IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CANEGUIDEX: SMART OBSTACLE RECOGNITION AND VOICE ASSISTANT FOR THE BLIND

[1]Sunil Kumar K N, [2]Mahalakshmi K, [3]Gayathri T, [4]Amulya A

[1] Assistant Professor, [2] Student, [3] Student, [4] Student

Department of Information Science and Engineering,
Cambridge Institute of Technology, Bangalore, India

Abstract: "CaneGuideX" incorporates sophisticated equipment to assist the blind. It enhances safety and liberty for individuals exploring new situations by using smart obstacle detection algorithms to detect and evaluate their surrounds in real-time. It offers clear direction and thorough descriptions of barriers through voice aid, making navigation simple and effective. By providing thoughtful, proactive assistance, this ground-breaking technology transforms the way blind people use canes, allows them to move with assurance and independence. With its ability to extend the gap between those with blind people and the surroundings around them and promote greater diversity and autonomy in daily activities, CaneGuideX is a significant leap in accessibility technology.

***Index Terms* -** Obstacle Detection, YOLO, Deep Learning, Raspberry pi, Text-to-Speech(tts).

I. INTRODUCTION

Vision is God's beautiful gift to all living things. The human visual system plays an significant role in analyzing information about the environment. Because visual signals provide additional information than auditory information, visual cues are better than auditory signals when people recognize information. Blind people have difficulty identifying even the smallest details with healthy eyes. The blind people problem is the difficulty of adapting to an unknown external environment. Use of traditional headgear, guide dogs, and mobility training, including intelligence and support from professionals working in the field of orientation and mobility to assist the blind. A visually impaired assistive system is a piece of technology that makes navigating the environment is easier for those who can vision impairments. By giving blind individuals precise and up-to-date information about their environment, this system seeks to improve their safety and sense of independence. Many approaches are there to implement a visually impaired assistive system. Certain techniques involve the advantage of audio output and sensors like infrared or ultrasonic. An ultrasonic system may, for instance, employ high frequency sound waves to identify obstructions and communicate to the blind person via speakers. Similar to this, an infrared system can make use of sensors to find objects and provide information about their place and distance. A visually impaired assistive system's most important element is the output of audio. This is a outcome of the system's need to give the user clear information in an understandable manner. The usability of a visually impaired assistive system is another important factor. This should be easier for people to know how to use the system; it should be simple and intuitive. This is especially necessary for elderly users who may not has much expertise with technology. Making sure a this assistive system is accurate and dependable is one of the implementation problems. For instance, the system ought to be able to identify barriers even in difficult settings like congested city streets. The machine must provide accurate information about the location of objects such as curbs and stairs. Machine learning techniques are incorporated into many visually impaired assistive system to address these issues. Over time, these algorithms might adjust to the user's tastes and habits based on how they engage with the system. For instance, over time, the system might learn to offer more thorough details about items the user regularly interacts with, like their favorite coffee shop. The possibility for improving blind people's overall well-being of life is another benefit of visually impaired assistive system. These techniques can helped people to feel less stressed and more confident by giving them clear and reliable details about their surroundings.

Additionally, blind people may boon from using this assistive system Systems to help them access previously unattainable possibilities, like work or education. Many methods for the blind have advanced significantly with the introduction of visually impaired assistive system. These technologies can improve blind people's safety and independence by giving them accurate and trustworthy information about their environment. Furthermore, by giving visually impaired people new options, lowering stress levels, and boosting confidence, visually impaired assistive system have the efficiency to improve their overall well-being of life. Among the most fundamental activities in today's society is reading. It can be hard to find a good speed while visually impaired, for example, when reviewing the material. Encouraging those with severe visual impairments to read printed labels and item bundles will enhance their well-being of life and promote social and financial independence. Recent developments in PC vision, digital cameras, and small PCs are precisely helping people by producing camera-based products that combine PC vision innovation with other already-existing commercial products, including OCR frameworks. Pupils who struggle with vision have difficulty examining and comprehending the book's content, which requires close examination. When put up properly, it achieves amazing academic execution after being configured in associations. It has been noticed that the visually handicapped are having difficulty in the areas of preparation and business go ahead and drive. From the Internet to implementing breakthroughs in daily life, data and organized technology have advanced quickly. Object acknowledgment innovation, or object detection, is one of the technologies to take into account. This phrase refers to the ability to recognize the size and form of various objects while recording their position using the device's camera. Object detection in this research refers to the process of identifying instances of actual objects, such as cars, motorcycles, chairs, doors, trees, and people, in still images or films. It improves our overall comprehension of the picture by enabling us to identify, and detect a broad range of objects inside an image. Examples of applications include image retrieval, security, surveillance, and advanced driver assistance systems. Finding objects and surrounding them with distance-distant rectangular bounding box-like structures is the basic tenet of object detection. Applications for detection of object are emerging in many different fields, such as real-time sports applications, agricultural inspection, counting, and person recognition. The suggested smart cane with micro camera arrangement is made to enable legible text by turning text into audio that will be played through speakers. A Raspberry Pi module powers this portable, reasonably priced smart cane, and image processing software aids in identifying and extracting text from images. Ultimately, the text extract is transformed into speech, making it audible to individuals with visual impairments.

