ISSN 2277-2685
IJESR/Jan-Mar. 2025/ Vol-15/Issue-1s/357-365
Yandapalle Akshaya *et. al.*, / International Journal of Engineering & Science Research

AI Based Object Classifier For Blind People With Voice

[1]Ms B Jyothsna, [2]Yandapalle Akshaya, [3]Balagoni Anusha, [4]Grandhi Bhavana
[1]Assistant professor, Electronics and Communication Engineering, BRECW
[2]B.Tech Student, Department of Electronics and Communication Engineering, BRECW

ABSTRACT

As we can see, there are numerous blind persons nearby who encounter various challenges, such as difficulty in crossing roads and identifying objects in their environment. With the advancement in technology in several fields, human life is evolving to better standards. Unfortunately, those who are blind are unable to fully enjoy this kind of lifestyle. So, this project is one strategy for introducing blind individuals to a new way of living that makes them independent on others. The major goal of this project is to create a deep-learning algorithm that can be used to analyse the environment for people who are blind by using the rapidly evolving technology. We\'ll accomplish this using object detection and transform the data into speech alerts and warnings. Real-time object detection is one of the more challenging tasks since it requires continuous processing and takes a long time. The convolution neural network is the main backbone for any type of object detection (CNN). We can create algorithms based on photos and videos by employing a convolution neural network. We utilise the YOLO technique for object detection because it is simple and quick to process. In addition, for the voice warnings, we employed Text to Speech (TTS). The dataset used in this technique is the COCO dataset, which contains the names of things and objects in our daily lives. These algorithms have been thoroughly trained by the over 90 outdoor objects that we view every day in our daily lives.

1-INTRODUCTION

Living independently in a visually driven world presents significant challenges for individuals with visual impairments. While technological advancements have made leaps in improving accessibility, the need for seamless, real-time assistance remains critical. An AI-based object classifier for the blind with voice assistance addresses this gap by offering an intuitive and practical solution that enhances day-to-day independence.This system leverages advanced computer vision and deep learning technologies to identify objects in real-time, such as household items, personal belongings, or outdoor obstacles. Once identified, the device provides immediate audio feedback, describing the object to the user in a clear and concise manner. The solution combines the power of machine learning algorithms with robust speech synthesis, ensuring accurate recognition and intuitive interaction.Equipped with a compact, user-friendly design, this tool is not only portable but also highly adaptable to various environments, such as homes, offices, or public spaces. Users can benefit from enhanced spatial awareness, making tasks like identifying groceries, distinguishing doorways, or locating personal items significantly more manageable. Beyond functionality, this innovation promotes inclusivity, independence, and confidence for individuals with visual impairments. By harnessing the capabilities of AI, this solution is poised to transform how visually impaired individuals interact with their surroundings, fostering an equitable and empowering experience for all.

2-LITERATURE SURVEY

The development of an AI-based object classifier for the visually impaired with voice assistance draws from various advancements in computer vision, machine learning, natural language processing, and accessibility technologies. Below is a summary of the relevant literature and technologies explored in the project:

- **(Redmon et al., 2016) YOLO (You Only Look Once)**: Real-time object detection models such as YOLO provide high-speed and accurate object classification, making them suitable for real-time applications for visually impaired users.
- **(Ren et al., 2015) Faster R-CNN**: Known for its accuracy in object detection, this model uses region proposal networks for efficient object classification.
- **MobileNet and EfficientNet**: Lightweight models like MobileNet and EfficientNet are optimized for deployment on portable devices with limited computational resources, crucial for user-friendly solutions.

3- ARTIFICIAL INTELLIGENCE (AI)

Artificial Intelligence (AI) refers to the simulation of human intelligence in machines designed to think, learn, and problem-solve autonomously. It enables computers and software to perform tasks such as speech recognition, visual perception, decision-making, and language translation. At its core, AI includes key areas such as machine learning, which allows machines to improve through experience; natural language processing (NLP) for understanding and generating human language; and computer vision, which enables machines to interpret and analyze visual data.

