e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

OBJECT, COLOUR AND DISTANCE DETECTION SYSTEM FOR VISUALLY IMPAIRED PEOPLE

Nishad Kadam*1, Varun Singh*2, Shashank Singh*3,

Prof. Ashwini Phalke*4

*1,2,3,4Department Of Information Technology, Vasantdada Patil Pratishthan's College Of Engineering And Visual Arts College In Mumbai, Maharashtra, India.

ABSTRACT

Computer Vision aims at developing systems to interpret or understand visual data. Image processing which is the core element of computer vision focuses on processing and enhancing the raw visual data and making it suitable for various operations. Recently, there is a lot of development in the field of computer vision and image processing. Due to the introduction of deep learning, many computer vision problems can be solved easily. Object detection which has a variety of use cases like face detection, object tracking, and video surveillance, has made rapid progress due to deep learning. Using these technologies everyday life of visually impaired people can be enriched. We have proposed an object, colour and distance detection system that will help visually impaired people to detect or identify the object which is placed in front of them and will also help them to identify the attributes of the detected object like colour and distance. We hope this system will improve the quality of life for blind people by allowing them to identify objects and the attributes of that object.

Keywords: Smart Glasses, Visually Impaired People, Object Detection, Color Detection, Distance Detection, Text-To-Speech.

I. INTRODUCTION

Walking in and around the house and even on the street is quite easy and effortless for normal people but, can you imagine how the life of someone who is blind could be? According to a survey conducted globally, among 7.79 billion people living in 2020, 49.1 million are blind. Moreover, India has one-third of the world's blind population Many of them can't walk without the assistance of the white cane that they have to carry to navigate around and sometimes need the constant help of someone to take them to the place that they wish to go. Navigating around their own house is difficult sometimes as they don't know which object is in front of them and at what distance. Locating crucial things is also a challenging task for them. The increasing number of individuals with disabilities in the world attracts the majority of researchers to use the latest technologies to build systems, hoping that these systems can assist the disabled people in closing their tasks in way of life like normal people.

So, we wanted to build something for visually impaired people that helps them to identify and locate objects present in their surroundings and also provide some additional data to the blind about the object.A smart glasses project is what we propose to make. These smart glasses take inputs through a camera module and provide sound feedback to the person through headphones. For detecting the object which is present in the surrounding of the blind, we have used SSD (Single Shot Detector) object detection model which is used to detect objects in real-time quickly. This model has been trained on the COCO dataset and has the ability to detect 80 different objects that we use in our day to day life. For colour detection, we have used OpenCV which is a library that aims to solve real-time computer vision problems. To calculate the distance of the object from the camera module we have used the focal length of the camera and triangle similarity. And then finally we convert this information about objects from text to voice to inform the blind person about their surroundings.

II. LITERATURE SURVEY

For the proposed problem, products are designed for solving the issues by detecting the item which is placed in the front of the blind with the assistance of deep learning algorithms like R-CNN (Region-based Convolutional Neural Network), Faster R-CNN, YOLO (You Only Look Once) and more. Only some of those products along with detecting the object also detect the attributes of an object like the colour and distance of the object from the blind person. Moreover, the object, colour and distance detection models employed by these products demand

a good amount of computing power which is a constraint on devices like Raspberry Pi. This increases the response time of the system which is unacceptable in the real-time environment.

Jyun-You Lin, Chi-Lin Chiang and Meng-Jin Wu [1] proposed a solution to the problem of the blind. They came up with Smart Glasses Application for visually impaired people based on Deep Learning. In their project, the blind person wears the smart glasses and takes pictures of the scenes with the help of a camera module, the pictures captured are stored on an SD card and then are uploaded to the backend server designed by them. Then with the help of the object detection model, the object is detected and then the output which is in text format is given to the TTS (Text-To-Speech function) which outputs the detection result in a voice manner so that the blind person can know the object which is in front of him. This project uses YOLO v3 deep learning object detection system based on Tensorflow and Keras version. This project provides a good foundation for solving the problems of blind . This system achieves results within 4 seconds which is appreciable, but the drawback of the system is that it just detects the object which is placed in front of the blind and not its attributes like colour and distance of the object from the blind. The computing or processing devices like RAM, CPU, GPU used by the system has a very high computing power (RAM-16GB, CPU-Intel Core i7-8700K 3.40GHz) which enables the system the generate output quickly but this makes the system difficult to deploy in a wearable device for blind.

