Object Detection for Blinds

DIPTI BANGDE[1], AFIYA MOHAMMAD[2], ADITI CHARLAWAR[3], ANUKSHA CHIKTE[4], DR. RAHILA SHEIKH[5]

[1, 2, 3, 4] *Students, Department of Computer Science & Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India*

[5] *Professor, Department of Computer Science & Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India*

Abstract*— Object recognition technology has transformed a number of industries, including industrial facilities and driverless cars. The people who need this technology the most—those who are visually impaired—haven't been able to use it well, though. Consequently, a machine learning-based object identification system created especially for the blind community is presented in this work. The system combines text-to-speech (TTS) technology to deliver a voice-guidance technique that transmits information about the objects around users, and it incorporates the Yolo V3 (You Only Look Once) algorithm for object detection. The suggested system's primary goal is to enable visually impaired people to recognize items in a given location on their own, without the need for outside assistance. Experiments are used to carefully examine the system's effectiveness and performance to verify its precision and efficiency. The suggested method uses technical jargon like object recognition, machine learning, the Yolo V3 algorithm, and image classification techniques to identify and extract features from video frames and classify them into the appropriate groups. The COCO dataset is used. In summary, the proposed object detection system combines voice assistance and state-of-the-art deep learning techniques to let visually impaired people identify objects on their own. Experiments are used to assess the system's performance and show off its usefulness and potential in practical settings.

***Index Terms*— Yolo V3, Voice Feedback, pyttsx3, Object Detection, COCO dataset.**

I. INTRODUCTION

The swift advancement of Information Technology (IT) has incited multiple research initiatives with the goal of addressing daily annoyances and offering personal conveniences to people. Though there are still many barriers for visually impaired people to overcome, the two biggest ones are traversing inside settings and obtaining object information. The development of object detection technologies for the blind is a crucial field of study that aims to enhance their independence and quality of life. For those who are blind or visually impaired, object detection is essential because it allows them to safely and freely navigate their environment. This can involve identifying items and providing details about them, including their name, through the use of cameras and computer vision algorithms. In order to give the user aural input and enable safer and more independent navigation and interaction with their surroundings, voice guidance technology is also integrated. In order to get precise object information and locate objects, this study suggests a deep learning object recognition technique. Using a camera, things are detected using the You Only Look Once (YOLO) architecture, a cutting-edge object identification machine learning model. This dissertation describes a deep learning-based system that analyzes exact object location and information. With the use of the pyttsx3 Python module, audio guiding is added to improve the system's usability for visually impaired people. This gives the user quick access to pertinent information about their environment by enabling the system to synthesis verbal announcements of object detections in real-time. The suggested system is a useful tool for helping the visually impaired in their daily lives because it provides them with more independence and safety when navigating their surroundings thanks to voice direction.

II. LITERATURE SURVEY

1.This study included the standard datasets along with an explanation of the current detection model approaches. This study examined many detectors, including one- and two-stage models, which aided in the analysis of various object detection techniques and also gathered some novel and conventional

applications. A few branches pertaining to object detection were also mentioned. Additionally, several development trends were highlighted to are simultaneously predicted. This makes it especially appropriate for real-time uses, such helping people who are blind or visually handicapped. help follow the state-of-the-art algorithm and advance the process.

2.In this paper, a completely convolution network based on regions was given. For accurate and effective object detection, R-FCNN is employed. For object detection, this article can therefore readily incorporate ResNets like fully convolution image classifier backbones. An effective yet straightforward RFCNN framework for object detection was presented in this paper. When compared to the quicker R-FCNN, this approach obtains the same accuracy. This aided in the adoption of the cutting-edge image classification frameworks.

3. This challenge is regarded as the standard for object detection and classification. This study categorized and identified over a million photos and over 100 object categories. The large-scale data collection procedure was presented in this publication. He talked about the success and failure of various algorithms in addition to describing the most effective approach for this set of data.

4. The study's findings demonstrated that, for human recognition, grids with directed gradient outperform the available feature set.

III. PROPOSED SYSTEM

The deep learning system known as "YOLOv3," or "You Only Look Once version 3," is renowned for its quickness and precision in object detection in photos. The input image is divided into a grid, and each grid cell's bounding boxes and class probabilities are simultaneously predicted. This makes it especially appropriate for real-time uses, such helping people who are blind or visually handicapped.

The YOLOv3 algorithm will be executed in real-time by the system using the processing capabilities of a laptop. The YOLOv3 model will process a live video feed that is captured by the laptop's camera of the user's surroundings. The algorithm will determine the bounding box coordinates and labels (e.g., "person," "chair") as items are recognized.