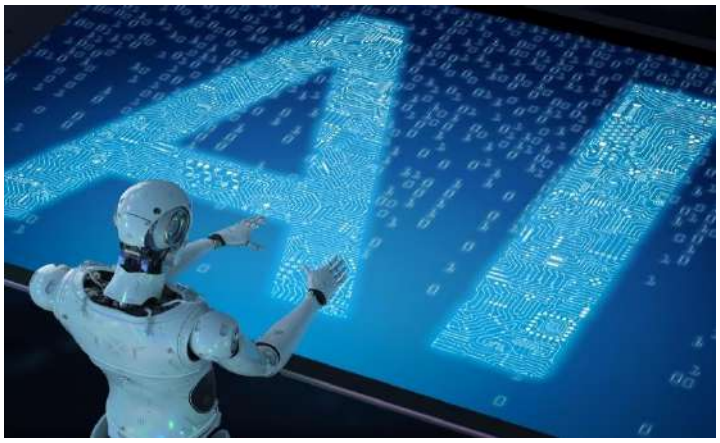

Fig 2.1: Artificial Intelligence (AI)

AI systems operate through various techniques like neural networks, deep learning, and reinforcement learning, allowing them to solve complex problems and adapt over time. With the growth of big data, AI has become essential for analyzing vast amounts of information and extracting meaningful insights. Industries like healthcare, automotive, education, and entertainment are leveraging AI to automate processes, enhance user experiences, and optimize decision-making.

Despite its potential, AI also raises concerns regarding job displacement, data privacy, security, and ethical considerations, such as the risk of algorithmic bias. Addressing these issues while advancing AI technology remains a critical focus. As AI continues to evolve, it is expected to

revolutionize countless aspects of daily life, from personalized services to complex problem-solving, marking a transformative shift in how we interact with technology.

4- WORKING

The flow chart and working of the system provide a comprehensive understanding of how the AI-powered object classification solution functions to assist visually impaired individuals. The flow chart visually represents the step-by-step process, starting from capturing video input to delivering audio feedback to the user. Each component is interconnected, showcasing how the system integrates advanced AI technologies like object detection models and real-time processing.

The working process highlights the seamless operation of the system, beginning with video capture, followed by object detection using pre-trained models such as COCOSSD or YOLO, and concluding with audio conversion for user accessibility. This structured approach ensures clarity, efficiency, and real-time responsiveness, making the system practical and user-friendly. The flow chart serves as a guide to understanding the intricate yet intuitive mechanism driving the project, while the working description elaborates on the technical and functional aspects.

Existing system

Existing systems designed to assist visually impaired individuals face significant challenges in providing a holistic solution for navigation and object recognition. Traditional aids, such as walking canes and guide dogs, are limited in their scope. While they assist in detecting obstacles, they do not offer detailed information about the surrounding environment or identify specific objects. These tools also rely heavily on the user's physical effort or a guide dog’s training, which may not always be adequate in complex or unfamiliar scenarios.

Fig: Person with White canes and Guide dog

Modern electronic aids, such as ultrasonic sensors, tactile feedback devices, and simple obstacle detection systems, have introduced some technological advancements. However, these systems are often confined to detecting proximity rather than providing object-specific identification. For instance, while they may alert a user to the presence of an object, they cannot distinguish whether it is a chair, a table, or a hazard like a sharp object. This lack of contextual understanding limits their effectiveness in enabling visually impaired individuals to fully interact with their environment. AI-based solutions have started to address these gaps, but they are far from perfect. Many existing AI systems require sophisticated and expensive hardware, such as smart glasses or wearable devices, which are not accessible to most users due to high costs. Additionally, these systems

frequently rely on cloud-based processing for tasks like object detection and classification, which introduces latency and dependency on stable internet connections. This reliance can be a major drawback in rural or low-resource settings where internet connectivity is unreliable. Moreover, current AI systems often struggle with real-time performance. Delays in processing video input or detecting objects can lead to missed opportunities for users to react promptly, particularly in dynamic and fast-changing environments. Accuracy is another concern; many systems may fail to differentiate between similar objects or perform poorly in low-light conditions, which are common in real-world use cases.

