AI Powered Glasses for Visually Impaired People Using Object Detection

Shivani Jangle[1], Divya Hutke[2], Tanmay Hiwanj[3], Ayush Kahar[4]

[1]Shivani Jangle Department of Information Technology, SSGMCE Shegaon
[2]Divya Hutke Department of Information Technology, SSGMCE Shegaon.
[3]Tanmay Hiwanj Department of Information Technology, SSGMCE Shegaon.
[4]Ayush Kahar Department of Information Technology, SSGMCE Shegaon

---***---

Abstract - The advancement of artificial intelligence and machine learning has opened new possibilities for assistive technologies aimed at enhancing the lives of visually impaired individuals. This paper focuses on developing AI-powered smart glasses using Raspberry Pi and YOLOv8 object detection to provide real-time environmental awareness. The system incorporates a USB camera to capture surroundings, and a machine learning model for face recognition, which identifies known individuals and announces their names via an audio output. The audio signal is either played at a low volume or through earphones for private listening. The Raspberry Pi processes input data, detects objects and faces, and converts the visual information into speech, improving the user's spatial understanding. This innovation fosters independent navigation and social interaction, offering a cost-effective solution for visually impaired individuals. By leveraging real-time object detection, the system enhances accessibility, enabling users to identify obstacles, recognize people, and interact more confidently with their surroundings.

Key Words: Object Detection, YOLOv8, AI, Face Recognition.

I. INTRODUCTION

Visual impairment poses significant challenges in daily life, making navigation, interaction, and social recognition difficult for affected individuals. Traditional mobility aids, such as white canes and guide dogs, offer some assistance but lack the ability to convey detailed environmental information. With recent advancements in artificial intelligence, object detection, and face recognition technologies, there is an opportunity to develop smart assistive devices that enhance the autonomy of visually impaired users.

This paper introduces an AI-powered wearable smart glasses system that leverages Raspberry Pi and YOLOv8 object detection to provide real-time object and face recognition. The system uses a USB camera to capture live surroundings, processes the data using a deep learning model, and announces detected objects and recognized faces via an audio signal. The user receives feedback through a speaker or earphones, ensuring privacy and accessibility.

The coupling of yolov8 (You Only Look Once version 8) simplest appearance as soon as model eight helps fast precise item detection permitting customers to navigate challenging environments with safety moreover the face recognition version lets in customers to decide recognized people by way of saying their names out loud improving social interplay this era-based totally answer seeks to empower visually impaired humans by last the gap between belief and interaction in their daily lives.

PROBLEM STATEMENT

Visually impaired humans come across difficulties in perceiving their surroundings, identifying limitations and spotting people, proscribing their mobility and autonomy. traditional solutions, which includes canes and guide dogs, aren't very useful and do not offer longer situational facts about surroundings, what required is a smart, actual-time useful resource device that facilities navigation, item detection and face reputation for the visually impaired individuals

OBJECTIVE

- To create AI-powered smart glasses with Raspberry Pi and YOLOv8 for real-time object detection and recognition.
- To implement face recognition for identifying acknowledged humans and announcing their names via an audio signal.

- To ensure efficient and transportable audio feedback through low-volume alerts or earphones.
- To provide a cost-effective and wearable assistive device that improves navigation, object recognition, and social interaction for visually impaired individuals.

II. LITERATURE REVIEW

Ho et al. [1] advise an impediment detection technique that makes use of depth statistics to allow the visually impaired to keep away from boundaries after they pass in unusual surroundings. The system is composed of three elements: scene detection, obstacle detection, and a vocal declaration. This takes a look at and proposes a new technique to cast off the floor plane that overcomes the over-segmentation hassle. This system addresses the over-segmentation problem by doing away with the threshold and the initial seed function problem for the location growth technique using the connected component method (CCM).

Lin et al. [2] suggested a navigation framework the usage of cellphone applications that may be used with a picture recognition system. The stated framework works in one of the two feasible modes, online or offline, relying on network availability. Whilst the system is becoming on, the smartphone captures a photograph and sends it to the server for processing. To differentiate between character boundaries, the server uses a deep learning algorithm. The primary drawback of the gadget is high strength intake.

Lock et al. [3] investigated a multimodal person interface that makes use of sound and vibration alarms to transmit navigation statistics to goal users. The principal downside is that you need to run Augmented Reality core, which isn't supported on all smartphones.

