IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 2, December 2025

Impact Factor: 7.67

Blind Assistance System Using YOLOv8

Sinchana M[1], Moulya S U[2], Moksha K H[3], Geetha N[4], Prof. Chethan Kumar T[5]

[1-4]UG Student, Department of Information Science and Engineering

[5]Assistant Professor, Department of Information Science and Engineering

Kalpataru Institute of Technology, Tiptur

Abstract: *For real-time object detection and speech alarms, Research offers a sophisticated blind aid solution that smoothly combines the YOLOv8 algorithm with OpenCV's DNN (Deep Neural Network) module. The main objective is to improve the safety and freedom of people with visual impairments by offering quick object recognition and audio feedback. The YOLOv8 method, which is tuned for real-time inference, is used by the system to precisely recognize objects after using a webcam to record live video input. It further determines the distance and direction of detected objects, counts multiple instances, and provides real-time voice alerts when any object is too close to the user. Then, using audio description, it produces speech notifications that provide crucial details about the things it has spotted. Because of its exceptional versatility, the research can provide speech outputs in the user's preferred language, increasing its usability and accessibility. Its versatility is further demonstrated by its capacity to precisely handle a variety of object classes, making it a priceless tool for greatly enhancing the lives of those who are blind or visually impaired. Challenges in real-time object detection include occlusion, scale variations, and cluttered environments. Researchers must navigate the trade-offs between accuracy and speed. Real-time object detection is pivotal in computer vision, enabling intelligent systems across diverse applications. The continuous evolution of deep learning algorithms and hardware capabilities pushes the boundaries of this field, making it a dynamic research domain in artificial intelligence.*
Additionally, this project incorporates a focus-based detection feature that prioritizes objects directly in the user's path, ensuring more meaningful alerts. These enhancements collectively strengthen the system's effectiveness as a reliable real-time assistive tool.

Keywords: yolov8, Real-time object detection,Speech alerts

I. INTRODUCTION

Sight plays a fundamental role in human perception, enabling individuals to interpret their surroundings, make decisions, and navigate safely. For visually impaired individuals, the absence or limitation of vision creates significant challenges in daily mobility, environmental understanding, and independent living. Although several assistive tools such as white canes, braille systems, GPS- based devices, and ultrasonic sensors exist, many of these technologies offer only partial support, lack real-time environmental understanding, or provide insufficient contextual information. This gap highlights the growing need for intelligent, robust, and accessible solutions powered by modern computational techniques.

Recent advancements in artificial intelligence (AI), computer vision, and deep learning have opened new possibilities for developing assistive systems that provide more meaningful support to visually impaired users. YOLOv8 (You Only Look Once, version 8), one of the latest object detection models, demonstrates state-of-the- art performance in terms of speed, accuracy, and real- time inference, making it well-suited for wearable and portable applications. By leveraging its enhanced feature extraction capabilities, it becomes possible to design systems that can detect multiple objects simultaneously while maintaining high efficiency.

The "Blind Assistance System using YOLOv8" utilizes these technological strengths to create a real- time vision-aid framework capable of detecting objects, estimating their distance, identifying their direction, and delivering this information through speech output. Developed using Python and integrated with text-to- speech engines, the system aims to bridge the gap between visual perception and auditory understanding, thereby enhancing the user's spatial awareness. Unlike traditional assistive devices that merely detect obstacles, this system provides detailed contextual information such as object type, proximity, and relative position, enabling more informed decision-making.

Moreover, the addition of a focus-based detection feature allows the system to highlight objects that are most relevant to the user's immediate path, reducing unnecessary auditory load and improving clarity. Real- time processing ensures that users receive instantaneous alerts in dynamic environments such as streets or crowded public areas. The multilingual speech capability further enhances the system's adaptability, making it suitable for users from diverse linguistic backgrounds.

This introduction outlines the foundational motivation behind the project, emphasizing the importance of integrating modern AI capabilities into assistive technology. It also provides an overview of the existing drawbacks in conventional systems, the significance of real-time detection using YOLOv8, and the potential impact of this work on the quality of life for visually impaired individuals. The chapter concludes by presenting the aims, purpose, and scope of the proposed system, establishing the basis for the subsequent sections that discuss methodology, system design, implementation, experimental results, and future research directions.

II. LITERATURE REVIEW

Gupta et al. [1] introduced an advanced guide cane designed to support blind individuals in indoor and outdoor environments using obstacle detection and GPS navigation. Although effective in basic mobility assistance, the system relies heavily on ultrasonic sensors and lacks real-time object recognition or contextual awareness. This limitation emphasizes the need for more intelligent, vision-based solutions.

