INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Virtual Assistant for the Blind

Isshita Borkar[1], Prof Asma Shaikh[2], Snehal Jadhav[3], Vaishnavi Khandade[4], Bhakti Nagpure[5]

[1, 2, 3, 4, 5]Computer Engineering Department, Marathwada Mitra mandal's college of engineering , Maharashtra,India

Abstract: In this world, certain people are visually impaired and they face a lot of problems while doing there day-to-day work. Physical movement is one of the biggest challenges for the visually impaired. People with complete blindness or low vision often have a difficult time in self-navigating unfamiliar environments. Hence this work is aiming for developing a device which will help them as their personal assistant. The goal of the present project is to model an object detector to detect objects for visually impaired people and other commercial purposes by recognizing the objects at a particular distance. The object recognition deep learning model utilizes the You Only Look Once(YOLO) algorithm and a voice announcement is synthesized using text-to-speech (TTS) to make it easier for the blind to get information about objects. This system proposes conglomeration of technologies like Image processing, Speech processing etc, so the problems faced by blind people can be reduced to certain extent. Object recognition methods in computer vision, Image processing, Text to Speech conversion can be embedded in a single object : SMART GLASSES (spectacles).
Keywords: YOLOV3, Convolutional neural network(CNN),Object detection, Distance estimation

I. INTRODUCTION

In today's advanced hi-tech environment, the need for self-sufficiency is recognized in the situation of visually impaired people who are socially restricted. They are in an unfamiliar environment and are unable to help themselves. Because most tasks require visual information, visually impaired people are at a disadvantage. This is because the crucial information about their surroundings is unavailable. Visually challenged people perform day to day activities. These activities contain various obstacles due to unknown surroundings. It is high time now that we contribute to our society and assist specially abled people in order to make their work obstacle free. It is now possible to extend the support provided to people with visual impairments thanks to recent advancements in inclusive technology. This project proposes to use Artificial Intelligence, Machine Learning, Image and Text Recognition to assist persons who are blind or visually impaired.

Visually impaired people need assistance in various day to day activities. Through the virtual assistant we can help them in detecting and identifying the object in front of them. This virtual assistant can help the blind as well as partially blind people. Virtual assistant is the need of the hour. Our project might help the visually impaired people in understanding the surrounding, that is understanding about the objects in front of them.Our virtual assistant includes features such as voice assistant, image recognitionthrough raspberry Pi camera, and achieve the distance hopefully using ultrasonic sensors. The virtual assistant used the YOLOv3 algorithm to detect and identify the objects captured in real time.

II. LITERATURE SURVEY

Rutuja kukade et al. have created a Virtual personal assistant system which can multitask such as read emails maintain a diary get weather forecast and it uses voice hat to process the commands and talk back to the user. They have used raspberry pi with voice hat and Google TTS and STT module.

Ankush Yadav et al. have done a comparative study of object recognition algorithms and their advantages as well as accuracies were compared. They have proposed a new system "Android-Based Object Recognition for the Visually Impaired" using Faster RCNN which had an accuracy of approximately 95%.

Freddy Poly et al. have developed a device equipped with an ESP32 camera and multiple ultrasonic sensors. A real time system has been created which performs object detection on the live streamed data, generated by the camera. In accordance with the data from the sensors, the object and its proximity are identified. It contains of two parts 1) VI Spectacle 2) Mobile Application. VI spectacle is with the one having the camera module and sensors and all the computations takes place in the mobile application. The mobile application performs a variety of tasks like object detection, route planning and guiding the user.

Vipul Sharma et al. have developed system comprises several Deep learning models; some of them are object detection, face recognition, speech recognition. The website is built on the backbone of flask, which serves the purpose of providing connectivity between the python code and the HTML.

The main aim of this system is to build an automatic text reading assistant which combines small-size, mobility and low cost price. The disadvantage identified was slower MVP development in most cases.