II. RESEARCH METHODOLOGY

The creation of visually impaired assistive system is making greater use of artificial intelligence, deep learning, and machine learning. These innovations enable these systems to perform better overall and increase accuracy and reliability.
Artificial intelligence is the ability of a robot to perform tasks that usually involve human brain function, such as distinguishing between consecutive numbers, predicting the future, and choosing the best one based on all available information.
Machine learning is a part of artificial intelligence, which is defined more broadly as the development of algorithms that allow computers to improve their performance through observation. For example, deep learning is a branch of machine learning that uses neural networks designed to process complex patterns found in large data sets.

ML and AI are utilized in visually impaired assistive system to create algorithm that can identify obstacles and give the user correct information. To identify items in the surroundings and provide the user with feedback, machine learning methods, for an instance, can be trained on substantial datasets of sensor data, like an ultrasound or infrared readings. Furthermore, algorithms that can pick up on user preferences and behaviours over time by learning from their reciprocity with the system can be created using AI and machine learning. Additionally, the development of visually impaired assistive system is greatly aided by deep learning. To provide a more thorough depiction of the user's surroundings, deep learning algorithms can be used to process photos and other sensory inputs. These algorithms can identify obstacles with a high degree of accuracy even in complicated and dynamic surroundings since they have been trained on massive datasets of photos and other sensory data. In conclusion, the primary technologies utilized in the creation of visually impaired assistive system is artificial intelligence, machine learning, and deep learning. These technologies enable these systems to function better overall, with increased precision, dependability, and user-friendliness. Visually impaired assistive system is being utilized to increase the freedom and safety of visually impaired people by utilizing AI, ML, and DL capabilities. These systems give visually impaired people the details they wanted to confidently navigate their surroundings.

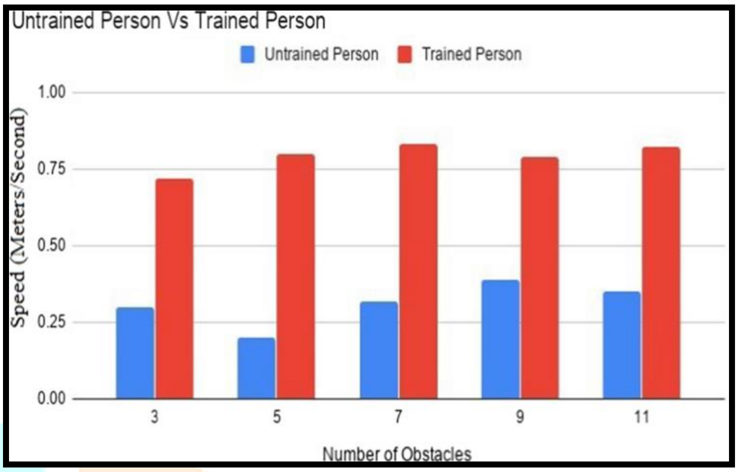

Fig.1. Obstacles Classification

I. PROPOSED SYSTEM

The proposed system will use three ultrasonic sensors to provide a panoramic field of view. Additionally, it will have a torch built in, which will activate upon detecting an object and provide a far sharper image. Additionally, The YOLO method should be used at a quicker pace of 45 frames per second.

ADVANTAGES: The Visually impaired assistive system enables people with vision impairments to perform their daily tasks without the assistance of guide dogs or other outside aids. This would assist those unlucky individuals in regaining some kind of normalcy and in leading independent lives.

DISADVANTAGES OF THE EXISTING SYSTEM: The only ultrasonic sensor used by the current systems can only detect obstacles in one direction. Moreover, object detection cannot capture a crisp picture at night or in the dark.

