TE-IT(VI), Mini Project Report

On

“Third Eye for the Visually Impaired”

Submitted in fulfillment of the requirement of University of Mumbai

<u>TE-IT SEM VI for subject: Mini Project</u>

For the degree of

Bachelor of Engineering

(Information Technology)

By

Rishikesh Bhor - TU4F2122024

Ayan Joshi - TU4F2122030

Arya Sawant - TU4F2122032

Aarya Paradkar - TU4F2122067

Under the guidance of

Dr. Sujata Kadu

DEPARTMENT OF INFORMATION TECHNOLOGY

TERNA ENGINEERING COLLEGE

UNIVERSITY OF MUMBAI

SH-2023 (2023-2024)

APPROVAL SHEET

The Lab Journal submitted **Rishikesh Bhor - TU4F2122024 , Ayan Joshi - TU4F2122030, Arya Sawant - TU4F2122032, Aarya Paradkar - TU4F2122067** by is approved for the partial fulfillment of the requirement for the award of the Semester VI- ML Mini Project for degree of **"Bachelor of Engineering"** in **"Information Technology"** from University of Mumbai.

…………………………………………..

Internal Examiner

Dr. Sujata Kadu

…………………………………………

External Examiner(s)

Date:

Place: Terna Engineering College, Nerul

CERTIFICATE

This is to certify that **Rishikesh Bhor - TU4F2122024 , Ayan Joshi - TU4F2122030, Arya Sawant - TU4F2122032, Aarya Paradkar - TU4F2122067** has submitted the "Project Report" of "ML Mini Project" to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of "Bachelor of Engineering" in "Information Technology"

Dr. Sujata Kadu

Supervisor/Guide

Dr. Vaishali Khairnar

Head of Department

Dr.L.K. Ragha

Principal

TABLE OF CONTENTS

TABLE OF CONTENT

Appendix

- **Plagiarism**

ACKNOWLEDGEMENT

No project is ever complete without the guidance of those expert who have already traded this past before and hence become master of it and as a result, our leader. So, we would like to take this opportunity to take all those individuals how have helped us in visualizing this project.

We express our deep gratitude to our project guide ***Dr. Sujata Kadu*** for providing timely assistant to our query and guidance that he gave owing to his experience in this field for past many years. He has indeed been a lighthouse for us in this journey.

We extend our sincerity appreciation to all our Professors form Terna Engineering College for their valuable inside and tip during the designing of the project.

Their contributions have been valuable in so many ways that we find it difficult to acknowledge of them individual. We also great full to our HOD Dr. Vaishali Khairnar, Principal Dr. L. K. Ragha for extending their help directly and indirectly through various channel in our project work.

ABSTRACT

Visual impairment presents significant challenges to the independence and mobility of individuals, profoundly impacting their ability to perceive and interact with the surrounding environment. The limitations imposed by visual impairment can hinder daily activities and diminish quality of life. Recent years have witnessed remarkable advancements in machine learning (ML) techniques, offering promising solutions to address these challenges and enhance the autonomy of the blind and visually impaired community.

One groundbreaking solution emerging from these advancements is the development of ML-based third eye systems. These systems harness the power of computer vision algorithms to interpret and provide real-time auditory or tactile feedback about the surrounding environment. By leveraging ML technology, these systems aim to revolutionize the way visually impaired individuals navigate and perceive their surroundings, offering a new level of independence and mobility.

The primary objective of ML-based third eye systems is to provide individuals with low vision real-time information about their surroundings. Through the use of Bluetooth headphones, these systems can identify objects in the user's vicinity, aiding them in crossing roads and performing other daily activities safely. Additionally, these systems aim to measure distances between objects, providing users with enhanced spatial awareness. This functionality helps users understand the layout of their surroundings and navigate more effectively.

2.1 Introduction:

The project is dedicated to empowering visually impaired individuals by leveraging cutting-edge machine learning algorithms, thereby fostering their independence and confidence. Visual impairment often poses significant challenges, limiting the ability of individuals to perceive and interact with their surroundings. However, recent advancements in machine learning offer promising solutions to enhance the autonomy of the blind and visually impaired community.

