

# Object Detection, Convert Object Name to Text and Text to Speech

Ms. Sathya G<sup>1</sup>, Mr. Kamesh S<sup>2</sup>, Mr. Rishi Kumar S<sup>3</sup>, Mr. Saai Sabapathi<sup>4</sup>

<sup>1</sup>Assistant Professor, Department of Computer Science and Engineering

<sup>2 3 4</sup>Students, Department of Computer Science and Engineering

SRM Valliammai Engineering College, Chennai, Tamil Nadu, India

**Abstract:** *Visually impaired persons (VIPs) comprise a significant portion of the population, and they are present around the globe and in every part of the world. In recent times, technology proved its presence in every domain, and innovative devices assist humans in their daily lives. This work presents a smart and intelligent system designed to aid visually impaired persons (VIPs) in mobility and safety. Utilizing real-time navigation through automated voice assistance, VIPs can sense and understand their surroundings, facilitated by a deep learning model for object detection and recognition. The system includes a hardware component that automatically alerts family members and shares the VIP's location and incident details in case of abnormal object recognition. This ensures the safety of VIPs while upholding their privacy. This innovative solution enables VIPs to visualize their environment, empowering them with increased security. The employed deep learning model demonstrates high accuracy in object detection and recognition, enhancing the overall effectiveness of the system.*

**Keywords:** Visually impaired persons

## I. INTRODUCTION

The World Health Organization (WHO) has reported that 285 million of the world's population is blind or visually impaired. Out of these, 39 million are blind. The major diseases that cause visual impairments include refractive error, glaucoma, trachoma, corneal, opacities, cataracts, diabetic retinopathy, and unaddressed presbyopia. Visually impaired persons (VIPs) face difficulties in performing activities of daily living (ADLs) e.g. opportunity of work and schooling, moving in their surroundings, ability to interact with the environment, and searching for common objects (indoor/outdoor) on their own or even with some assistance[1].

The main challenges for VIPs are/object detection and recognition, currency identification, textual information (sign, symbol) and translation, mobility/navigation, and safety. One of the most important organs in human body are eyes. We enjoy the beauty of nature, various types of books, and many other aspects of our lives. We can go anywhere independently and have fun with friends and family . What if we are blind? Forgot about enjoying, what if we don't even do our own work independently? What if we must depend on some others for regular daily works? It is difficult to think and imagine these kinds of situations. However, some of us in society are visually impaired. They must depend on others for their regular work. The ability to visualize the surroundings is a gift.

Huge number of research and developments are going in the domain of machine learning and object detection .Large number of new kind of tools are also came in to the existence .Few of those developments are similar to our idea .But all those projects implementation has distinctions and differences in object detection like using of different algorithms and different libraries for the processing . The dataset contain nearly 90 object names which are useful and observed by a common man in our day to day life , Which is enough for the real time object detection. we use YOLO algorithm for object detection and Text to Speech conversion technique for voice alerts.

## II. METHODOLOGY

The methodology for the proposed Detection of object would involve the following steps:

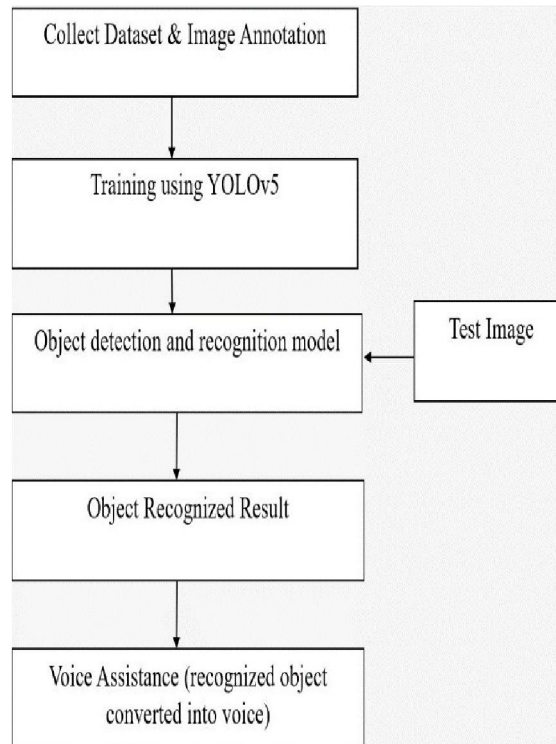


Fig. 1. Architecture diagram of Detection of Object Detection, Recognizing Text and Speech

### Dataset Collection & Annotation

Datasets are collected from public sources such as Pascal VOC and Kaggle.

After collecting the dataset, annotation will be done using the makesense.ai online tool. Makesense.ai is a free-to-use online tool for labeling images for object detection.