A. HARDWARE AND SOFTWARE REQUIREMENTS

When the analysis task is finished, the software requirements specification is produced. The software's assigned role and performance within the system engineering framework are currently going through a rigorous refinement process. This entails defining the system's performance requirements, design constraints, and appropriate validation criteria additionally to producing an extensive information description that accurately depicts the system's behaviour.

- A Raspberry Pi, a power supply, Pi camera, Ultrasonic Sensor
- Monitor (with micro-HDMI adaptor) (optional)
- A USB keyboard and mouse (optional)
- Headphones or speakers (optional)
- An ethernet cable (optional)
- Operating system: Linux(preferable) or windows 10
- Raspberry Pi OS, installed using the Raspberry Pi Imager

Operating System built especially for Raspberry Pi single-board computers, Raspberry Pi OS is a open-source and free operating system. Pre-installed apps, software development tools, and a graphical user interface are all included in this Linux distribution that in light of the Debian brand. Cost-effectiveness, small size and minimal power consumption, high degree of customization, and a large community for troubleshooting and assistance are just a few of the benefits of selecting Raspberry Pi over its competitors.

Ultrasonic sensors use high-frequency sound waves to create sound waves, which they then utilize to find echoes returning from surrounding objects. This can be implemented to close-quarters object detecting situations, such parking sensors in cars. Using computer vision techniques, details about the objects in digital photos or videos is extracted through analysis. Algorithms like edge detection, feature extraction, and template matching can be used for this. In general, the context, the technology being utilized, and the required level of accuracy and precision all affect object detection approaches.

B.ARCHITECTURE DIAGRAM

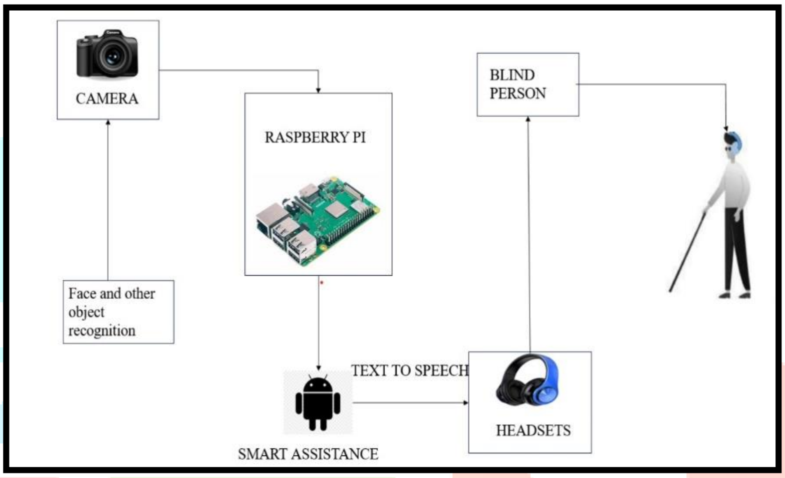

Fig. 2. Architecture Diagram

Real-time video footage of the surroundings is captured by the Pi camera. The video data is pre-processed, and pertinent features are extracted using the OpenCV package. For object detection, the pre-processed video data is exposed to the YOLO algorithm. The Yolo algorithm's output is processed, and additional picture and tasks involving recognition are completed using the TensorFlow framework. The ultrasonic sensor detects obstacles when they are close to the user. To locate items in the user's environment, the Raspberry Pi combines the object detection data with the output from the ultrasonic sensor. The technology helps the user navigate their environment more skilfully by giving them auditory feedback depending on the objects it has spotted.

a. Dataset

The COCO (Common Objects in Context) dataset is an extensive collection widely used in computer vision tasks, includes object detection, segmentation, and captioning. With over 200,000 images spanning 80 common object categories, it offers rich annotations such as pixel-level segmentation masks, bounding boxes, and labels. This dataset presents a diverse array of real time scenarios and contexts, making it invaluable for training models to handle varied visual challenges. Its annotations provide detailed information on object instances, aiding tasks like instance segmentation. Due to its size, detailed annotations, and diversity, COCO serves as a benchmark for evaluating and training state-of-the-art models, making it a cornerstone in computer vision research, competitions, and industry applications.