Proposed system

The proposed system, "AI-Based Object Classifier for Blind People with Voice," is a groundbreaking solution designed to assist visually impaired individuals by translating visual data from their surroundings into descriptive audio feedback. This system seeks to bridge the gap between the visual and auditory senses, enabling users to gain an understanding of their environment and empowering them with greater independence in their daily lives. At the heart of the system is a real-time object detection mechanism powered by advanced machine learning and computer vision technologies. It begins by capturing video input through an intuitive interface built using ReactJS, offering ease of use and accessibility. The video data is processed using TensorFlowJS, which integrates powerful pre-trained object detection models

like COCOSSD and YOLO. These models enable the system to recognize and classify objects within the video frames with exceptional speed and accuracy, even in complex or dynamic environments. The identified objects are organized into a queue for systematic processing, ensuring that no detection is missed or overlooked, even when multiple objects are identified simultaneously. After processing, the system converts the visual information into audio descriptions that are delivered to the user through a voice output module. This transformation of visual input into auditory cues is designed to provide clear, concise, and contextually relevant information, helping users navigate and interact with their surroundings effectively.

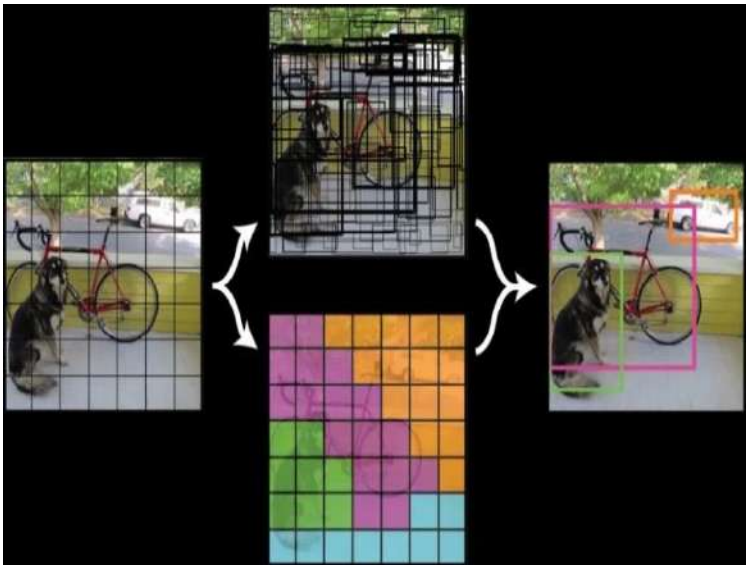

Fig: Shows Object detection Using Yolo

The integration of OpenCV further strengthens the system by offering robust tools for video and image processing, ensuring that the system performs efficiently under various conditions, such as changes in lighting or background noise. This project combines the power of AI, machine learning, and accessible design to create a solution that is not only technologically advanced but also practical for everyday use. By delivering a seamless and intuitive experience, this system has the potential to significantly enhance the quality of life for visually impaired individuals, giving them the tools they need to independently explore and understand the world around them.

5-IMPLEMENTATION OF SOFTWARE

The project involved analyzing the design of few applications so as to make the application more users friendly. To do so, it was really important to keep the navigations from one screen to the other well ordered and at the same time reducing the amount of typing the user needs to do. In order to make the application more accessible, the browser version had to be chosen so that it is compatible with most of the Browsers.

Python IDLE:

Python IDLE (Integrated Development and Learning Environment) is a simple, user-friendly environment bundled with Python installations, designed primarily for beginners to write, test, and debug Python code. It provides a graphical user interface (GUI) that includes an interactive Python shell, where users can execute Python commands one at a time, and a script editor to write and save longer programs. IDLE comes with essential features such as syntax highlighting, auto-completion, and debugging tools, including stepping through code and setting breakpoints. While it's lightweight and straightforward, making it ideal for learning and small projects, IDLE is not typically used for large-scale development, where more powerful integrated development environments (IDEs) like PyCharm or VS Code are preferred. Nonetheless, IDLE remains a valuable tool for beginners and for quickly testing small snippets of code.

Django

Django is a high-level, open-source web framework written in Python, designed to help developers build robust, scalable, and secure web applications quickly. Released in 2005, Django follows the "Don't Repeat Yourself" (DRY) principle and promotes the reuse of code, making development efficient and maintainable. It provides a clean and pragmatic design, handling much of the complexity involved in web development tasks like URL routing, database management, and user authentication out of the box. Django comes with an extensive set of built-in features such as an admin interface, form handling, and security mechanisms, including protection against common web vulnerabilities like SQL injection and cross-site scripting (XSS). It is based on the Model-View-Template (MVT) architectural pattern, which helps organize code in a modular way. Django's flexibility, combined with its scalability, makes it a popular choice for small and large projects alike, powering websites like Instagram, Pinterest, and Mozilla.