Weal A. Ezat et al. proposed an experimental work for train the deep learning model based on YOLOv3 architecture. The training process had been done using the data-set PASCAL VOC 2007 and data-set PASCAL VOC 2012 and using The Adaptive Moment Estimation Optimizer. The YOLOv3 is efficient with average precision rate of 80.07% The only disadvantage is that the method can detect only selected objects that are defined in the dataset

V. Balaji et al. proposed a model to detect object for visually impaired people and commercial purposes by recognizing the object at a particular distance. Already trained model caffemodel is used. The Accuracy gained was up-to 95%. One of the major drawback is that the system might find little difficult to detect object when there are some background disturbances or if the images are blurred

Dr. B. Harichandana et al. have proposed a system which transforms the various forms of text into digital form. It can be known as OCR. Various API will appear for different platforms which are utilized for OCR implementation. The accuracy they have reached is about 90%. It is simple and fast.

Author	Year	Techniques used	Advantages/ accuracy
Rutuja V. Kukade et al.	2018	Raspberry pi with Voice hat, GoogleTTS API, STT Module	This is a system which can multitask such as read emails, maintain a diary, get weather forecast and it uses voice hat to process the commands and talk back to the user.
Ankush Yadav et al.	2020	Google TTS API	This system allows user to search anything on the internet and thus give audio output as well as it prints the text output.
Freddy Poly et al.	2020	YoloV3,COCO dataset, ultrasonic sensors	The mobile application performs a variety of tasks like object detection, route planning and guiding the user. The system is able to detect and recognize objects in real time. Its pretty balanced and user friendly system.
Vipul Sharma et al.	2020	The system developed comprises several Deep learning models; some of them are object detection, face recognition, speech recognition. The website is built on the backbone of flask, which serves the purpose of providing connectivity between the python code and the HTML.	The main aim of this system is to build an automatic text reading assistant which combines small-size, mobility and low cost price.
Weal A. Ezat et al.	2021	This paper describes an experimental work for train the deep learning model based on YOLOv3 architecture. The training process had been done using the data-set PASCAL VOC 2007 and data-set PASCAL VOC 2012 and using The Adaptive Moment Estimation Optimizer.	The YOLOv3 is efficient with average precision rate of 80.07%
V.Balaji et al	2020	This present project is a model an object to detect object for visually impaired people and commercial purposes by recognizing the object at a particular distance. Already trained model caffemodel is used.	Accuracy upto 95%.
Dr. B. Harichandana et al.	2022	This technology transforms the various forms of text into digital form. It can be known as OCR. Various API will appear for different platforms which are utilized for OCR implementation.	Accuracy upto 90%. It is simple and fast.

III. PROBLEM STATEMENT

To develop a system to detect objects, identify and gain an audio output containing the object description anddistance between the blind person and the object.

A. Project Objective

The objective of this project is to help the blind people in various day to day activities such as – identifyingthe physical objects in front of them, giving the description about the various objects by scanning the QR code, information about various hospitals, and allowing them to give input to all functionalities implement.
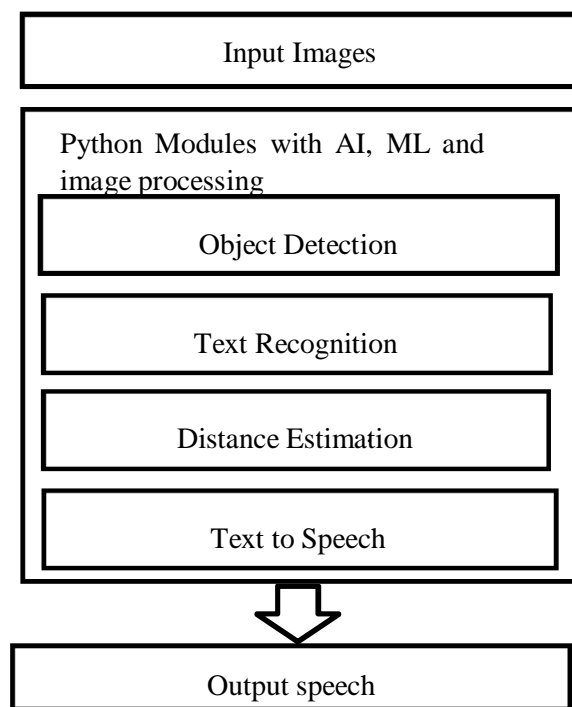
B. Proposed System