b. Pyttsx3 Library

A Python package called Pyttsx3 offers cross-platform text-to-speech (TTS) capabilities. It lets programmers add realistic-sounding voice output to their apps by instantly translating textual text to speech. A variety of TTS engines are backed by the library, including third-party choices like eSpeak and Festival in addition to the built-in voices for Windows and Mac OS X. With its straightforward API for managing voice selection, loudness, and speech rate, Pyttsx3 is an easy-to-use tool. Since it is open-source software and available under the MIT licence, anyone can use and alter it can be used for both personal and business purposes without cost. The following functions can be performed using YOLO, pyttsx3, OpenCV, ultrasonic sensors, and the algorithm when used on a Raspberry Pi for object detection:

YOLO Algorithm: This real-time object identification technique is capable of detecting objects. YOLO is a programme that may be employed to detect objects in photos or live video streams on a Raspberry Pi. Because of its processing complexity, optimising the YOLO technique for the Raspberry Pi can be difficult. There are many methods to increase the accuracy of the YOLO object detection model:

1. Expand the training dataset: More objects, variations, and backgrounds can be recognised by the system with the aid of a larger dataset. Make sure there are variety of angles, lighting settings, object sizes, and forms in the collection.

2. Adjust the pre-trained model: YOLO is pretrained on the OID dataset, a sizable collection of different objects. To help the model learn more about the particular items in your photographs, you can fine-tune it using your own dataset.

3. In order to assist the model learn more about the objects in your dataset and increase accuracy, increase the instances of training iterations.

4. Make use of data augmentation methods: These methods, which include rotation, flipping, and scaling, can aid to broaden the training data set and enhance the model's object recognition capabilities.

5. Use a larger network design: The model is capable of learning more complex features and increase accuracy by using a larger network architecture, such as YOLOv5 or YOLOv8.

Using a range of techniques, including regularisation techniques, anchor boxes, NMS, training on a big dataset, fine-tuning the model, and using high-quality photos, can enable YOLO achieve high reliability in object detection. This is another benefit of employing YOLO.

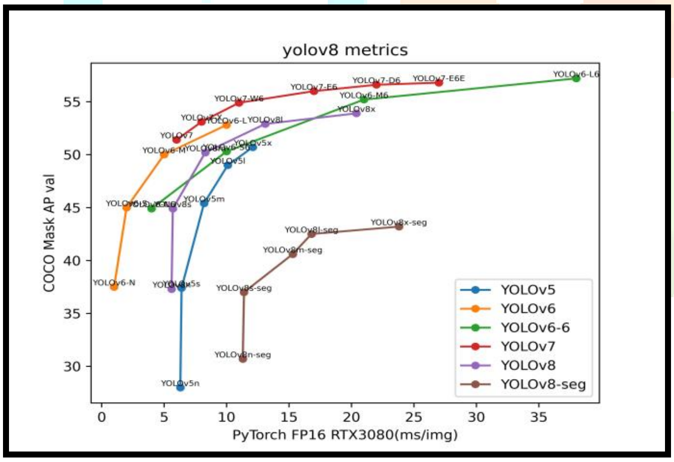

Fig. 3. YOLOv8 Algorithm Accuracy

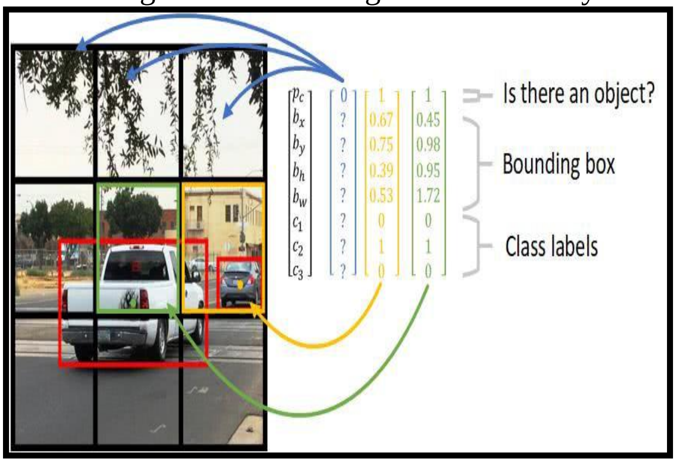

Fig. 4. Bounding Boxes

OpenCV: OpenCV, or Open Source Computer Vision Library, is an open-source computer vision and machine learning software library. It was originally developed by Intel and later maintained by the OpenCV community. OpenCV is written in C++ and has bindings for various programming languages, including Python, Java, and more, making it accessible to a extensive range of developers.

Ultrasonic Sensor: The distance between the Raspberry Pi and nearby objects can be determined using the ultrasonic sensor. By doing so, the Raspberry Pi can prevent collisions with neighbouring items and enhance the object detection process.