Flow Chart

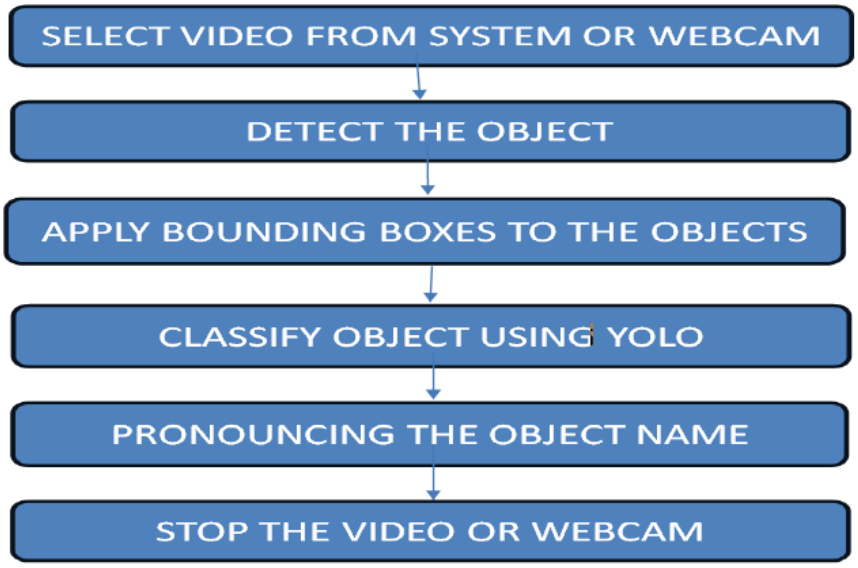

Fig 4.1 Flow Chart of AI based object classifier for blind with voice

6-ADVANTAGES, DISADVANTAGES AND APPLICATIONS

Advantages

1. Real-Time Accuracy: With TensorFlowJs and the COCOSSD model, the system can detect and identify multiple objects in real-time, providing accurate and immediate feedback to users.
2. Enhanced Accessibility: The user-friendly interface built with ReactJs ensures that the system is easy to navigate, making it accessible for users of all ages and technological proficiency levels.
3. Portability: As a web-based application, the system can be accessed on various devices, including smartphones, tablets, and computers, allowing users to utilize it wherever they go.

Disadvantages

1. Complexity Technical Issues:

- Battery Life: Continuous use of camera and voice synthesis can drain battery quickly, which is problematic for mobile or wearable devices.
- Connectivity: Some systems rely on internet connections for cloud-based processing, which might not be available in all areas.

2. Privacy Concerns:

- Public Spaces: Announcing detected objects in public can reveal private information or attract unwanted attention.
- Sensitive Situations: Voice feedback might be inappropriate in certain situations, like quiet environments (libraries, meetings) or confidential settings.

3. Dependence on Voice:

- Noisy Environments: In noisy settings, voice output can be difficult to hear and distinguish, reducing the effectiveness of the system.

Applications

1. **Daily Navigation:** Helping users navigate through familiar and unfamiliar environments by identifying objects, obstacles, and landmarks in real-time.

2. **Indoor Navigation:** Assisting in indoor environments like homes, offices, shopping malls, and airports by providing detailed information about the surroundings.
3. **Shopping Assistance**: Identifying products, reading labels, and recognizing prices in stores, allowing users to shop independently.

7-RESULTS

In this chapter, we will discuss about the results of the AI Based Object Classifier for Blind People with Voice. Presenting the results from implementing the system, including object detection accuracy, system performance in various conditions, and real-world test cases. This chapter will also contain analysis on how the system improves the user experience for visually impaired individuals.