Prasad and Shanmuganantham [2] proposed a robot- assisted navigation system incorporating a Raspberry Pi and onboard camera for obstacle detection. Their approach shifts from traditional wearables to a rover-based model, offering hands-free guidance through voice commands and Bluetooth-controlled navigation. However, the system's portability and user convenience remain concerns, making it less practical for continuous personal use.

Building on sensor-based solutions, Prutha et al. [3] designed a smart cane integrating ultrasonic and water sensors, controlled by Arduino and Raspberry Pi modules. While effective in detecting obstacles and puddles, the system still depends on short-range sensors and cannot identify object types or provide directional context, which are crucial for advanced situational awareness.

Shandu et al. [4] addressed the challenges faced by visually impaired individuals by developing an AI-based pilot system to prevent slips, falls, and unawareness of surrounding objects. Although the system highlights the importance of AI, it does not incorporate real-time deep learning models capable of multi-object classification.

Venkatraman et al. [5] developed another smart cane capable of detecting moving obstacles and potholes, contributing to safer mobility. However, like previous cane- based systems, it does not integrate advanced computer vision algorithms, limiting its ability to interpret complex environments.

The work by Hegde et al. [6] explores a face-recognition system using Raspberry Pi for security and COVID-19 temperature detection. While not directly aimed at visually impaired assistance, it emphasizes the effectiveness of computer vision in embedded systems, reinforcing its potential application in mobility aids.

Satam et al. [7] proposed a smart stick solution that assists blind individuals in navigating daily environments. The design enhances user support but remains grounded in traditional sensor technologies, lacking contextual or semantic understanding of the user's surroundings.

Williams et al. [8] investigated the interaction between auditory context and visual perception, demonstrating how naturalistic sounds influence interpretation of visual objects. Their findings provide theoretical support for audio–visual integration in assistive systems, validating the use of speech-based feedback in real-time detection systems. This insight reinforces the importance of combining sound cues with visual data to enhance user understanding in assistive technologies.

Ali et al. [9] explored AIoT and Big Data applications in healthcare, highlighting the shift toward intelligent, connected devices. These technologies inform the development of more advanced assistive systems that can process, analyze, and respond to data in real time.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

III . PROPOSED SYSTEM

A. SYSTEM OVERVIEW

The proposed system is an real-time Blind Assistance System designed to enhance environmental awareness for visually impaired individuals through advanced computer vision and speech-based feedback. The system integrates the YOLOv8 object detection algorithm with OpenCV's Deep Neural Network (DNN) framework to identify and classify multiple objects in the user's surroundings with high accuracy and minimal latency. A live video feed captured through a camera is processed frame-by-frame, enabling continuous monitoring of the environment. Beyond basic object detection, the system incorporates additional modules for distance estimation, direction prediction, and focus- based prioritization. These components allow the system not only to detect objects but also to determine their spatial relevance—such as whether an object is directly ahead, to the left, or to the right of the user. Distance calculation ensures timely alerts when an object is too close, improving safety during navigation.

To provide seamless interaction without visual cues, the system employs a text-to-speech engine that converts detection results into real-time audio alerts. This includes information about object type, proximity, and position, allowing users to develop an accurate mental map of their environment. The speech feedback can also be configured to support multiple languages, enhancing accessibility for diverse users.Overall, the proposed system offers a robust and scalable solution by combining deep learning, spatial analysis, and audio communication. It addresses the limitations of traditional sensor-based assistive devices by providing rich contextual awareness and real-time decision support, ultimately empowering visually impaired individuals to navigate more independently and confidently

B SYSTEM ARCHITECTURE

Object detection is a fundamental task in computer vision with numerous applications spanning across various domains such as surveillance, autonomous vehicles, robotics, healthcare, and more. The utilization of advanced deep learning techniques has significantly enhanced the accuracy and efficiency of object detection systems. Among these techniques, YOLOv8 (You Only Look Once version 8) has emerged as a prominent algorithm due to its real-time processing capabilities and robust performance. Its ability to detect multiple objects simultaneously with high precision makes it suitable for dynamic and complex environments. Its lightweight architecture allows deployment on portable and embedded devices, enabling real-time applications such as assistive technologies.