Sonay DUMAN, Abdullah ELEWİ, Zeki YETGİN [2] proposed a system to solve the problems of blind people. In their proposed system they used a CNN (Convolution Neural Network) based object detection model called YOLO to detect objects with the help of a single camera mounted on a Raspberry Pi board. They have used a non-linear regression model, random forest for distance estimation. The system provides great accuracy (98.86%) in distance estimation. An essential attribute of the object like the colour of the object is not provided by the system. In a project, undertaken by Dr Sunitha M.R, Fathima Khan, Gowtham Ghatge R and Hemaya S [3] gives the blind ability to estimate whether the obstacle placed in front of them is a human or not. Ultrasonic waves are used to estimate the distance of the human from the blind person and it also triggers the identification system. This system also uses Raspberry Pi. Only whether the obstacle is a human and its distance from the blind person is known with help of this system. In the paper published by B.N. Krishna Sai and Sasikala T [4], the Faster R-CNN algorithm is used for detecting threatening objects using the camera. This algorithm even though provides accurate results, doesn't detect objects quickly and isn't ideal for real-time use cases.

III. PROPOSED SYSTEM

The proposed system is composed of different modules, each of which has a specific function. The block diagram of the system is shown below. Python is used as the programming language for this project because of its simplicity and ease of use. Firstly, with the help of a USB camera that is attached to the Raspberry Pi 4, the video of the surroundings of the blind person is captured. As video is nothing but a collection of frames, with the help of OpenCV and Numpy each frame of the video is converted from BGR (Blue, Green and Red) format to HSV (HUE, Saturated and Value) format. Then with the help of Quantized SSD (Single Shot Detector) v1 TensorFlow lite object detection model, the object is detected. The class name of the object (for example, if the object detection model is used to detect different fruits then Apple, Banana, and Mango can be class names) is stored inside of a variable. With the help of the pixel width of the object, which is calculated from the bounding box coordinates provided by the object detection model, the calculated focal length of the camera and Triangle Similarity, a distance of the object from the camera module of the system is calculated and is stored inside of another variable. Colour range boundaries have been set in the system which helps to identify 5 different colours which are Black, Yellow, Red, Green, Blue. Hue, Saturation and Value values of three different pixels which are located on the detected object, are compared with the colour range boundaries set in the system and the colour of the object is identified. The identified colour string is stored in a varible. These three variables are then passed to Text-To-Speech function which convert the text to a voice format which can be heard by the blind person with the help of earphones attached to Raspberry pi 4 thus providing the user with object class , distance and colour of the detected object.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