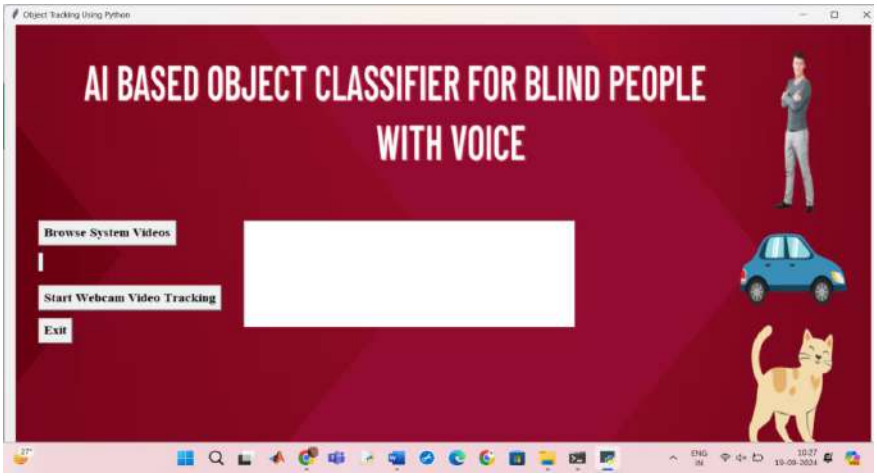

Fig 6.1: Proposed Method GUI

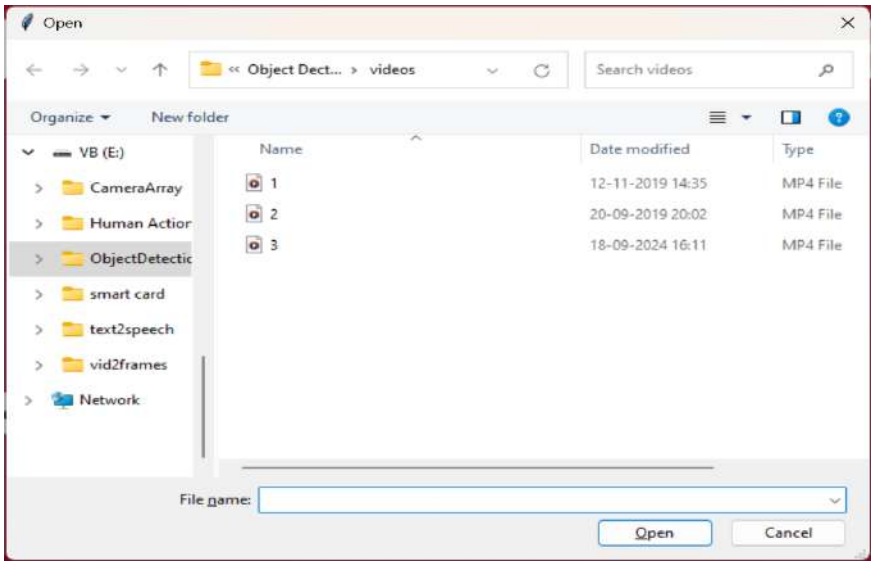

Fig 6.2: Browse the video from user

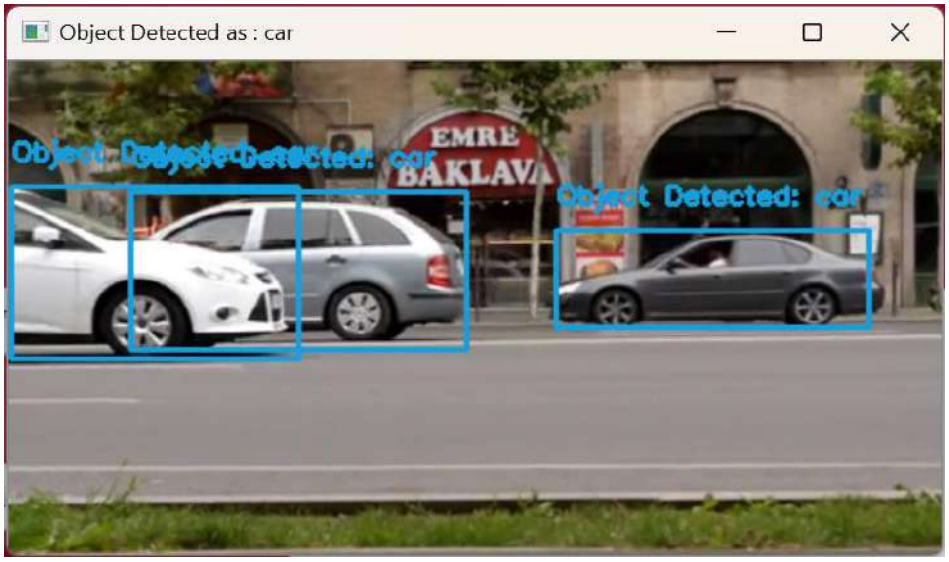

Fig6.3: In above video Object detected as: car

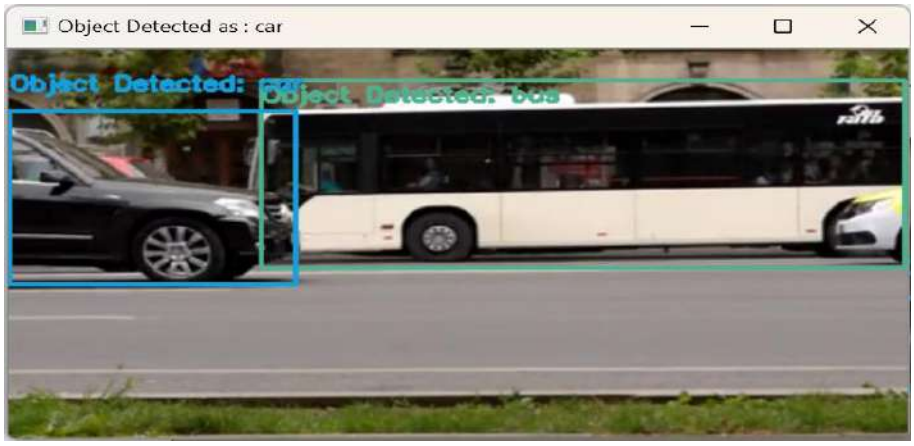

Fig 6.4: Above video Object detected as: Bus and Car

Fig 6.5: In above video Object detected as: Person and car

8-CONCLUSION

In this project we used image recognition, voice generation modules for the development of the project. As of now accuracy is good but in case if we want to increase the accuracy we have to train the model with more object/images in the dataset. This project is a small experiment which is useful for blind persons, can be able to find the objects which are surrounded by them, and they are in a position of taking care of themselves when they are outside. The ability of the blind person to stand alone and carry out tasks independently makes this blind assistance device useful for object detection by voice warnings. The device\'s camera serves as the blind person\'s virtual eye, capturing every detail of their environment. The voice alerts keep the person informed about the surroundings so that accidents are decreased. Reduced rely on other parties. There are so many people present in the world who are visually impaired and illiterate from different parts of the