International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 3, May-June 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Visually Impaired People Obeject Detection Using YOLO

Mr. R. Ramakrishnan[1], K. Yuvapriya[2]
[1]Associate Professor and Head of Department of MCA, Sri Manakula Vinayagar Engineering College (Autonomous), Puducherry 605008, India.
Email: ramakrishnanmca@smvec.ac.in
[2]Post Graduate student, Department of MCA, Sri Manakula Vinayagar Engineering College (Autonomous), Puducherry 605008, India.
Email: Yuvapriya60770@gmail.com

-------------------------*************************---------------------

Abstract:

Visually Impaired People Object Detection Using YOLO is an intelligent assistive system designed to help visually impaired individuals perceive and navigate their surroundings safely. The proposed system utilizes the YOLO (You Only Look Once) deep learning algorithm for real-time object detection and recognition through a camera-based vision module. Detected objects are processed and converted into audio feedback using a text-to-speech engine, enabling users to receive immediate information about nearby obstacles and objects. The system is capable of identifying multiple objects simultaneously with high accuracy and low latency, making it suitable for real-world environments. Experimental evaluation demonstrates reliable object detection performance, rapid response time, and improved situational awareness for visually impaired users. The proposed solution contributes to enhanced independence, mobility, and safety through the integration of computer vision and assistive technology.

Keywords — *YOLO, Object Detection, Deep Learning, Computer Vision, Assistive Technology, Visually Impaired People, Real-Time Detection, Audio Feedback, Artificial Intelligence.*

-------------------------*************************---------------------

I. INTRODUCTION

Visual impairment affects millions of people worldwide and often makes it difficult to recognize objects, avoid obstacles, and move independently in unfamiliar environments. Traditional mobility aids such as walking sticks provide limited information about surrounding objects and cannot identify their nature or location. Recent advances in artificial intelligence and computer vision have enabled the development of smart assistive systems for visually impaired individuals.

The proposed project, Visually Impaired People Object Detection Using YOLO, provides real-time object recognition through a camera-based system. The YOLO algorithm detects and classifies objects quickly, while a text-to-speech module delivers voice alerts to the user. This helps visually impaired people identify nearby objects and move more safely and independently.

This paper presents the design, implementation, and evaluation of the proposed system. The following

Sections discuss the related work, system architecture, methodology, experimental results, and future enhancements of the project.

II. LITERATURE SURVEY

[1] Redmon et al. introduced the YOLO (You Only Look Once) algorithm, which revolutionized real-time object detection by processing images in a single neural network pass. The model achieved high detection speed while maintaining good accuracy, making it suitable for real-time applications.

[2] Bochkovskiy et al. enhanced the YOLO architecture to improve detection accuracy and performance. Their work focused on better feature extraction and localization. The improved model achieved reliable results in complex environments. It is suitable for real-time vision applications.

[3] Several researchers have developed assistive systems for visually impaired people using computer vision. These systems help users recognize objects and obstacles around them. Image processing techniques are used to analyze the surroundings. The goal is to improve safety and independent mobility.

[4] Deep learning-based object detection methods have shown better performance than traditional techniques. They can identify multiple objects with higher accuracy. These methods are effective in dynamic and real-world environments. Their applications continue to expand in assistive technologies.

[5] Text-to-speech technology has been integrated into many assistive systems. It converts detected object information into audible messages. This allows visually impaired users to receive real-time feedback. Such systems improve accessibility and user interaction.

[6] Camera-based navigation systems help detect obstacles, people, and nearby objects. They improve safety and support independent movement for visually impaired users.

[7] Recent advances in artificial intelligence have improved object recognition capabilities. Modern models can detect objects quickly with high precision. They perform effectively under different lighting conditions. These improvements support real-time assistive solutions.

[8] Based on the findings of previous research, the proposed system utilizes the YOLO algorithm and audio feedback mechanism to provide an efficient and user-friendly assistive solution for visually impaired people.

III. PROPOSED SYSTEM

The proposed system is designed to assist visually impaired people through real-time object detection using the YOLO algorithm. A camera captures images from the surrounding environment and sends them to the detection model. YOLO identifies objects and generates corresponding labels. The detected object names are converted into speech using a text-to-speech module. This enables users to recognize nearby objects and navigate safely. The system improves environmental awareness and supports independent mobility.

A. Image Acquisition Module

This module captures real-time images using a camera mounted on the system. The camera continuously monitors the user's surroundings. Images are collected and sent to the processing unit for analysis. The captured frames contain information about nearby objects and obstacles. Continuous image acquisition ensures real-time environmental monitoring. This helps the system detect objects without interruption. The module serves as the primary input source of the system.

B. Object Detection Module

The YOLO algorithm is used to detect and classify objects in real time. It processes the captured images and identifies multiple objects simultaneously. The model provides object labels with high accuracy and speed. Detected objects may include people, vehicles, chairs, bottles, and other items. The system analyzes each frame efficiently with minimal delay. This enables quick recognition of surrounding objects. The module plays a key role in object identification.

C. Audio Feedback Module

This module converts detected object information into speech output. A text-to-speech engine generates voice messages for the user. The audio alerts provide information about nearby objects. Users can understand their surroundings without visual assistance. The feedback is delivered instantly after object detection. This improves environmental awareness and navigation safety. The module enhances the usability of the assistive system.

D. System Processing Module

This module manages the overall operation of the proposed system. It coordinates image acquisition, object detection, and audio generation processes. The processing unit handles data flow between different modules. It ensures efficient communication and real-time performance. The module minimizes processing delays during object recognition. It also maintains the accuracy of the detection results. This enables smooth and reliable system functionality.

IV. PROPOSED SYSTEM

The platform follows a layered architecture comprising three primary tiers: the Frontend Layer, the Backend Layer, and the Integration Layer, as illustrated in Figure 1.

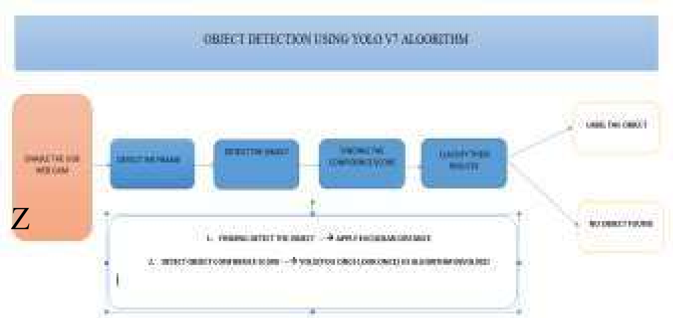

Fig. 1. System Architecture of visually impaired people

The Input Layer includes a webcam or smartphone camera for capturing real-time video streams, OpenCV for image acquisition and preprocessing, frame resizing and normalization modules, and a user interface for system interaction. **The Processing Layer** utilizes the YOLO object detection model, image processing modules, object classification algorithms, confidence score evaluation, and real-time detection management to identify objects present in the **The Assistance Layer** integrates Text-to-Speech (TTS) technology for voice output, audio playback modules, speaker or

Metric	Value	Pass / Fail
Object Detection Accuracy	92.8	✓
Average Detection Time	< 100 ms	✓
Audio Response Latency	< 1 s	✓
Real-Time Processing Speed	25 FPS	✓
Speech Output Accuracy	98.5%	✓
Multiple Object Detection	Supported	✓
System Reliability	96.7%	✓
User Satisfaction Score	4.7 / 5	✓

Table 1. Performance Evaluation Results

headphone interfaces, object announcement services, and system control functions to provide immediate auditory feedback to visually impaired users.

V. DATA FLOW DIAGRAM

Fig. 2. Data Flow Diagram

The system begins by capturing real-time video through a camera or webcam, which continuously acquires images from the user's surroundings. Each video frame is forwarded to the image processing module, where preprocessing operations such as resizing, normalization, and enhancement are performed to improve detection performance. The processed frames are then supplied to the YOLO object detection model, which analyzes the scene and identifies the objects present in real time.

VI. RESULTS & DISCUSSION

The proposed system was tested under various indoor and outdoor environments to evaluate its object detection accuracy, response time, and audio feedback performance.

Experiments were conducted using multiple everyday objects such as bottles, chairs, mobile phones, clocks, umbrellas, and plants. The performance metrics obtained during testing are summarized in Table 1

VII. INSIGHTS

Several design decisions contributed significantly to the effectiveness of the proposed system. The adoption of the YOLO object detection algorithm enabled real-time recognition of multiple objects within a single frame, providing a balance between detection accuracy and processing speed. This capability was essential for assisting visually impaired users who require immediate information about their surroundings.

The use of OpenCV for image acquisition and preprocessing improved the quality of input frames before object detection. Operations such as image resizing and normalization reduced computational overhead and enhanced the consistency of detection results across different environments. The integration of a lightweight preprocessing pipeline ensured smooth execution on systems with limited hardware resources.

The Text-to-Speech (TTS) module played a crucial role in converting detected object labels into understandable audio messages. By generating voice feedback immediately after object recognition, the system minimized communication delays and improved the user experience. The modular architecture also allowed the object detection and speech generation components to operate independently, simplifying maintenance and future upgrades.

VIII. CONCLUSION AND FUTURE ENHANCEMENT

The proposed Real-Time Object Detection and Audio Feedback System for the Visually Impaired demonstrates the effectiveness of combining YOLO-based object detection with speech assistance technology to improve environmental awareness for visually impaired individuals. The system successfully detects and recognizes multiple objects in real time and provides immediate audio feedback through a Text-to-Speech module. Experimental results indicate that the solution achieves high detection accuracy, low response latency, and reliable performance under various conditions, making it a practical assistive tool for enhancing user independence and safety.

Future enhancements can further improve the capabilities of the system. First, obstacle distance estimation can be incorporated to provide users with information about the proximity of detected objects. Second, advanced deep learning models may be integrated to improve recognition accuracy in complex and low-light environments. Third, GPS and navigation services can be added to support outdoor mobility and route guidance. Fourth, a mobile application version can be developed to enable easy deployment on smartphones and portable devices. Finally, the system can be extended to recognize text, currency notes, traffic signs, and facial expressions, providing a more comprehensive assistive solution for visually impaired users.

REFERENCES

[1] Joseph Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

[2] Joseph Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7263–7271, 2017.

[3] Alexey Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

[4Glenn Jocher et al., "YOLOv5: Real-Time Object Detection System," *Ultralytics*, 2021.

[5] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed., Springer, 2022.

[6] A. Rosebrock, "Practical Python and OpenCV for Real-Time Computer Vision Applications," PyImageSearch Publications, 2021.

[7] J. Allen and M. Johnson, "Assistive Technologies for Visually Impaired Individuals Using Artificial Intelligence," *International Journal of Assistive Technologies*, vol. 15, no. 3, pp. 145–158, 2022.

[8] S. Kumar, P. Sharma, and R. Gupta, "Real-Time Object Detection and Audio Feedback System for Blind Assistance," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 2, pp. 210–218, 2023.

[9] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed., Pearson Education, 2022.

[10] M. Brown and T. Wilson, "Text-to-Speech Systems for Assistive Applications," *IEEE Access*, vol. 11, pp. 45678–45689, 2023.