One such solution is the development of ML-based systems that use image recognition techniques to provide real-time analysis of the user's surroundings. These systems process visual data and offer instantaneous verbal descriptions through a seamless voice interface. By providing auditory feedback about the environment, these systems enrich the user's understanding, enabling them to interact independently with everyday objects and navigate with newfound autonomy.

The core objective of this project is to bridge the gap between visual information and auditory feedback, enabling individuals with visual impairments to navigate the world more confidently. By providing real-time descriptions of their surroundings, these systems empower users to make informed decisions and navigate safely. This initiative not only offers practical assistance but also aims to instill a sense of empowerment and self-reliance among visually impaired individuals, contributing to a more inclusive and welcoming society for all.

2.2 Objectives:

- ***Enhanced Navigation:***
  Develop a system that allows visually impaired individuals to navigate their surroundings with increased independence and confidence, leveraging advanced machine learning algorithms.

- ***Real-time Object Recognition:***
  Implement state-of-the-art image recognition techniques to enable the system to accurately identify and interpret visual information in real-time.

- **Multi-language Support:**
  Integrate a multi-language support feature to cater to a diverse user base, ensuring accessibility for individuals with different linguistic backgrounds.

- **Object Detection with Auditory Feedback:**
  This feature will enable the system to audibly announce the detected object, providing real-time auditory feedback to the user. By leveraging advanced sound synthesis techniques, the system will generate distinct and recognizable sounds corresponding to different objects

LITERATURE SURVEY

SR No.	Name	Description	Author	Date
1.	Third Eye: Object Recognition and Speech Generation for Visually Impaired	Detecting and recognizing the objects and generating speech about the objects helps visually impaired in a great way in understanding their surroundings. Required a mechanism to assist the visually impaired person to travel independently with the ability to identify objects in their path, and the ability to generate speech describing the objects detected in the scene. This can be achieved with the help of YOLOv5 image detection model and text to speech converters such as gTTS and pyttsx3 modules in python.	Koppala Guravaiah	203
2.	Third Eye: Assistive Technology for the Visually Impaired	Third Eye project is a wearable aid device for assisting anyone with visual impairments. It contains multiple sonar sensors to track approaching objects and a camera that identifies faces according to a user-entered database. Facial detection is used in order to store people in a contact list and inform the user who is approaching. The intent is for the visually impaired individual to know who they are talking to or who is approaching them.	Tasmiyah Qazi	2019
3.	A third eye with human-computer interaction for the visually impaired	In order to reduce dependency of the visually impaired people on others resulting from their visual impairment, an ergonomic, portable, vocalizable and refreshable Braille reader was developed. Texts from a scanned document or any text imported from a computer are transmitted to the device developed, enabling visually handicapped persons to read a text in the Braille format by using their hands.	Fatih Basciftci	2017
4.	The Third Eye: An AI Mobile Assistant for Visually Impaired People	Today, many hardware products and mobile devices are developed to provide useful features, but their effectiveness and practicability do not fully help visually impaired people. They are dependent on people around them for their daily needs. If no one is available next to them, they become helpless to do many simple tasks independently. Thus, this paper is intended for visually impaired people, who can become more self-reliant with the help of the Third Eye android mobile application. They will be able to identify objects around them and find the object they need with the YOLO object detection algorithm.	R.S. Hariharan	2023

PROBLEM STATEMENT

Visual impairment significantly impacts individuals by limiting their ability to perceive and interact with their surroundings, affecting their independence and mobility. The inability to see obstacles, read signs, or discern spatial cues independently restricts their freedom to move and explore. This reliance on others for navigation and daily tasks can lead to feelings of frustration, dependency, and reduced autonomy.

Visually impaired individuals often rely on external assistance, such as sighted guides or guide dogs, to navigate through unfamiliar or complex environments. While these aids are essential, they are not always readily available and can act as a temporary solution. This constant need for support and intervention can hinder spontaneous decision-making and self-reliant mobility, further emphasizing the challenges faced by visually impaired individuals in achieving independence.