The system will produce voice output to notify the user of the things it The proposed system aims to assist visually impaired individuals in detecting objects in their surroundings using YOLOv3, a state-of-the-art object detection algorithm, implemented on a laptop. This system will provide real-time auditory feedback to the user, enabling them to navigate their environment more safely and independently.

The deep learning system known as "YOLOv3," or "You Only Look Once version 3," is renowned for its quickness and precision in object detection in photos. The input image is divided into a grid, and each grid cell's bounding boxes and class has detected once they have been identified. The laptop's speakers or headphones will be used to transmit this audio input. For instance, if the system detects a human nearby, it will announce "person." In a similar manner, it will notify the user when other items are spotted, facilitating better navigation and interaction with their surroundings.

The system will give priority to responsiveness and efficiency in order to offer the user real-time support. Optimisation methods that lower the amount of computing power needed for inference, like model quantisation, can be used to improve laptop hardware performance.

The system's implementation will entail writing software that combines the laptop's audio output and video feed with the YOLOv3 algorithm. This software will be compatible with a variety of laptops because it is made to run on widely used operating systems like Windows.

YOLOv3 on a laptop, as part of the suggested object detection for the blind system, provides a workable and efficient way to improve the safety and freedom of people with visual impairments. The device gives users the confidence to explore and engage with their environment more confidently by giving them real-time aural input about their surroundings.

IV. METHODOLOGY

It should be extremely quick to detect, recognize, and communicate objects to the user. The response time of the suggested system is important because it is meant for people who are visually impaired. It will be ineffective to wait to inform the user of the obstacles. The suggested solution makes use of the Yolo V3 algorithm to get around problems with processing speed and delay. Compared to several other real-time object detection methods, the Yolo V3 approach is faster.

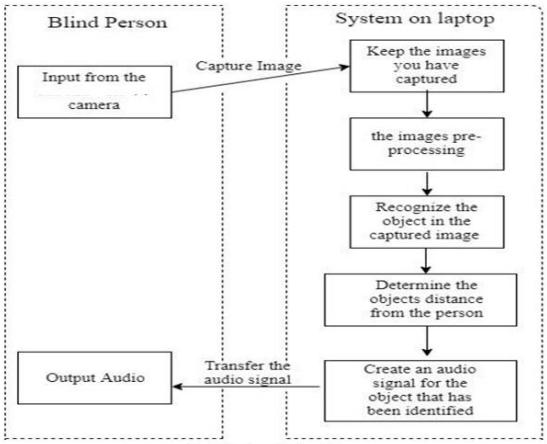


Fig 1. System Methodology

Using YOLOv3 on a laptop, the suggested system for object detection for the blind with speech output can be put into practice in a few simple steps.

1.Research and Preparation: The first stage entails a detailed investigation of current object detection algorithms that are appropriate for real-time applications, as well as the hardware implementation of these algorithms on laptops. This entails learning about the features and capabilities of the YOLOv3 algorithm as well as comprehending the unique requirements and difficulties that visually impaired people encounter. Acquiring the required software development tools, a laptop with the right specs, and any extra hardware, including cameras and audio output devices, is another aspect of preparation.

2. Data Collection and Annotation: An extensive dataset of photos featuring a range of things pertinent to the user's surroundings must be gathered in order to train and optimize the YOLOv3 model for object detection. Bounding boxes specifying the location and class of each object featured in these photographs must be added. Particular consideration should be paid to incorporating elements that visually impaired people frequently come across, such as pedestrians, obstructions, and landmarks.

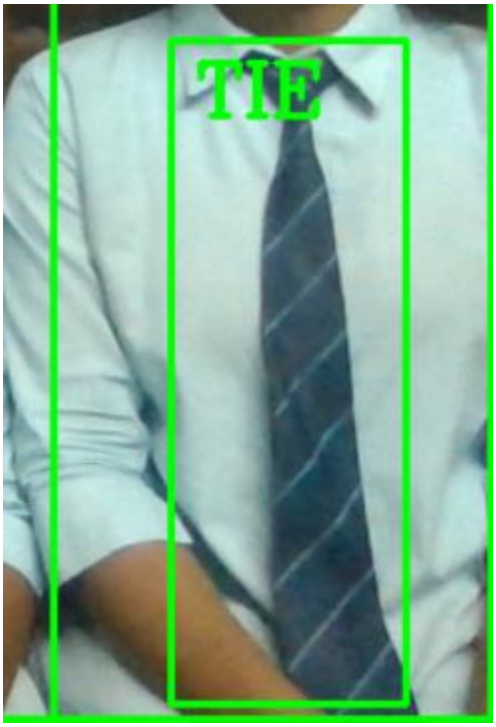
3. Software Development: The creation of the user interface and integration between the laptop camera feed, audio output capabilities, and YOLOv3 model will commence concurrently with the model training process. Programming modules are used in this to record live video from the laptop's camera, process the video frames for object detection using the YOLOv3 model that has been trained, and produce appropriate voice output in response to the items that are recognized.