### Train YOLOv5

After annotation, it will export into YOLOv5 format. For YOLOv5 format conversion, the online tool roboflow is used, which generates a correctly formatted custom dataset. Roboflow includes dataset splitting, image resizing, and image augmentation. We provide a unique deep-learning framework for object detection and recognition in this work. In this work, YOLOv5 deep learning model is used to detect and recognize the object. In this network YOLOv5 model takes the 11 extracted features as input from the CSPDarknet53 backbone model. This backbone is comprised of 53 layers.

### Object Detection and Recognition

Once the model has been trained, an input image will be provided and then given to the YOLOv5 detection module to detect and recognize the object based on the trained model.

### Voice Assistance

The recognized object is converted into voice, providing live guidance and descriptive information to visually impaired persons (VIPs) about their surroundings. This feature enhances their mobility and safety by offering real-time assistance and valuable insights, ensuring a more informed and secure navigation experience

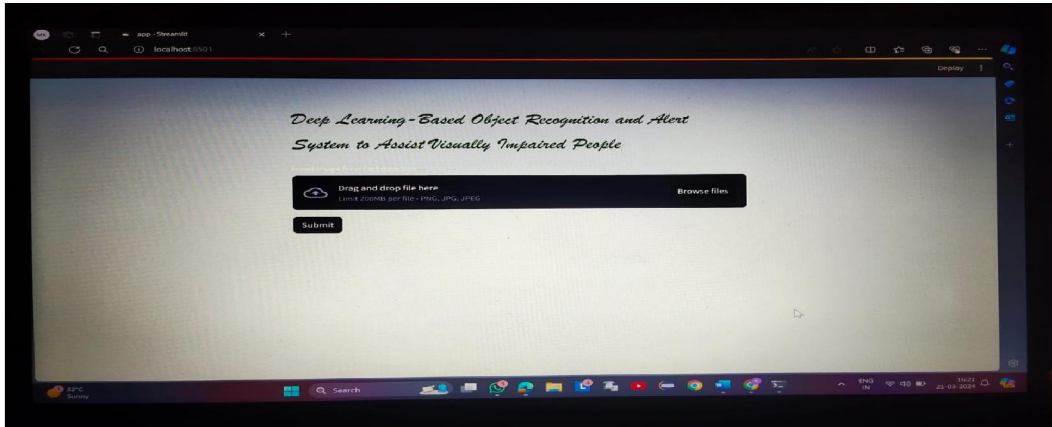


Fig. 2. The User Interface of Object detection convert Object Name to Text and Text to Speech

### III. RESULTS

The implementation of the Object Detection has yielded compelling results, showcasing the efficacy of advanced deep learning techniques in accurately identifying and classifying the lesions.

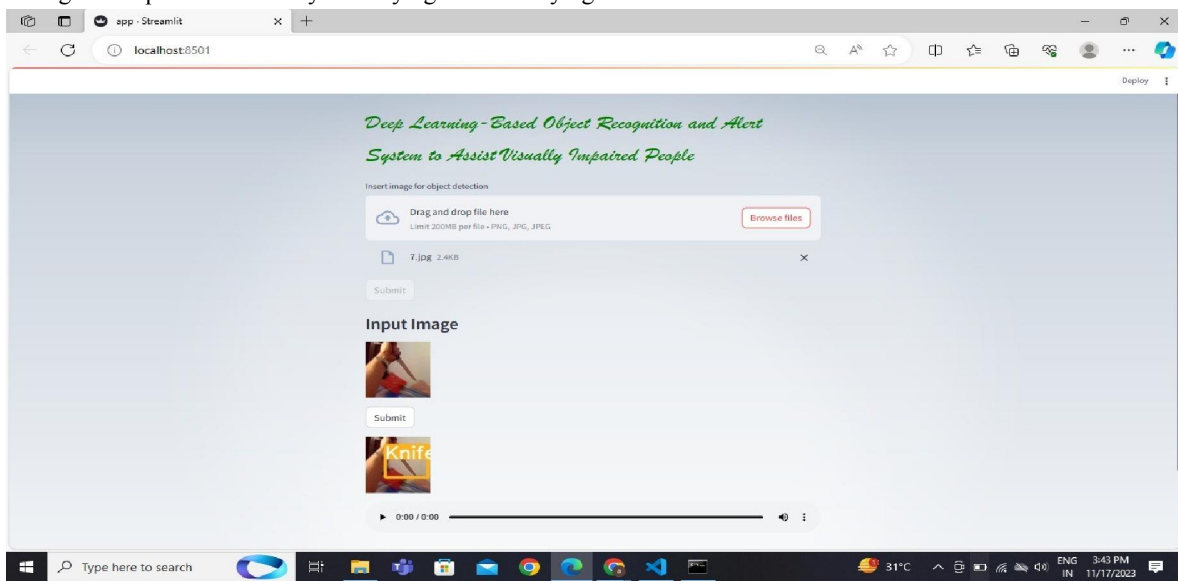


Fig. 3. Detection and voice of the object name

Furthermore, with the help of Natural Language Processing, the model achieved an impressive and robust performance in distinguishing the object name and voice. The model has image segmentation capabilities and can be able to identify and label the object name and speech recognition.

For implementation, you can use programming languages like Python and libraries such as OpenCV for image processing, TensorFlow or PyTorch for object detection, and libraries like pyttsx3 or gTTS for text-to-speech conversion. Integration with APIs like Google Cloud Vision for object recognition or Google Text-to-Speech for TTS is also an option

This project aligns with the broader goal of enhancing object name, then convert object name to text and text to speech. This feature was helpful to visually impaired people to understand the name of the object

### IV. CONCLUSION

This Deep Learning based object recognition and alert system to assist visually impaired people by generating voice alerts will be helpful to the blind person because they can stand on their own, and work on their own. The voice alerts

keeps the person warned about the things that he/she is surrounded so that accidents get reduced. Depending on third person will be reduced. By this assistance system a blind person can easily take a step forward and move into the society.

They can recognize objects around them and sense the natural environment using YOLOv5-based Object detection Algorithm. Moreover, a web-based application is developed to ensure the safety of VIPs. The user of this application can detect and recognize the object with the voice assistance. The user can interact with the system using the web application. This application gets the input image and detects the given test image with voice output. The experimental analysis shows that the proposed system provided satisfactory results and outperformed other devices in terms of supported features.

### REFERENCES

- [1] K. Marapalli, A. Bansode, P. Dundgekar And N. Rathod, "Aiger An Intelligent Vehicle For Military Purpose", Pp. 1052-1057, 2022
- [2] L.-B. Chen, J.-P. Su, M.-C. Chen, W.-J. Chang, C.-H. Yang, And C.-Y. Sie, "An Implementation Of An Intelligent Assistance System For Visually Impaired/Blind People," In Proc. Ieee Int. Conf. Consum. Electron. (Iccec), Pp. 1-2, Jan. 2021.
- [3] W.-J. Chang, L.-B. Chen, C.-Y. Sie, And C.-H. Yang, "An Artificial Intelligence Edge Computing-Based Assistive System For Visually Impaired Pedestrian Safety At Zebra Crossings," Ieee Trans. Consum. Electron., Vol. 67, No. 1, Pp. 3-11, Feb. 2021
- [4] M. A. Rahman And M. S. Sadi, "Iot Enabled Automated Object Recognition For The Visually Impaired," Comput. Methods Programs Biomed. Update, Vol. 1, May 2021.
- [5] M. A. Sayeed, S. P. Mohanty, E. Kougianos, And H. P. Zaveri, "Eseiz: An Edge-Device For Accurate Seizure Detection For Smart Healthcare," Ieee Trans. Consum. Electron., Vol. 65, No. 3, Pp. 379-387, Aug. 2020.
- [6] M. M. Islam, M. S. Sadi, K. Z. Zamli, And M. M. Ahmed, "Developing Walking Assistants For Visually Impaired People: A Review," Ieee Sensors J., Vol. 19, No. 8, Pp. 2814-2828, Apr. 2020.
- [7] M. M. Rahman, M. M. Islam, S. Ahmmed, And S. A. Khan, "Obstacle And Fall Detection To Guide The Visually Impaired People With Real Time Monitoring," Social Netw. Comput. Sci., Vol. 1, Pp. 1-10, Jul. 2020.
- [8] W.-J. Chang, L.-B. Chen, M.-C. Chen, J.-P. Su, C.-Y. Sie, And C.-H. Yang, "Design And Implementation Of An Intelligent Assistive System For Visually Impaired People For Aerial Obstacle Avoidance And Fall Detection," Ieee Sensors J., Vol. 20, No. 17, Pp. 10199-10210, Sep. 2020.
- [9] R. Tapu, B. Mocanu, And T. Zaharia, "Wearable Assistive Devices For Visually Impaired: A State Of The Art Survey," Pattern Recognit. Lett., Vol. 137, Pp. 37-52, Sep. 2020
- [10] Z. Bauer, A. Dominguez, E. Cruz, F. Gomez-Donoso, S. Orts-Escolano, And M. Cazorla, "Enhancing Perception For The Visually Impaired With Deep Learning Techniques And Low-Cost Wearable Sensors," Pattern Recognit. Lett., Vol. 137, Pp. 27-36, Sep. 2020.