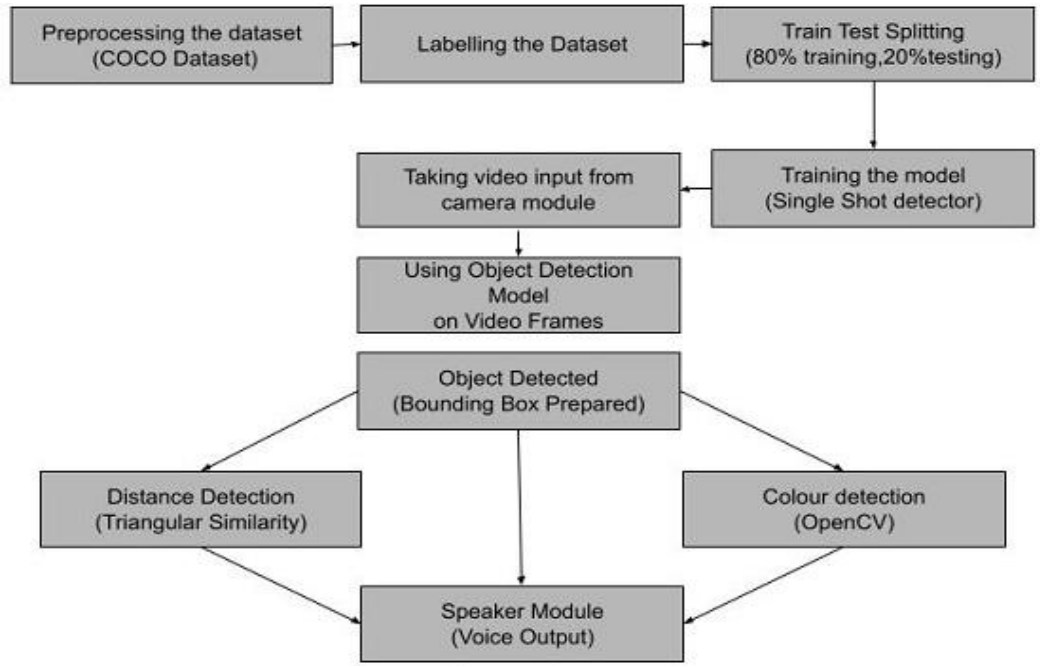

Figure 1: Block Diagram of the System

1. Object Detection

In this project for detecting objects, we've used Tensorflow. TensorFlow is a library developed by Google for deep learning and machine learning. It provides an ecosystem of tools and libraries that help programmers to build ML applications quickly. Tensorflow object detection API aims at solving object detection problems and is built on top of TensorFlow. It contains a variety of pre-trained models that can be used to detect objects. Here the models are pre-trained on huge datasets like the COCO dataset and KITTI dataset. In this project out of the various object detection models available we have used the SSD (Single-shot detector) MobileNet V1 object detection model. As the name suggests SSD models take just one shot to detect different objects present in the images or video frames. SSD model is faster than other models. Models like R-CNN and Faster R-CNN find a region of interest or region proposals using an algorithm and then create boundary boxes based on those regions. Then image classification is carried out on those boundary boxes to identify or classify objects. While this approach provides high accuracy, it is slow and time-consuming, the process runs at 7 frames per second, making these models unsuitable for real-time applications. SSD is designed for object detection in real-time, it speeds up the process by eliminating region proposal algorithms. As this object detection model is supposed to run on resource constraint devices like Raspberry Pi we have used the SSD object detection TensorFlow Lite model. TensorFlowLite models are optimised to run on resource constraint devices.

2. Object Distance Detection

For detecting the distance from the camera module to the object Triangle Similarity concept is used. With the help of triangle similarity, we can find distance using a single camera. The triangle similarity concept is something like this, let's say we have an object of known width W, then we place this object at some known distance D from the camera module and measure the width of the object is pixels P. Using these three parameters and the formula:

$$F \text{ (Focal Length)} = (P \times D) / W$$

we can calculate the focal length of the camera. Then when we move the camera module closer or far away from the object and we can use triangle similarity to know the distance of object from the camera module using formula:

$$D^* \text{ (Updated Distance)} = (W \times F) / P$$

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

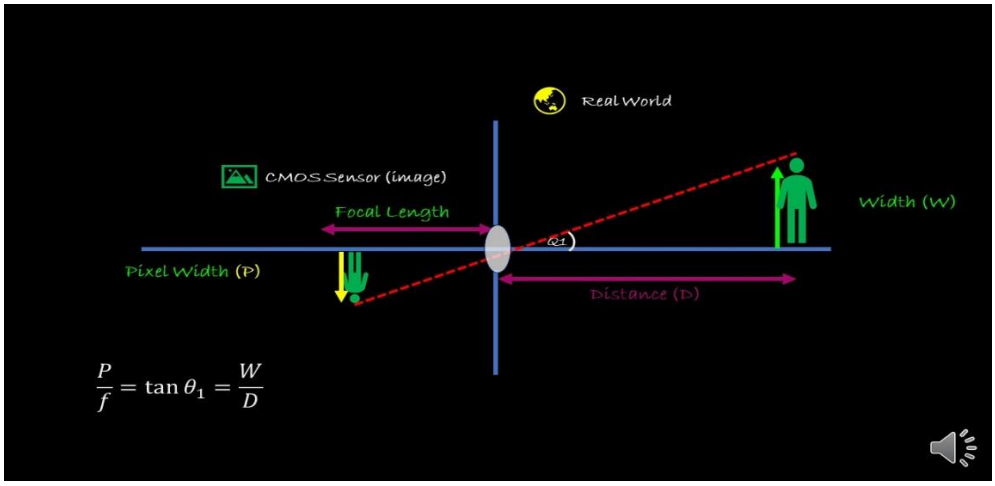

Figure 2: Calculating Focal length of Camera

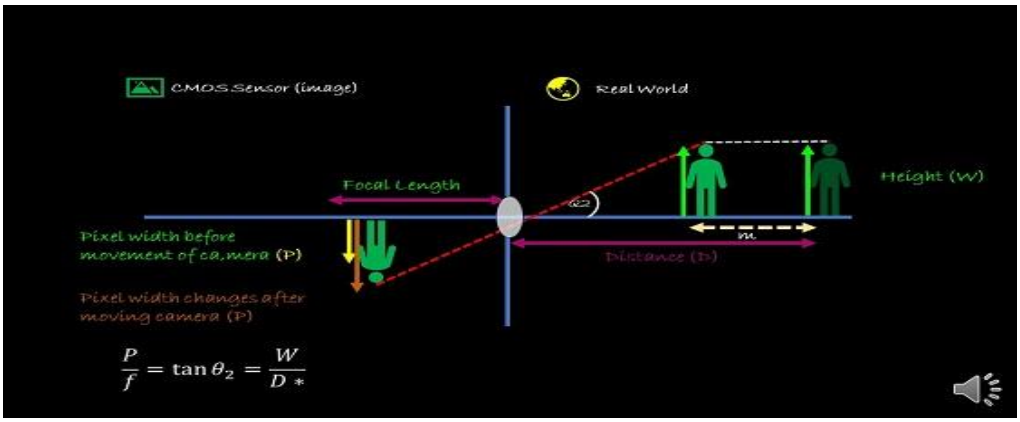

Figure 3: Finding Distance after changing camera position using Focal length

For this particular project what we have done is, we have recorded the actual width W of the crucial objects that are to be detected for the blind person. Then we keep the object in front of the camera at a known distance D and let our SSD TensorFlow Lite object detection model detect the object. For each object, it provides bounding box coordinates of which left and right coordinates (both are x coordinates) are used to find the width of the object in pixels P. Then using all this data and the above-mentioned formula focal length is calculated for all the objects. There is a minor difference between the different focal lengths calculated for every object. The mean of all these focal lengths is calculated and the resulting focal length is used to find the distance of the object from the camera.

3. Object Colour Detection

Video is a collection or series of frames, each frame is equivalent to an image. Pixels are the primary element in an image. We have used the OpenCV python library to capture video streams from the surrounding of a blind person. Our system is capable of identifying basic 5 different colours including Black, Green, Red, Blue, Yellow OpenCV loads the images or captures the video data in BGR (Blue Green Red) format. We convert the frames from BGR format to HSV (Hue, Saturated, Value) format. In this format, each pixel is represented using 3 numeric values namely Hue, Saturation and Value. We have defined colour range boundaries for the above-mentioned colours in our system. For instance, if the Hue value of a pixel on the detected object falls between 1 to 8 then the pixel is of Red colour, if the hue value is between 9 to 33 the pixel is of yellow colour, if the hue value is between 34 to 69 the pixel is green in colour and if the hue value is between the range 70 to 117 the pixel is blue in colour. We locate three pixels on the body of the detected object and then compare their Hue, Saturation, Value values of these pixels with the ranges mentioned above and detect the 5 basic colours.

4. Raspberry Pi 4

Raspberry Pi 4 Model B is used in this project to run different project modules. Raspberry Pi 4 Model B is the latest model in computers provided by Raspberry Pi. It has a better processing speed, storage, memory and

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

connectivity as compared to previous Raspberry Pi models. Raspberry Pi single-board computers provide great features that attract developers to use Pi in their products. Here are some specifications of the Raspberry Pi 4 Model B.

1. Quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz

2. 2GB LPDDR4 Ram

3. Micro SD card slot for operating system and data storage 4. 2.

4 GHz and 5.0 GHz IEEE 802.11b/g/n/ac wireless LAN, Bluetooth 5.0, BLE Gigabit Ethernet ,2 × USB 3.0 ports, 2 × USB 2.0 ports.

