5] Shubham Melvin Felix, Sumer Kumar, and A. Veeramuthu, "A Smart Personal AI Assistant for Visually Impaired People". Proceedings of the 2*nd International Conference on Trends in Electronics and Informatics (ICOEI 2018).*

[6] Durgadevi, K.Thirupurasundari, C.Komathi and S.Mithun Balaji,"SMART MACHINE LEARNING SYSTEM FOR BLIND ASSISTANCE", *2020 2nd International Conference on Power, Energy, Control and Transmission Systems.*

[7] Mrs. Subhasini Shukla, Shreya Pimple, Shraddha Shetke, Namrata Gaikwad, Krutika Patil, "Smart Cap for Visually Impaired Person using Raspberry Pi", *International Research Journal of Engineering and Technology (IRJET)*, pp- 2805-2809

[8] Agarwal A, D.Kumar, and A. Bhardwaj, "Ultrasonic Stick for Blind". *International Journal of Engineering and Computer Science, 2015*

[9] Andreas Wachaja, Pratik Agarwal Mathias Zink, Miguel Reyes Adame Knut Moller, Wolfram Burgard,"Navigating Blind People with a Smart Walker", *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Congress Center Hamburg, pp- 6014-6019