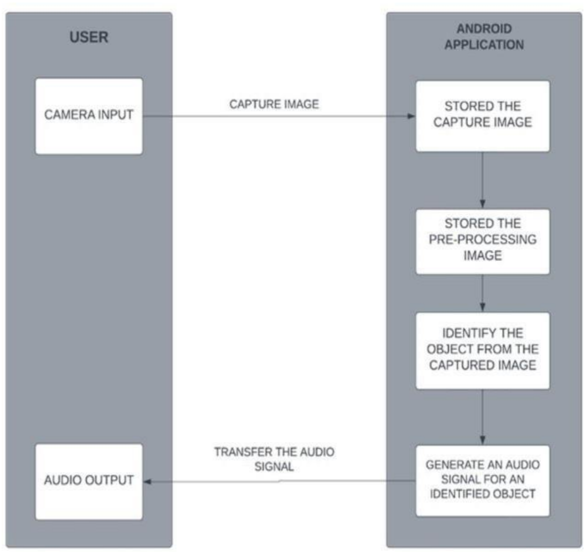

Fig 1 Block diagram of the system

Figure 1 illustrates the overall workflow of the proposed Blind Assistance System. The camera module captures real-time video input, which is then processed using the YOLOv8 object detection algorithm integrated with OpenCV. The system identifies objects within the frame and estimates their distance (close, medium, far) and direction (left, center,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

right) based on bounding box size and position. These processed outputs are then converted into audio feedback through a speech synthesis module, enabling the visually impaired user to understand their surroundings. The feedback includes the object's name, direction, and relative distance, allowing the user to make informed movement decisions. This mechanism provides real-time assistance for safe and independent navigation by combining computer vision, deep learning, and voice guidance in a single integrated system.

C. MODULES

The proposed Blind Assistance System consists of three main modules: Object Detection, Converting Detected Objects into Speech, and Depth Estimation. These modules work together to provide real-time assistance, helping visually impaired users navigate safely and understand their surroundings.

Object Detection

This module is the core of the system. It identifies objects in the user's environment using a pre-trained dataset, specifically the COCO dataset. The system captures live video through a webcam and detects objects in real time. YOLOv8, along with bounding boxes, is used to locate objects and classify them into specific categories. Once an object is detected, the system recognizes its name by comparing it with the pre-trained dataset. This ensures that users receive accurate information about everyday objects such as chairs, doors, or people.

Converting Detected Objects into Speech After detecting an object, the system converts its name and location into audio feedback. This is critical for visually impaired users, allowing them to "hear" what is around them. The system uses pyttsx3, a Python library, to generate clear and understandable voice notifications. This module provides real-time guidance, including object type and position, enabling users to navigate safely and confidently.

Depth Estimation

The depth estimation module calculates the distance between the user and detected objects. Once an object is identified, the system draws a bounding box around it and estimates how far it is from the user. If an object is too close, the system provides an alert, helping the user avoid collisions. By combining object detection with distance calculation, this module enhances the reliability of the system and improves overall safety and usability.

D. METHODOLOGY

The proposed Blind Assistance System is designed to assist visually impaired users in navigating their surroundings safely by detecting objects, estimating their distance and direction, and providing real-time audio feedback. The system begins with capturing live video through a camera, which is then preprocessed to ensure compatibility with the YOLOv8 object detection model. Preprocessing involves resizing, normalization, and formatting of frames, which enhances detection accuracy and reduces computational load, ensuring smooth real- time performance even in dynamic environments.After preprocessing, the YOLOv8 model detects and classifies objects within the captured frames. Leveraging deep convolutional neural networks, YOLOv8 predicts object locations and categories in a single pass, drawing bounding boxes around each detected object. Confidence scores are assigned to each prediction to maintain reliable detection. This information is passed to the spatial analysis module, which calculates both the distance and direction of each object relative to the user. Objects are categorized as near, medium, or far and positioned as left, center, or right, providing crucial spatial awareness for safe navigation.

Following detection and spatial analysis, the system converts the processed information into speech using the pyttsx3 Python library. The audio feedback includes details about the object type, its direction, and proximity, allowing users to form a mental map of their surroundings and make informed movement decisions. The system is designed to operate seamlessly in real time, continuously capturing frames, detecting objects, analyzing spatial data, and generating speech output. the proposed system offers an intelligent, reliable, and user-friendly assistive tool that significantly enhances the independence and safety of the visually impaired individual.

ISSN: 2581-9429
Impact Factor: 7.67

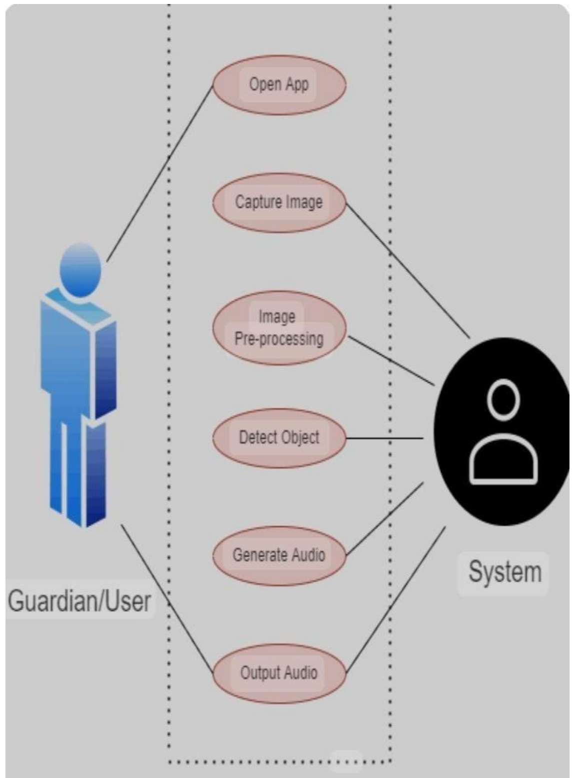

Fig 2 Block diagram of the system

This diagram represents the workflow of a system designed to assist a user or guardian in detecting objects. First, the user opens the application and captures an image using the system. The captured image undergoes pre-processing to enhance quality and prepare it for analysis. Next, the system detects objects present in the image using object detection algorithms. After identifying the objects, the system generates corresponding audio information. Finally, this audio is output to the user, providing real-time feedback about the detected objects. Overall, the diagram shows a step-by-step interaction between the user and the system to convert visual information into audio output for easier perception.

IV. RESULTS

The system was tested on various everyday objects to evaluate its performance and accuracy. It successfully detected objects of different sizes and shapes, such as chairs and bottles, in real time. The snapshots show clear identification with bounding boxes, demonstrating the reliability of the object detection module. Overall, the results confirm that the system can effectively assist visually impaired users by recognizing objects in practical environments. The proposed system was evaluated through a series of real-time object detection tests to verify its accuracy, responsiveness, and practical applicability.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

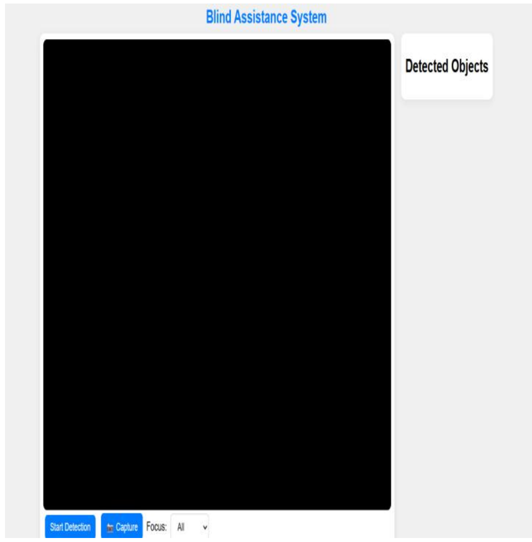

Fig 3 Front Page of the system

Figure 3 illustrates the front page of the application, which serves as the main user interface. The interface is designed for simplicity and real-time feedback, displaying detected objects along with relevant information such as object labels and confidence scores.

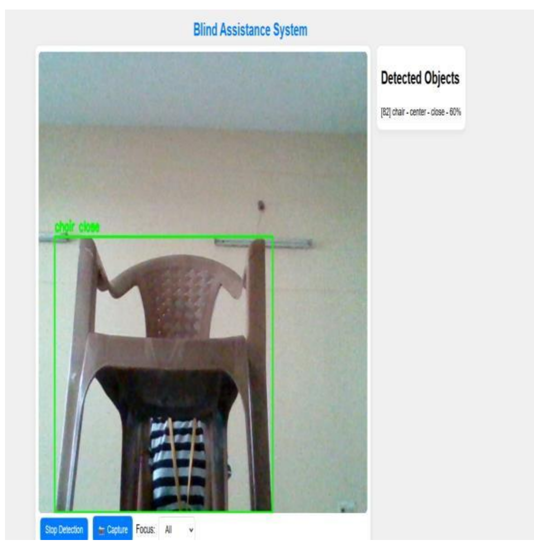

Fig 4 System detecting Chair

Figure 4 shows the system detecting a chair in an indoor environment. The detection is precise, with a clearly marked bounding box around the object, demonstrating the model's ability to accurately identify larger, stationary objects. This indicates that the system can effectively recognize common furniture items, which are essential for navigation assistance in indoor spaces.

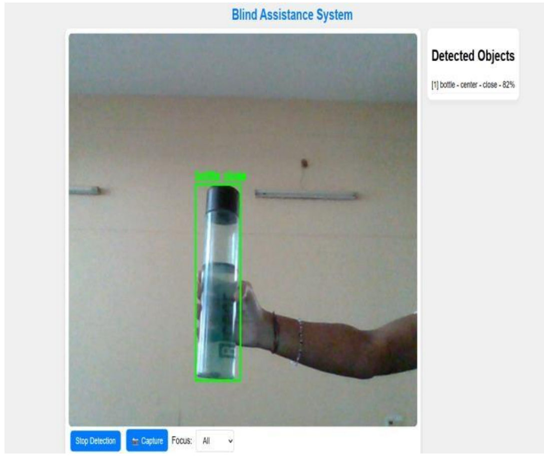

Fig 5 System detecting Bottle

Figure 5 presents the detection of a bottle. Despite being smaller and having a different shape compared to a chair, the system successfully identifies the object, showing versatility in detecting objects of varying sizes and geometries. The accuracy of detection in diverse scenarios suggests that the system can generalize well across different object categories and environmental conditions.

V. CONCLUSION

The "Blind Assistance System" is a useful and efficient tool for the blind and visually impaired that effectively combines sensor and computer vision technology. Users benefit from a greater awareness of their surroundings and are able to avoid impediments thanks to the system's ability to identify and distinguish things in the environment. The method ensures that users may make informed navigational decisions by providing crucial information about surrounding items through audio feedback. Although the system's performance is encouraging, there is still room for development. These include increasing accuracy in different lighting scenarios, lowering false positive rates even further, and including more sophisticated capabilities like distance sensing and obstacle tracking. Its broad adoption will also depend on adding more user-friendly interfaces and making sure the battery lasts a long time.

In the conclusion, this project shows how object detection and assistive technology can be used to enhance the quality of life for people with visual impairments by promoting more mobility and independence . Additionally, researchers are working on developing more robust and accurate algorithms that can better handle different object scales, orientations, and lighting conditions. Another area of focus is the integration of object detection with other technologies. Object detection plays an important role in scene understanding, which is popular in security, transportation, medical, and military use cases.

REFERENCES

[1] Choi D., and Kim M. (2018). Trends on Object Detection Techniques Based on Deep Learning, Electronics and Telecommunications Trends, 33(4): 23-32.

[2] Dai Jet al., (2016). R-FCN: Object Detection via Region-based Fully Convolutional Networks. Conf. NeuralInform. Process. Syst., Barcelona, Spain, Dec. 4-6, p. 379-387.

[3] Dalal N. and Triggs B., Histograms of oriented gradients for Human Detection (2015). IEEE Comput. Soc.Conf. Comput. Vision Pattern Recogn.,San Diego, CA, USA, June 20-25, p. 886- 893.

[4] Russakovsky O et al., (2015). ImageNet Large Scale Visual Recognition Challenge, Int..J.Comput. Vision, 115(3): 211-252.

[5] Rajeshvaree Ravindra Karmarkar (2021). Object Detection System for the Blind with Voice Guidance, International Journal of Engineering Applied Sciences and Technology,6(2): 67-70.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 2, December 2025

Impact Factor: 7.67

[6] M. Hiromoto, H. Sugano and R. Miyamoto, "Partially Parallel Architecture for AdaBoost- Based Detection With Haar-Like Features", IEEE Trans. Circuits and Systems for Video Technology, vol. 19, pp. 41-52, Jan 2009.

[7] Singh, B.B.V.L. Deepak, Tanjot Sethi and Meta Dev Prasad Murthy, "Real Time Object Detection and Tracking Using Color Feature and Motion", IEEE Int. Conf Communication and Signal Processing, 2015.

[8[Joseph Redmon, Santosh Divvala, Ross Girshick, "You Only Look Once: Unified, Real-Time Object Detection",The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

[9] YOLO Juan Du1,"Understanding of Object Detection Based on CNN Family",New Research, and Development Center of Hisense, Qingdao 266071, China.

[10] Matthew B. Blaschko Christoph H. Lampert, "Learning to Localize Objects with Structured Output Regression", Published in Computer Vision – ECCV 2008 pp 2-15.

[11] Wei Liu, Dragomir Anguelov, Dumitru Erhan, "SSD: Single Shot MultiBox Detector", Published in Computer Vision – ECCV 2016 pp 21-37.