Fig 1 System architecture of Virtual assistant for blind

C. Object Detection Module

We have used this algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images. YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects. This means that prediction in the entire image is done in a single algorithm run. Entire architecture into two major components: Feature Extractor and Feature Detector (Multi-scale Detector).The image is first given to the Feature extractor which extracts feature embeddings and then is passed on to thefeature detector part of the network that spits out the processed image with bounding boxes around the detectedclasses
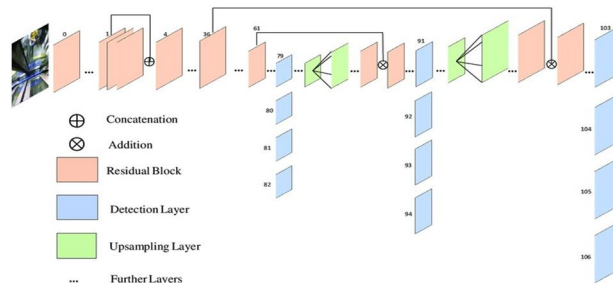

Fig 2. Architecture of YOLOV3

D. Text to Speech Algorithm

1) Prepare a dataset of known/recognized objects.
2) Reading this dataset in our program.
3) Loading any image format (.bmp, .jpg, .png) from given source.
4) Detecting the objects present in the images using Edge detection, line detection, and pattern detectionalgorithms.
5) Extracted objects are than compared with the know dataset.
6) Decision is based on this comparison if the object is recognized or not.
7) Once the object is recognized its respective description is extracted.
8) Text to Speech conversion using known description of object and output is provided through earpiece.

IV.   IMPLEMENTATION

Virtual assistant consists of raspberry pi camera and a text to speech api. The camera captures the images in real time. COCO dataset is ued to train the images. These images are processed and detected using YOLOv3. The object s are identified and a bounding box is formed around the object. The google text to speech API helps in reading the image containing text. The GTTS module gives voice based assistance to the visually impaired. An image input is given to the system and an audio output is obtained. The virtual asistant processes the image,identify the object and calculate an approximate distance of the object from the user. We estimate the distance between object and user using ultrasonic sensors. Finally the audio output is given to the user from the system.

V.   CONCLUSION AND FUTURE SCOPE

A modular solution is presented in this paper for improving the day to day activities by identifying the variousobstacles in front of the blind person and guiding them accordingly. Virtual assistant successfully helps in detecting the objects using yolov3.The system contains four modules namely- object recognition, text recognition, Distance estimation and text to speech that are currently implemented. Our Virtual assistant in the future will help the visually impaired or someone with low vision issues identify the objects in front of them aswell as give them an estimation of distance between object and the person as an audio output using text to speechmodule.