Moreover, the reliance on external assistance can also lead to social and psychological impacts, as individuals may feel a loss of control over their lives and a sense of isolation. The inability to navigate independently can limit their social interactions and participation in activities, affecting their overall quality of life. Addressing these challenges requires innovative solutions that empower visually impaired individuals to navigate and interact with their environment autonomously, enhancing their independence and mobility.

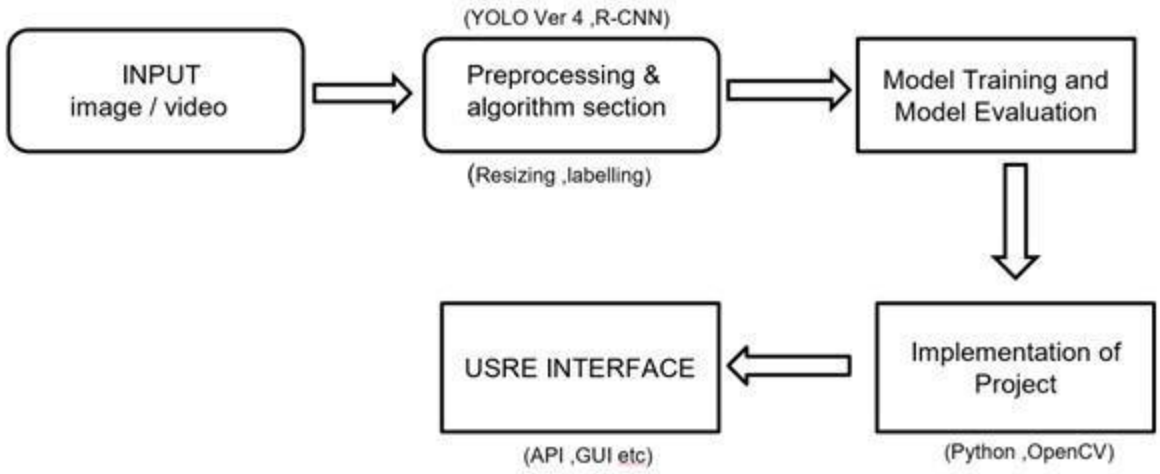
PROPOSED METHODOLOGY

The system is designed to empower visually impaired individuals by fostering a newfound sense of autonomy. It achieves this by enabling independent recognition and interaction with everyday objects, ultimately allowing users to navigate confidently and interact with their environment without constant assistance. A key feature of the system is its ability to provide auditory feedback, which plays a crucial role in enhancing the user's understanding of their surroundings.

In addition to auditory feedback, the system also offers tactile feedback to further enrich the user's interaction with their environment. This tactile feedback provides an additional layer of information, allowing users to more effectively navigate and interact with objects around them. Through a voice interface, the system provides instantaneous verbal descriptions of the user's environment. This real-time feedback not only enriches the user's understanding but also enables them to navigate with enhanced independence.

To achieve these functionalities, the system leverages state-of-the-art machine learning algorithms and advanced image recognition techniques. By processing visual data from a camera, the system provides real-time analysis of the user's surroundings. This capability allows visually impaired individuals to gain a comprehensive understanding of their environment and navigate with increased confidence. Overall, the system is designed to enhance the independence and mobility of visually impaired individuals, empowering them to lead more autonomous lives.

BLOCK DIAGRAM:



1. Input: This represents the initial input to the system, which could be visual data captured by a camera or another sensor. The input is essential for the system to process and analyze the user's surroundings.

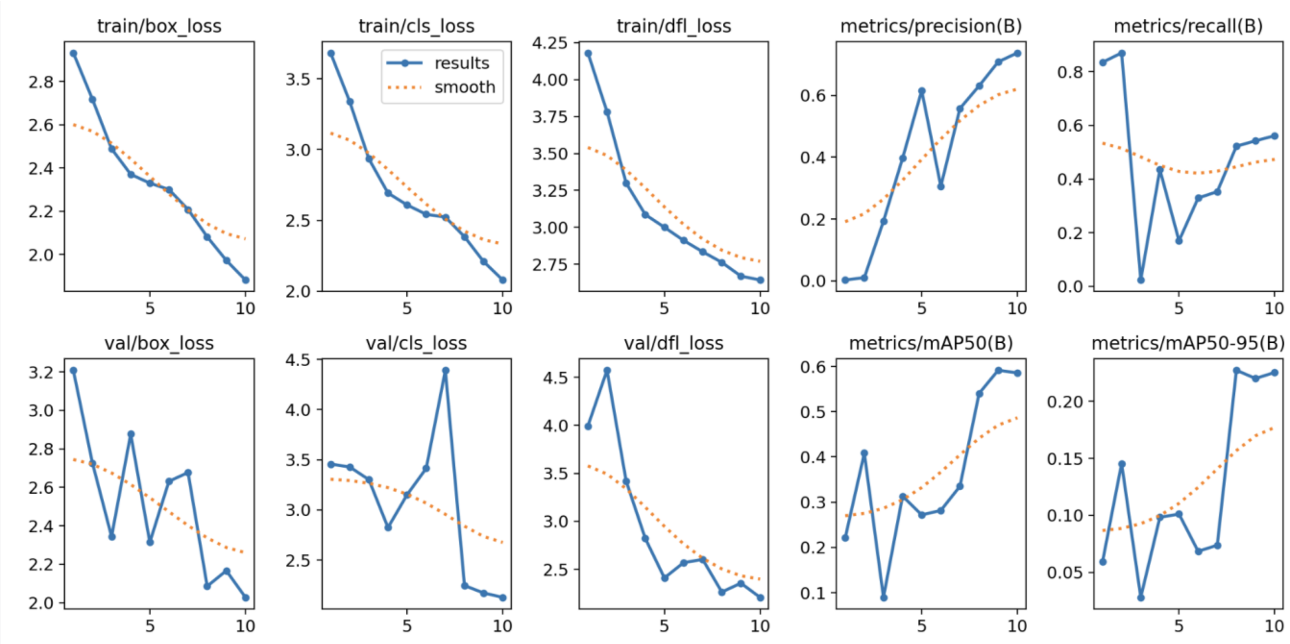
2. Preprocessing and Algorithm (YOLOv5): In this stage, the input data undergoes preprocessing to enhance its quality and prepare it for further analysis. YOLOv5, a state-of-the-art object detection algorithm, is then applied to detect objects in the preprocessed data. YOLOv5 stands for "You Only Look Once," and it is known for its efficiency and accuracy in real-time object detection.

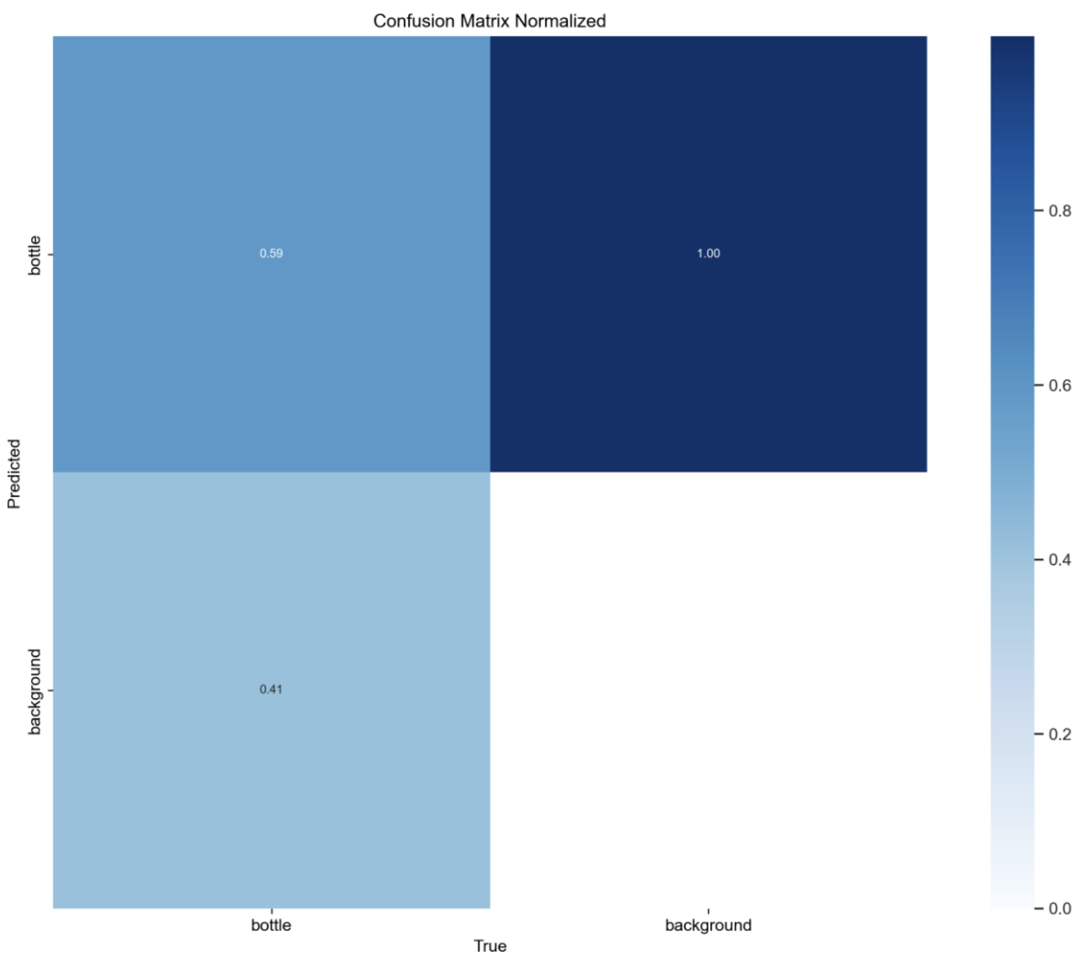
3. Model Evaluation and Training: This stage involves evaluating the performance of the YOLOv5 model on the detected objects. It may also include fine-tuning or training the model with additional data to improve its accuracy and adaptability to different environments. Model evaluation and training are crucial for ensuring that the system can accurately identify objects in various scenarios.

4. Implementation of Project (Python and OpenCV):The implementation of the project involves using Python, a popular programming language, and OpenCV (Open Source Computer Vision Library) to develop the system. Python provides a robust and flexible environment for developing machine learning applications, while OpenCV offers a wide range of tools and functions for image processing and computer vision tasks.