Pyttsx3: A Python package for literal to audio output conversion is called Pyttsx3. Pythonsx3 can be used to create audio alerts or notifications depending on the outcomes of the object detection algorithm when it is used for detection of object on a Raspberry Pi.

III. RESULTS AND DISCUSSION

A few trails were conducted to assess the precision and effectiveness of the project's design. The findings presented in this research demonstrate the application of our efforts to develop an accurate and reasonably priced travel aid that allows those who are blind people or visually impaired to interact with their daily surroundings. As was already mentioned, the embedded sensors gather data about the environment. Using a Pi camera and a Raspberry Pi 4b, the system was tested with the YOLOv8 object detection method. Chairs, tables, doors, people were among the many objects tm could accurately identify. The system was able to give precise bounding boxes encircling each object and overall showed good accuracy in detecting a range of objects. A few items were too tiny or partially obscured, though, and the system occasionally failed to identify them.

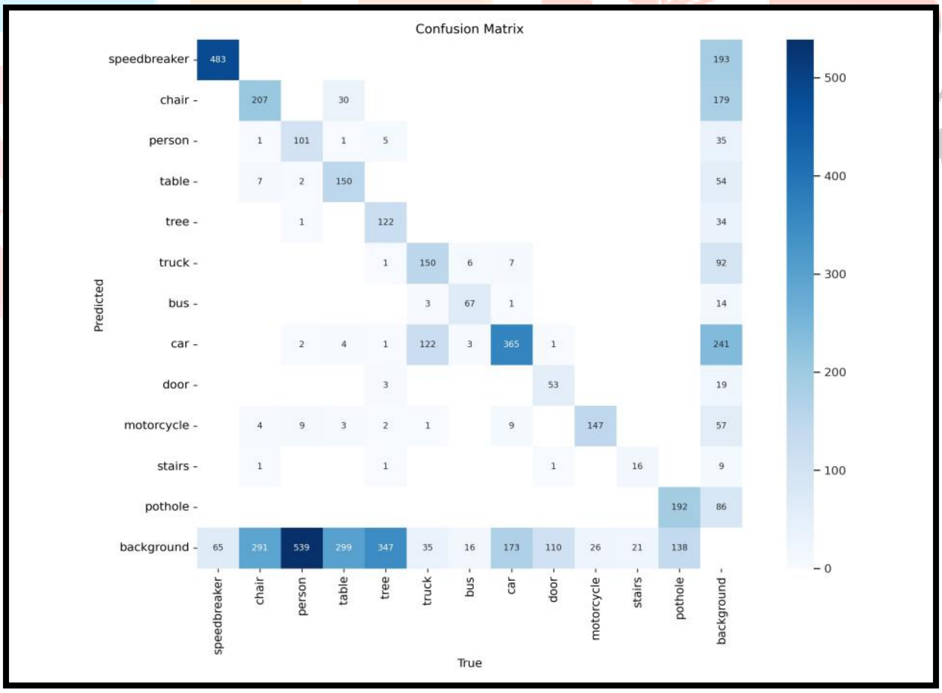

Fig. 5. Confusion matrix

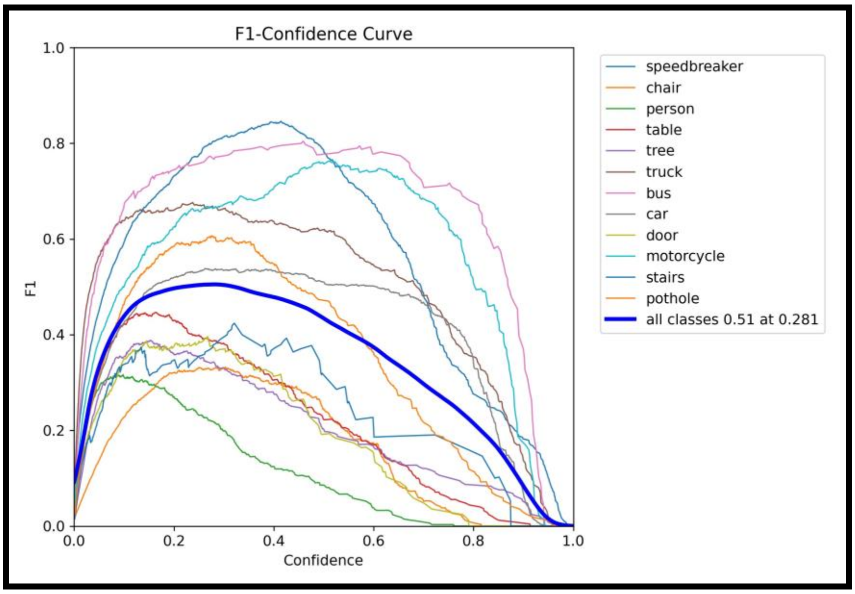

Fig. 6. F1-Curve

IV. CONCLUSION

A Raspberry Pi has been used to implement a picture-to-discussion change system. Using an amount of cases, the equipment yield has been tested and the reproduction results have been successfully verified. Our computation clearly forms the image and lays it forth. For people who lack vision or are blind, this is an useful and effective device. Upon analyzing multiple images, we have determined that our computation produces the desired result. This device is friendly to most people and conservative. Audio feedback helps educate the client and makes it easy to identify visually impaired people. It also uses an ultrasonic sensor to aid evade snags. For future enhancement a camera with more megapixels should be the goal of the project that will follow up on this article. We can create an application for user convenience rather than implementing the aforementioned project on a stick. In order to provide this device more functions, we may also try to install more sensors, like a gyroscope and a water sensor. Future project success will depend greatly on cost reduction because the raspberry pi is a costly component that makes it useless for those with little resources. Additionally, hiring a translator could make it easier for a blind individual to travel to a foreign country. To offer further features and personalization choices, the system may potentially be connected with a smartphone app.