REFERENCES

[1]   Rutuja V. Kukade, Ruchita G. Fengse, Kiran D. Rodge, Siddhi P. Ransing, Vina M. LomteVirtual"Personal Assistant for the Blind", IJCST Vol. 9, Issue 4, October - December, Pune, Maharashtra
[2]   Ankush Yadav, Aman Singh, Aniket Sharma, Ankur Sindhu, Umang Rastogi,Desktop Voice Assistant for Visually Impaired,International Journal of Recent Technology and Engineering (IJRTE), Volume-9 Issue-2, July 2020
[3]   Mrs. J. Meenakshi, Dr. G. Thailambal,Object Recognition by Visually Impaired using machine learning:A study, International Journal of Mechanical Engineering,Vol. 6 No. 3 December, 2021
[4]   Freddy Poly1, Dipak Tiwari2, Varghese Jacob,VI Spectacle – A Visual Aid for the Visually Impaired,International Journal of Engineering Research & Technology (IJERT),Vol. 8 Issue 05, May-2019
[5]   Isha S. Dubey, Isha S. Dubey, Ms. Arundhati Mehendale, "An Assistive System for Visually Impaired using Raspberry Pi", International Journal of Engineering Research & Technology (IJERT),Vol. 8 Issue 05,May-2019.
[6]   Sagar Agrawal, Mandar Agrawal, Prof. Sagar Padiya, "Android Application with Platform Based On Voice Recognition For Competitive Exam", International Journal of Advanced Research in Science & Technology (IJARST),Volume 5, Issue 5, May 2020
[7]   Amanda Lannan, "A Virtual Assistant on Campus for Blind and Low Vision Students" ,THE JOURNAL OF SPECIAL EDUCATION APPRENTICESHIP,Vol.8(2),September 2019.
[8]   Ronald Maryan Rodriques, "Speak2Code: A Multi-Utility Program based on Speech Recognition that Allows you to Code Through Speech Commands ",IJARSCT, Volume 3, Issue 1, March 2021.
[9]   Avanish Vijaybahadur Yadav, Sanket Saheb Verma, Deepak Dinesh Singh,"VIRTUAL ASSISTANTFOR BLIND PEOPLE", International Journal Of Advance Scientific Research And Engineering Trends, ||Volume 6 ,Issue 5 , May 2021.
[10] Sharma, Vipul and Singh, Vishal Mahendra and Thanneeru, Sharan, "Virtual Assistant for Visually Impaired",Department of Information technology, Pillai College of Engineering, University of Mumbai, April19, 2020.
[11] Yasir Dawood, Ku Ruhana Ku-Mahamud, Eji Kamioka, "Distance Measurement for Self-Driving Cars Using Stereo Camera"Proceedings of the 6 th International Conference on Computing and Informatics, ICOCI2017 25-27April, 2017
[12] Weal A. Ezat, Mohamed M. Dessouky, Nabil A. Ismail, "Evaluation of Deep Learning YOLOv3Algorithm for Object Detection and Classification " Menoufia J. of Electronic Engineering Research (MJEER), Vol. 30, No. 1, Jan.2021
[13] Giancarlo Iannizzotto, Lucia Lo Bello, Andrea Nucita, Giorgio Mario Grasso,   "A vision and speech enabled, customizable, virtual assistant for smart environments"2018 11th International Conference on Human System Interaction (HSI), 04-06 July 2018
[14] Jinqiang Bai, Shiguo Lian, Zhaoxiang Liu, Kai Wang, Dijun Liu, "Virtual-Blind-Road Following BasedWearable Navigation Device for Blind People", IEEE Transactions on Consumer Electronics, Volume: 64, Issue: 1, February 2018
[15] Pooja R. More, Puja S. Raut, Priydarshini M. Waghmode, "VIRTUAL EYE FOR VISUALLY BLINDPEOPLE" International Journal of Advance Scientific Research & Engineering Trends, Volume 5, Issue 12, June 2021

[16] V.Balaji, S.Kanaga Suba Raja, C.J.Raman, S.Priyadarshini, S.Priyanka, "Real Time Object Detection For Visually Imparied," European Journal of Molecular & Clinical Medicine, 2020, Volume 7, Issue 4

[17] Tejal Adep, Rutuja Nikam, Sayali Wanewe, Dr. Ketaki B. Naik, "Visual Assistance for Blind People using Raspberry Pi", international Journal of Scientific Research in Computer Science, Engineering and Information Technology Volume 7, Issue 3, May-June-2021

[18] Prof. Priya U. Thackeray, Kote Shubham, Pawale Ajinkya, Shelke Om, " Smart Assistance System for the Visually Impaired", International Journal of Scientific and Research Publications, Volume 7, Issue 12, December 2017

[19] Miss Rajeshwari Ravindra Karmarkar, "Object Detection System for the Blind with Voice Guidance", International Journal of Engineering Applied Sciences and Technology, Vol 6, Issue 2, 2021

[20] Dr. B. Harichandana, Dr. C. Krishna Priya, Dr. P. Sumalatha, "Speech_Based Virtual Assistant System For VIsually Impaired People", INTERNATIONAL JOURNAL OF MECHANICAL ENGINEERING, Vol.7No.5, May, 2022

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⓢ (24*7 Support on Whatsapp)