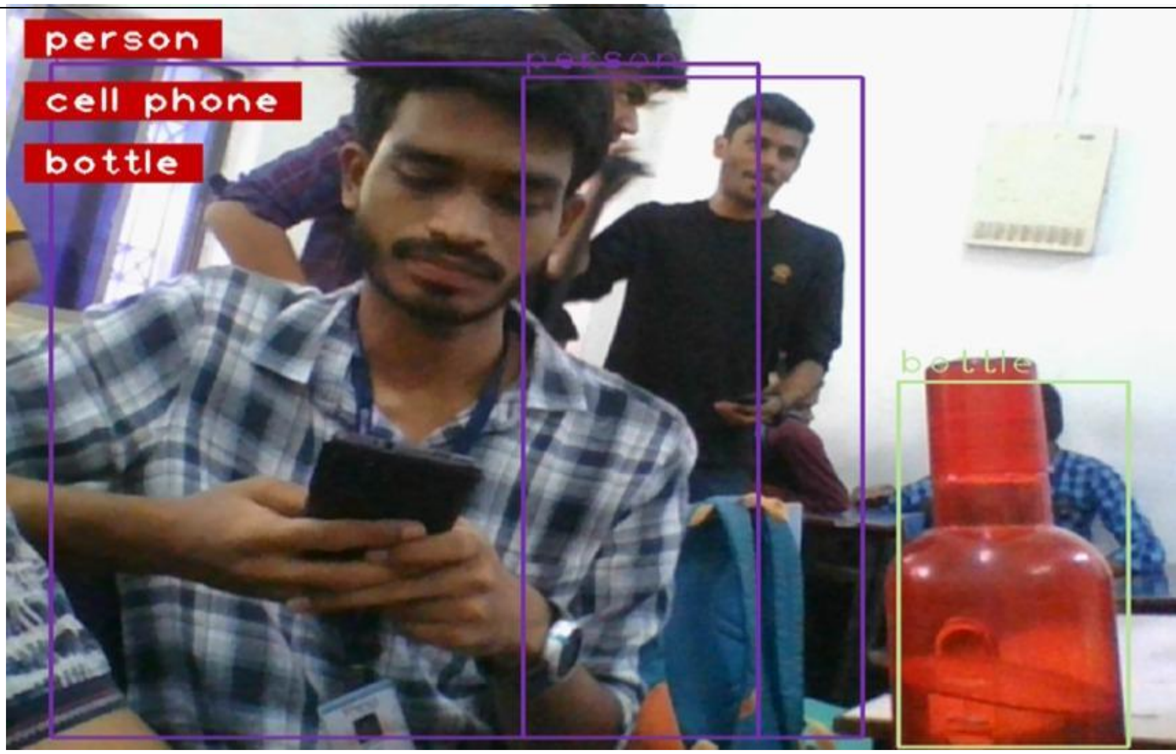
5. User Interface (GUI): The user interface (GUI) is the component of the system that allows users to interact with the application. It provides a graphical interface through which users can receive auditory or tactile feedback about their surroundings. The GUI plays a crucial role in enhancing the user experience and making the system more accessible to visually impaired individuals.

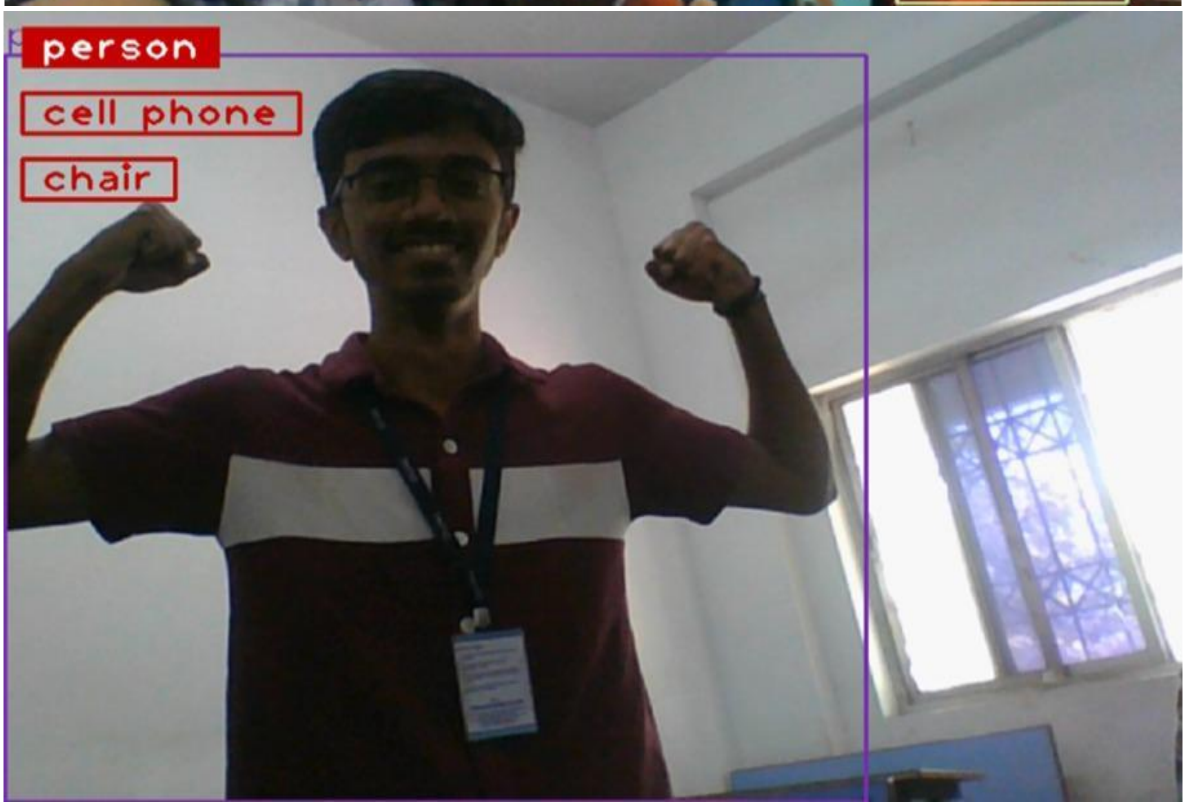
Implementation:





Result:


person
cell phone
bottle
person
bottle


person
cell phone
chair


person
cell phone
bottle
chair
fork
(497, 409)
cell phone


person
chair
cell phone
bottle
chair
fork

FUTURE SCOPE

1. **Enhanced Algorithm Development:**

Future research can focus on refining machine learning algorithms to improve accuracy and robustness in interpreting visual data, especially in challenging environmental conditions.

2. **Integration with Emerging Technologies:**

Integration with emerging technologies such as augmented reality (AR) and artificial intelligence (AI) can augment the capabilities of third eye systems, enabling more intuitive and context-aware feedback for users.

3. **Wearable Device Innovation:**

Further innovation in wearable devices can contribute to the development of more ergonomic and user-friendly solutions, ensuring greater comfort and convenience for individuals with visual impairments.

4. **Integration with edge computing devices**:

Devices such as smartphones and wearables, for real-time object detection and recognition in mobile applications.

CONCLUSION

ML-based Third Eye Systems leverage machine learning (ML) and computer vision to interpret visual environments in real-time, empowering individuals with visual impairments with increased autonomy and mobility. These systems face significant challenges, including improving algorithm accuracy across diverse conditions such as varying lighting and object textures. Ensuring reliable and useful information for users requires advanced algorithms capable of handling various visual scenarios effectively.

Another key challenge is optimising the computational efficiency of these systems. Given the need for real-time feedback, ML algorithms must process and interpret visual data swiftly. This demands not only powerful algorithms but also efficient hardware and software integration to minimise latency and ensure a seamless user experience.

Moreover, designing these systems for user-friendliness is paramount. Individuals with visual impairments should find them easy to use and understand. This involves intuitive interfaces, clear audio feedback, and customisable settings to cater to individual preferences and needs. Enhancing user experience and adoption relies heavily on the usability and accessibility of these systems.

In conclusion, ML-based Third Eye Systems hold immense potential in assisting individuals with visual impairments, but they must address challenges related to algorithm accuracy, computational efficiency, and user-friendliness to realise their full impact.

REFERENCES (in IEEE formats)

[1] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "**You only look once: Unified, real-time object detection**." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788. 2016

[2] Girshick, Ross. "Fast r-cnn." In *Proceedings of the IEEE international conference on computer vision*, pp. 1440-1448. 2015.

[3] Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pp. 21-37. Springer International Publishing, 2016.

[4] Law, Hei, and Jia Deng. "Cornernet: Detecting objects as paired keypoints." In *Proceedings of the European conference on computer vision (ECCV)*, pp. 734-750. 2018.

[5] Lin, Tsung-Yi, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. "Focal loss for dense object detection." In *Proceedings of the IEEE international conference on computer vision*, pp. 2980-2988. 2017.

[6] Dai, Jifeng, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. "Deformable convolutional networks." In *Proceedings of the IEEE international conference on computer vision*, pp. 764-773. 2017.

[7] Tian, Zhi, Chunhua Shen, Hao Chen, and Tong He. "Fcos: Fully convolutional one-stage object detection." In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9627-9636. 2019.

[8] Tan, Mingxing, Ruoming Pang, and Quoc V. Le. "Efficientdet: Scalable and efficient object detection." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10781-10790. 2020.