5. Standard 40-pin GPIO 6. 5V DC power supply via USB-C connector (minimum 3A), 5V DC power supply via GPIO header (minimum 3A)

IV. RESULT

A prototype of the system that was proposed has been implemented. The system was tested using 5 different objects, a chair, tv, cup, bottle, and keyboard, as mentioned in the table below. The name of the object and the colour of the object were known before the testing. To test the system the object was placed at a known distance from the camera module of the system. Then the system was activated and the object, colour and distance detections were carried out and the results were recorded in the table below (Table 1). Then the object was brought near to the camera module of the system to check whether distance detection works properly, the current distance between the camera and the object was recorded and then the system was again activated to detect the changed distance. Now after deactivating the system the object was placed at a distance greater than the two previously recorded distances. Again the distance between the system and the object was recorded and the system was activated and the detections were made. The results of detection are mentioned in the table below. The system detected the object and its class and the colour accurately. The distance that was detected by the system had a minor error. For instance, when the cup object was placed at a distance of 36 cm from the camera, the detected distance by the system was 34cm. Thus the system detected the object and the colour of the object and predicted the distance of the object from the camera with great accuracy.

Table 1: The detection result of the system

Actual Object	Detected Object	Actual Colour	Detected colour	Actual Distance in cm	Detected distance in cm
CUP	CUP	YELLOW	YELLOW	36	34
CUP	CUP	YELLOW	YELLOW	26	28
CUP	CUP	YELLOW	YELLOW	40	39
KEYBOARD	KEYBOARD	BLACK	BLACK	37	34
KEYBOARD	KEYBOARD	BLACK	BLACK	47	46
KEYBOARD	KEYBOARD	BLACK	BLACK	27	26
CHAIR	CHAIR	RED	RED	136	138
CHAIR	CHAIR	RED	RED	91	93
CHAIR	CHAIR	RED	RED	110	111
BOTTLE	BOTTLE	GREEN	GREEN	22	22
BOTTLE	BOTTLE	GREEN	GREEN	14	15
BOTTLE	BOTTLE	GREEN	GREEN	27	26
TV	TV	BLACK	BLACK	164	166
TV	TV	BLACK	BLACK	102	94
TV	TV	BLACK	BLACK	138	141

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

V. CONCLUSION

The proposed system has been developed and its main aim is to give the blind person a sense of the real world or his surroundings. The system uses a camera module to capture the surroundings of the blind and provides information to the blind person about the object, its colour and the distance of the object from the person in sound format with the help of headphones. This project uses Tensorflow object detection API which is an open-source framework based on TensorFlow, particularly SSD (Single Shot Detector) object detection model from Model Zoo which provides a variety of trained object detection models. TensorFlow Lite SSD model is used as it is optimized to work on resource constraint devices like Raspberry Pi. This project also uses the OpenCV library to get the visual data from the surroundings through the camera module, convert the video frames from BGR (Blue, Green, Red) format to HSV (Hue, Saturation, Value) and along with NumPy which is a library for multi-dimensional array processing, identify the colour of the object. With the help of the focal length of the camera, the actual width of the object and triangle similarity the distance of the object from the person wearing the glasses is calculated. This whole system is mounted on a Raspberry Pi 4 which has high processing power, sufficient enough to run object detection models. The camera module and earphones are connected to Raspberry pi. We hope that the system can help blind people to interact with their surroundings and help them to better understand environment.

VI. REFERENCES

[1] Jyun-You Lin, Chi-Lin Chiang, Meng-Jin Wu, Chih-Chiung Yao, Ming-Chiao Chen, Department of Computer Science and Information Engineering, “Smart Glasses Application System for Visually Impaired People Based on Deep Learning” (2020).

[2] Sonay DUMAN, Abdullah ELEWİ, Zeki YETGİN Computer Engineering Dept. Mersin University, Turkey “Design and Implementation of an Embedded Real Time System for Guiding Visually Impaired Individuals” (2019).

[3] Dr. Sunitha M.R. Professor, Computer Science and Engineering Adichunchanagiri Institute of Technology, Fathima Khan, Gowtham Ghatge R, Hemaya S. Bachelor of Engineering “Object Detection and Human Identification using Raspberry Pi” (2019).

[4] B N Krishna Sai, Sasikala T. Department of Computer Science and Engineering, Amrita School of Engineering, Bangalore, " Object Detection and Count of Objects in Image using Tensor Flow Object Detection API" (2019).

[5] Dola Das, Prottoy Saha and Nawshin Tabassum Department of Computer Science and Engineering, Bangladesh “Real Time Distance and Mobility Detection of Obstacles Using a Smart Multi-sensor Framework for Visually Impaired People” (2020).

[6] Qianjun Shuai, Xingwen Wu School of Information and Communication Engineering Communication University of China, CUC Beijing, "Object detection system based on SSD algorithm" (2020).