Tanveer et al. [4] developed a walking aid for the visually impaired based on a special smartphone-enabled wearable device. Whenever the location of an obstacle is identified, the smartphone app plays a voice in Bengali/English language. GPS is explored to find the user's location, and the blind person's location is tracked using Google Maps. The overall error rate reported is approximately 5% for concrete and floor tiles. However, this method does not work under certain conditions. A striking case is a room with a raised floor.

Chang et al. [5] present a system for detecting air obstacles and road falls. In the event of a fall, emergency alert notifications are sent to family persons or defined guardians. The proposed system comprises: i) wearable smart glasses, ii) smart wand, iii) mobile app, and iv) cloud-based information management platform that sends relevant alerts. Experimental results claimed an average fall detection accuracy of up to 98.3%. However, the system is unable to identify direct aerial and ground imagery such as road signs and traffic cones and does not mention data about the power, cost, and weight of the proposed solution.

Islam et al. [6] proposed the pedestrian guidance system, which can recognize obstacles in three directions—left, front, and right—and potholes in the road surface using ultrasonic sensors in combination with convolutional neural networks (CNN). It consists of: i) ultrasonic sensor, ii) Raspberry Pi, iii) Raspberry Pi camera, iv) headphones, and v) external power supply. The system is mounted on the user's head and receives feedback via sound signals. According to the authors, the system has an accuracy of 98.73% for the front sensor with an error rate of 1.26% (obstacle distance 50 cm), while the image classification's accuracy, precision, and recall attained are 92.67%, 92.33%, and 93%, respectively. However, the system requirements for headphones create problems for the blind and visually impaired, as headphones can potentially block out safety-threatening ambient sounds.

Lin et al. [7] suggested a deep learning-based framework with an RGB-D camera, a semantic map, and an obstacle avoidance engine that learns from pilot input tasks. It consists of i) smartphone, ii) headphones, iii) RGB-D stereo camera, iv) wearable terminal with sunglasses, and v) external PC. The system presents a voice interface to the user, weighs no more than 150 g, and achieves an accuracy of 98.7% in the daytime, 97% in daylight, and 9% at night. The authors do not include any information about power requirements and cost. One of the system's weaknesses is its form factor, which affects sensitivity and fit to different lighting scenarios. Efficient and accurate scene text (EAST) stands for an efficient and accurate scene text detector [21]. This method is a simple and robust pipeline that enables text detection in natural scenes and provides high accuracy and efficiency. Experimental results show that this method gives better results than previous methods in terms of accuracy and efficiency.

Long et al. [8] present a fusion system for perception and obstacle avoidance. It consists of a millimeter-wave radar

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

and an RGB depth sensor and also features a stereo user interface. Experiments with this system have shown that the effective detection range is increased to 80m compared to using only the RGB-D sensor. However, the proposed solutions are cumbersome and costly. Also, the system is limited to object detection, it doesn't recognize it, and it can't transmit because it's still running.

ENVISION [9] uses a special approach to reliably and accurately detect static and dynamic obstacles in real-time video streams captured by smartphones with average hardware. The system can be further improved if the obstacle detection and classification module can help the target user better navigate the environment.

Badrloo et al. [10] proposed a new approach to assist blind people with indoor and outdoor navigation by marking their location and guiding them to their destination. The system uses radio frequency identification (RFID) technology, which covers a distance of about 0.5 m, and the test results show that the accuracy of the proposed work is in the range of 1-2 m. However, the method(s) used to estimate the accuracy of the solution is not clearly defined.

Meliones et al. [25] presented an obstacle detection process as a component of a mobile application that analyzes real-time data received from an external sonar. Its prime task is to discover obstacles in the user's path and transmit information about the detected distance, size, and potential movement through a voice interface and advise the user on how to avoid the obstacles.

III. METHODOLOGY

The proposed system utilizes a Raspberry Pi equipped with a Class 10 SD card and a USB camera for real-time face recognition. The YOLOv8 model is implemented for efficient and accurate detection and identification of individuals. Upon recognizing a person, the system uses a Text-to-Speech (TTS) module to audibly announce their name. The audio output remains low without amplification, making earphones necessary for clearer sound. The system operates in a sequential pipeline, capturing images, processing them with YOLOv8, mapping detected faces to stored names, and converting text to speech for name announcement. This setup is ideal for lightweight edge computing and real-time applications.