world. Some of them do not understand other languages except their local language in their local accent. so, one of the future scope for this project is to develop the idea in which voice alerts in such a way that they can use their own local language.

REFERENCES

[1] Miss Rajeshvaree Ravindra Karmarkar, Prof.V.N. Honmane "Object Detection System For The Blind With Voiceguidance"- Published Online June 2021 in Ijeast

[2] Jigar Parmar, Vishal Pawar, Babul Rai, Prof. Siddhesh Khanvilkar "Voice Enable Blind Assistance System -Real time Object Detection"- IRJET, Apr 2022

[3] Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam "Real-Time Object Detection with Yolo"- IJEAT,Feb 2019

[4] Mayuresh Banne, Rahul Vhatkar, Ruchita Tatkare "Object Detection and Translation for Blind People Using Deep Learning"- IRJET, Mar 2020

[5] M. I. Thariq Hussan, D. Saidulu, P. T. Anitha, A. Manikandan and P. Naresh "Object Detection and Recognition in Real Time Using Deep Learning for Visually Impaired People"- IJEER,June 2022

[6] Priya Kumari, Sonali Mitra, Suparna Biswas, Sunipa Roy, Sayan Roy Chaudhuri, Antara Ghosal, Palasri Dhar, Anurima Majumder "YOLO Algorithm Based Real-Time Object Detection"- IJIRT, June 2021

BOOKS FOR AI and ML:

"Python Machine Learning" by Sebastian Raschka and Vahid Mirjalili.

"Deep Learning for Computer Vision with Python" by Adrian Rosebrock**.**