REFERENCES

[1] Woojin Chung, Gunbee Kim, Munsang Kim and Chongwon Lee, "Integrated navigation system for indoor service robots in large - scale environments," IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004, New Orleans, LA, USA, 2004, pp. 5099 -5104, doi: 10.1109/ROBOT.2004. 1302526.

[2] Denis Tudor, Lidia Dobrescu, Drago Dobrescu,“ Ultrasonic Electronic System for Blind People Navigation”, Grigore T. Popa University of Medicine and Pharmacy, Iai, Romania, November 19 -21, 2015.

[3] Kanchan M. Varpe, M.P. Wankhade,” Visually Impaired Assistive System” International Journal of Computer Applications (0975 – 8887), Volume 77 – No.16 , September 2013.

[4] A. Aladrén, G. López-Nicolás, L. Puig and J. J. Guerrero, "Navigation Assistance for the Visually Impaired Using RGB-D Sensor with Range Expansion," in IEEE Systems Journal, vol. 10, no. 3, pp. 922 -932, Sept. 2016, doi: 10.1109/ JSYST.2014.2320639.

[5] N. Giudice, B. Tjan, G. Legge, R. Roy and P. Beckmann, "Digital Sign System for Indoor Wayfinding for the Visually Impaired," in 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, San Diego, California, 2005 pp. 30.

[6] Deng, Jun, Xiaojing Xuan, Weifeng Wang, Zhao Li, Hanwen Yao, and Zhiqiang Wang. "A review of research on objec t detection based on deep learning." In Journal of Physics: Conference Series, vol. 1684, no. 1, p. 012028. IOP Publishing, 2020.

[7] Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., Shen, H., Ren, J., Han, S., Ding, E., & Wen, S. (2020). PP - YOLO: An Effective and Efficient Implementation of Object Detector.ArXiv.

[8] Bochkovskiy, Alexey, Chien- Yao Wang and Hong- Yuan Mark Liao. " YOLOv4: Optimal Speed and Accuracy of Object Detection." Ar Xiv abs/2004.10934 (2020): n. pag.

[9] Redmon, Joseph, Santosh Kumar Divvala, Ross B. Girshick and Ali Farhadi. " You Only Look Once: Unified, Real -Time Object Detection." 2016 IEEE Conference on Computer Vision an d Pattern Recognition (CVPR) (2015): 779 -788.

[10] Talele, Ajay, Aseem Patil, and Bhushan Barse. "Detection of real time objects using TensorFlow and OpenCV." Asian Journal For Convergence In Technology (AJCT) ISSN-2350 -1146 (2019).

[11] Nikhil Mishra. " Image Text to Speech Conversion using Raspberry PI and OCR Techniques." International Journal for Scientific Research and Development 5.8 (2017): 523 - 525.

[12] Yadav, Avanish Vijaybahadur, Sanket Saheb Verma, and Deepak Dinesh Singh. "Virtual Assistant for blind people." 2021 International journal of advance scientific research and engineering trends 6, no. 5 (2021).