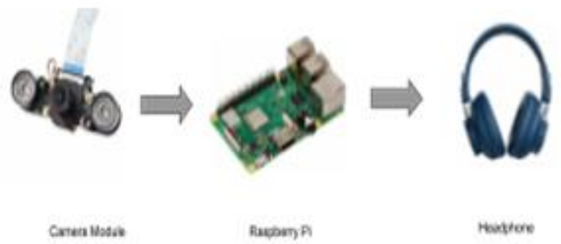

Fig.3.1: Block Diagram

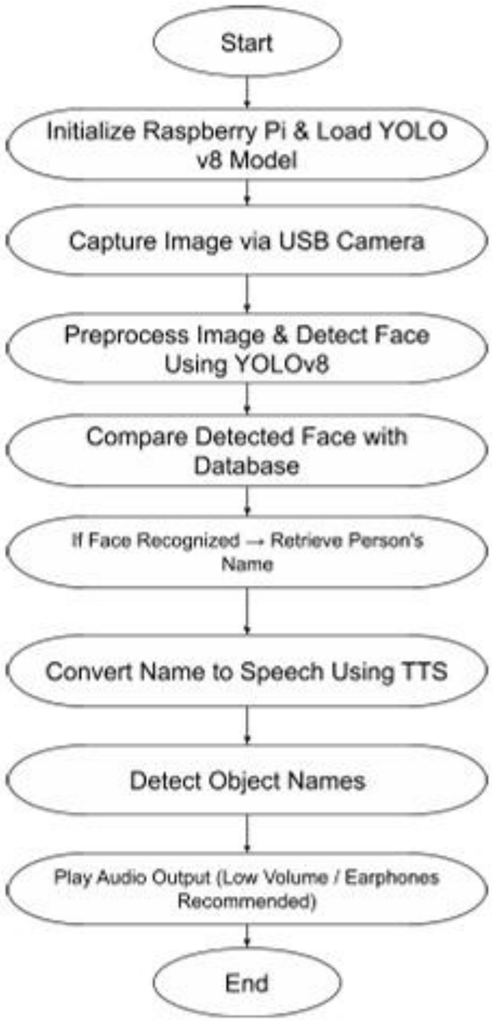

Fig. 3.2: Flow Chart

IV. WORKING

The proposed machine is designed to carry out real-time face recognition and object detection, utilizing a Raspberry Pi integrated with a USB camera, YOLOv8 model, and a Text-to-Speech (TTS) module. The camera continuously captures frames, which are processed by the YOLOv8 model to detect and identify human faces. If a recognized face matches a stored identity, the machine retrieves the corresponding name and converts it into speech using the TTS module. The audio output is performed via a low-amplitude speaker or headphones, ensuring that the identified name is audibly announced. This process is repeated continuously, making the device suitable for real-time recognition applications such as security, attendance systems, and smart assistants.

Further to face recognition, the system is also capable of object detection using the same YOLOv8 model. When the camera detects an object within its field of view, it classifies and recognizes the object based on a trained dataset. Once identified, the device generates an audio announcement of the detected object's name. The combination of face recognition and object detection with voice output makes this project a powerful AI-driven solution for enhanced accessibility, automation, and security monitoring.

4.1. System Requirement

❖ **Hardware Requirement**

1. RASPBERRY PI
2. IR-Cut Camera 5 Mp OV564
3. Headphone

❖ **Software Requirements**

1. Python

4.2. Experimental Setup & Result

Fig. 4.1: Experimental Setup

The implementation of real-time face recognition and item detection with voice output was effectively accomplished. The system successfully recognized human faces and detected objects while providing accurate voice-based announcements.

Fig 4.2: Shows the object Detection of the project

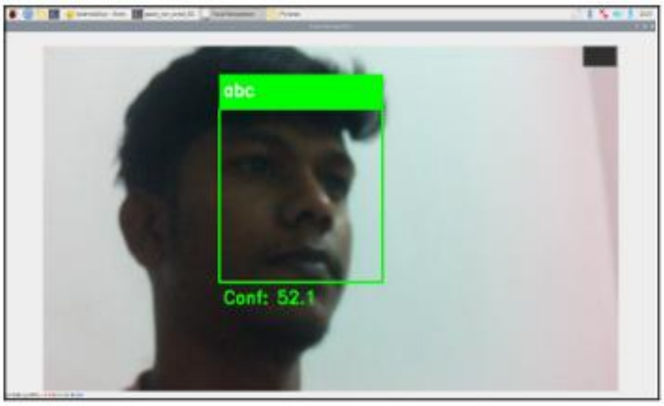

Fig.4.3: Shows the Face Recognition of the project

This AI-driven answer complements accessibility, automation, and safety monitoring, making it a feasible technique for programs such as clever assistants, surveillance, and computerized identification systems.