4. Testing and Evaluation: After the program is produced, extensive testing will be carried out to guarantee its accuracy, usability, and functionality. In order to assess how well the system performs in identifying different items and giving the user accurate audio input, testing entails modeling real-world events. Subsequent iterations of the system will be informed by identifying any flaws or areas for improvement based on user feedback and testing results.

5.Deployment and User Training: The technology will be made available to visually impaired people for usage in the real world following a successful testing and refinement phase. There will be user training sessions to acquaint users with the capabilities and functionalities of the system and make sure they can use it to navigate their surroundings. The effectiveness and usefulness of the system will be ensured by constant monitoring and support for any technical problems or user input.

To put it briefly, a thorough strategy that includes research, data collecting, model training, software development, testing, deployment, and user training is used in the methodology for putting the suggested system into practice. Developing a strong and user-friendly solution that enables visually impaired people to safely and independently explore their surroundings requires completing each stage.

V. OUTPUT



CONCLUSION

Numerous assstive technologies have been created recently to help the blind and visually impaired identify objects in their environment and navigate about more easily. However, the majority of currently available solutions are pricy, extremely complex, challenging to use, intended as specialized assistance, necessitate training, etc. Our main objective is to provide the visually impaired with an accessible, inexpensive, easy-to-use assistive solution that will enable them move around independently, comprehend their surroundings, and understand the world around them. The real time video stream is captured by the smartphone camera and then the label of the object detected using object detection algorithm is communicated to the user in speech through the speakers or headphones. The audio output communicated to the user would help them in performing their day to day activities and lead a quality life.

REFERENCES

[1] Dr. P. Boobalan1 , Bhuvanikha. S2 , Sivapriya. M3 , Sivakumar. R4, Object Detection with Voice Guidance to Assist Visually Impaired Using Yolov7, International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 11 Issue IV Apr 2023; 764.

[2] D. Ravi Kumar, Hiren Kumar Thakkar, Suresh Merugu, Vinit Kumar Gunjan, and Suneet K. Gupta, Object Detection System for Visually Impaired Persons Using Smartphone, January 2022; 1635-1638.

[3] Devi RA, Uthaman I, Shanthanam RR (2020) An IOT Security based electronic aid for visually impaired detection with navigation assistance system, p 6

[4] Jadhav PSP, Tomy S, Jayswal SS et al (2016) Object detection in android smartphone for visually impaired users. 5:332–334. https://doi.org/10.17148/IJARCCE.2016.51171

[5] Kuriakose B, Shrestha R, Sandnes FE (2020) Tools and technologies for blind and visually impaired navigation support: a review. IETE Tech Rev (Institution Electron Telecommun Eng India) 1–16. https://doi.org/10.1080/02564602.2020.1819893

[6] Chen LB, Su JP, Chen MC, et al (2019) An Implementation of an intelligent assistance system for visually impaired/blind people. In:

2019 IEEE International conference on consumer electronics ICCE 2019, pp 4–5. https://doi.org/10.1109/ICCE.2019.8661943

[7] Sait U, Ravishankar V, Kumar T et al (2020) Design and development of an assistive device for the visually impaired. Procedia Comput. Sci. 167:2244–2252. https://doi.org/10.1016/j.procs.2020.03.277

[8] Matusiak K, Skulimowski P, Strurnillo P (2013) Object recognition in a mobile phone application for visually impaired users. In: 2013 6th International conference on human system interaction (HIS), pp 479–484. https://doi.org/10.1109/HSI.2013.6577868

[9] Anand S, Kumar A, Tripathi M, Gaur MS (2019) Human face detection enabled smart stick for visually impaired people. Springer, Singapore

[10] Anitha J, Subalaxmi A, Vijayalakshmi G (2019) Real time object detection for visually challenged persons. Int. J. Innov. Technol. Explor. Eng. 8:312–314