V. CONCLUSION

The actual-time face reputation and object detection machine using Raspberry Pi, YOLOv8, and a textual content-to-Speech (TTS) module was successfully carried out, demonstrating efficient overall performance in figuring out human faces and detecting items with voice-based totally announcements. The seamless integration of hardware and software program ensured real-time processing with minimum latency, making the gadget suitable for applications such as safety, attendance monitoring, and smart assistants. The task efficaciously showcased the capability of AI-driven answers in enhancing accessibility, automation, and safety monitoring.

FUTURE SCOPE

The real-time face recognition and object detection system using Raspberry Pi, YOLOv8, and a Text-to-Speech (TTS) module was successfully implemented, demonstrating efficient performance in identifying human faces and detecting objects with voice-based announcements. The seamless integration of hardware and software ensured real-time processing with minimal latency, making the system suitable for applications such as security, attendance monitoring, and smart assistants. The project effectively showcased the potential of AI-driven solutions in enhancing accessibility, automation, and security monitoring.

REFERENCES

[1] H. W. Ho, C. D. Wagter, B. D. W. Remes, and G. C. de Croon "Optical-flow based self-supervised learning of obstacle appearance applied to MAVlanding," Robotics and Autonomous Systems, vol. 100, pp. 78-94, 2018, doi: 10.1016/j.robot.2017.10.004.

[2] B. S. Lin, C.-C. Lee, and P.-Y. Chiang, "Simple smart phone-based guiding system for visually impaired people," Sensors, vol. 17, p. 1371, doi: 10.3390/s17061371.

[3] J. C. Lock, G. Cielniak, N. A. Bellotto, "Portable navigation system with an adaptive multimodal interface for the blind," In Proceedings of the 2017 AAAI Spring Symposium Series, Stanford, CA, USA, pp. 27-29, March 2017.

[4] M. S. R. Tanveer, M. M. A. Hashem, and M. K. Hossain, "Android assistant eyemate for blind and blind tracker," In Proceedings of the 2019 18th International Conference on Computer and Information Technology (ICCIT), Dec. 2015, pp. 266-271, doi: 10.1109/ICCITechn.2015.7488080.

[5] 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 Sensor Journal, vol. 20, pp. 10199- 10210, 2020, doi: 10.1109/ICCE.2019.8661943.

[6] M. Islam, M. S. Sadi, and T. Braunl, "Automated walking guide to enhance the mobility of visually impaired people," IEEE Transactions on Medical Robotics and Bionics, vol. 2, pp. 485-496, 2020, doi: 10.1109/TMRB.2020.3011501.

[7] Y. Lin, K. Wang, W. Yi, and S. Lian, "Deep learning based wearable assistive system for visually impaired people," In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019, pp. 2549-2557, doi: 10.1109/ICCVW.2019.00312.

[8] X. Zhou et al., "EAST: an efficient and accurate scene text detector," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5551-5560, doi: 10.1109/CVPR.2017.283.

[9] N. Long, K. Wang, R. Cheng, K. Yang, W. Hu, and J. Bai, "Assisting the visually impaired: multitarget warning through millimeter wave radar and RGB-depth sensors," Journal Electron Imaging, vol. 28, p. 013028, 2019, doi: 10.1117/1.JEI.28.1.013028.

[10] Shankar et al., "ENVISION: assisted navigation of visually impaired smartphone users," Procedia Computer Science, pp. 128- 135, 2018, doi: 10.1016/j.procs.2016.09.132.

[11] S. Badrloo, M. Varshosaz, S. Pirasteh, and J. Li, "A novel region-based expansion rate obstacle detection method for MAVs using a fisheye camera," International Journal of Applied Earth Observation and Geoinformation, vol. 108, 2022, doi: 10.1016/j.jag.2022.102739.

[12] A. Meliones, C. Filios, and J. Llorente, "Reliable ultrasonic obstacle recognition for outdoor blind navigation technologies," vol. 10, 10.3390/technologies10030054. no. 54, 2022, doi: