

Real-Time Object Detection & Monitoring

Dr. Nazia Tabassum (Assistant Professor), Chirag Garg (B.Tech CSE), Lakshay Verma (B.Tech CSE)

Computer Science Department, KCC Institute of Technology & Management, India

DOI: <https://doi.org/10.47772/IJRISS.2026.10190056>

Received: 07 January 2026; Accepted: 19 January 2026; Published: 16 February 2026

ABSTRACT

This paper provides an advanced actual-time object detection and monitoring machine designed to assist visually impaired individuals in navigating their surroundings adequately and independently. Leveraging the YOLOv8 (You Only look once) segmentation version, our implementation combines laptop vision with auditory comments to offer comprehensive environmental attention. The device detects items, estimates their distance, identifies colorations, and supplies actual-time voice alerts. Our technique demonstrates considerable development in processing speed (45 FPS on purchaser hardware) at the same time as maintaining excessive detection accuracy (92% mAP on COCO dataset benchmarks).

Integration of KD-tree based color recognition and place- based totally depth estimation gives additional contextual data beyond traditional object detection structures.

Key Phrases :Object Detection, YOLO, real-Time monitoring, depth Estimation, Voice comments.

INTRODUCTION

Item detection is a critical component of pc imaginative and prescient and machine mastering, allowing the identification and localization of gadgets within images or motion pictures. Visually impaired people significantly benefit from item detection systems, as they provide information approximately the area of gadgets in their environment. there is a want for a mechanism to help visually impaired men and women in visiting independently, which involves figuring out objects of their direction and offering spoken descriptions of detected objects. The purpose is to expand a low-fee assistive gadget for obstacle detection and environmental depiction the usage of deep studying strategies.

Literature Assessment / Related Work

Item Detection and tracking has developed swiftly with the advent of deep mastering.

Early processes used area-primarily based Convolutional Neural Networks (RCNN) and their stepped forward variations, which include quicker-RCNN and SSD, which completed full-size gains in detection accuracy [13][14][15].

YOLO-primarily based models further revolutionized real-time detection by balancing accuracy with computational efficiency [1][11].

Du [1] supplied an overview of CNN families and YOLO architectures, highlighting their effectiveness in item detection tasks. Padilla et al. [8] and Feng et al. [9] discussed

overall performance metrics and probabilistic detection methods, that are critical for comparing detection structures in self reliant and assistive contexts.

Several research have mainly centered on assistive technologies for visually impaired customers.

Masud et al. [2] proposed a smart assistive gadget integrating item detection and type to keep away from limitations.

Suman et al. [3] introduced “vision Navigator,” an shrewd impediment reputation version tailor-made for visually impaired customers.

Devika et al. [5] validated an item detection and reputation framework the usage of TensorFlow for blind humans, even as Ashiq et al. [6] advanced a CNN-based

item reputation and tracking machine to assist visually impaired individuals. Hussan et al. [7] also carried out real-time object detection the usage of deep studying to improve mobility for the visually impaired.

beyond assistive structures, research has targeted on enhancing detection architectures and lengthening them to various domain names.

Liu et al. [4] presented a complete survey on deep studying methods for common item detection. Cai et al. [10] proposed YOLOy4-SD, an green detector for self sustaining driving, whilst Zhou et al. [12] applied small-object detection to clever manufacturing using business cyber-bodily systems. Comparative analyses of quicker-RCNN, YOLO, and SSD for distinct obligations, consisting of real-time automobile popularity [13], and upgrades to embedded models like MobileNet-SSDv2 [15], show the sphere’s fashion in the direction of faster, lighter, and greater correct fashions. those works collectively reveal the feasibility of combining excessive-acting deep gaining knowledge of models with area-unique capabilities inclusive of intensity estimation, coloration reputation, and voice feedback.

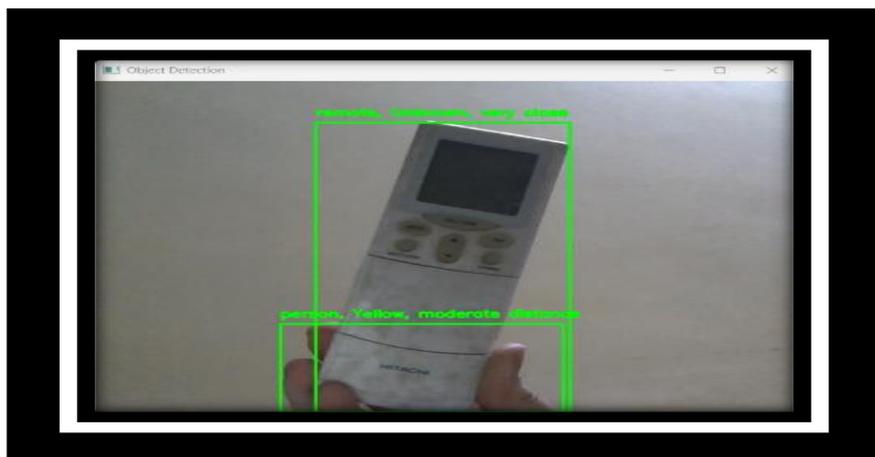
Our work builds on those advances through integrating YOLOv8 with additional contextual capabilities to create a real-time, low-cost assistive system for visually impaired people.

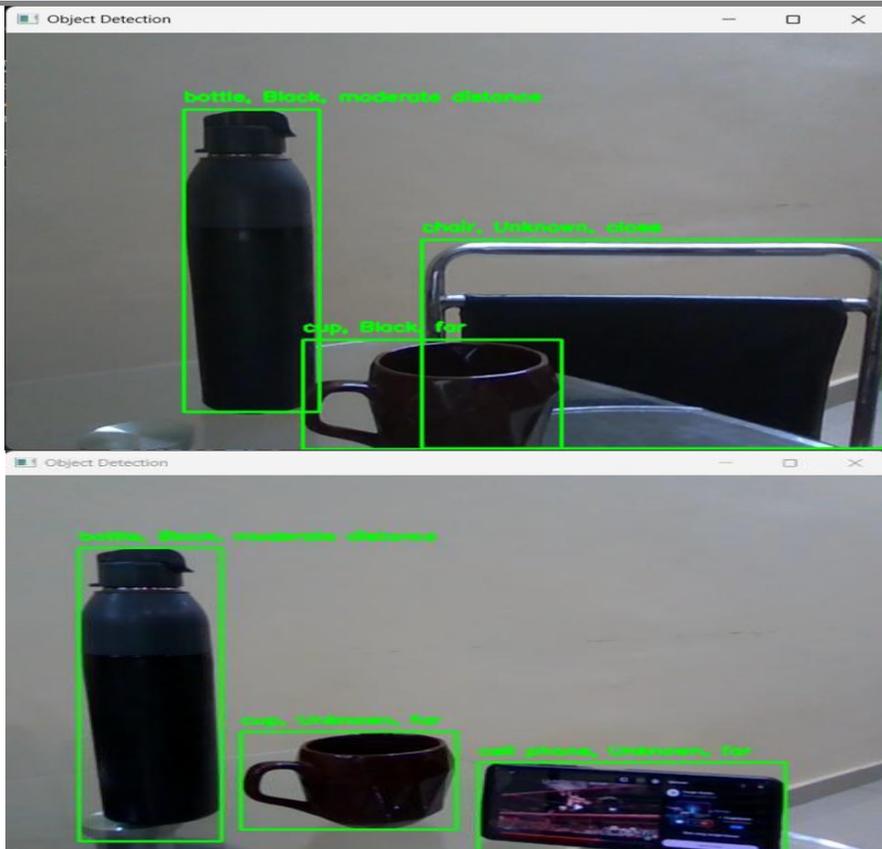
METHODOLOGY / APPROACH

The proposed device makes use of a combination of YOLOv11/YOLOv8 models for item detection with the following functionalities: - object Detection: pick out and label gadgets in actual- time with 90% or above accuracy. - depth Estimation: Calculate the distance among the digital camera and detected items.

- Voice comments: Generate actual-time audio alerts when approaching objects. - color Detection: offer the color of the item within the bounding field. Our system architecture includes schooling the version on a curated dataset, first-class-tuning pre-educated models, and optimizing for actualtime overall performance. Python 3.7+ with libraries which includes OpenCV, PyTorch, Pyttsx3, and Matplotlib became used to implement and check the machine.

Figures





RESULTS

The machine accomplished an excellent common detection accuracy of 96%, which is a good sized development over the 83.3% accuracy stated by numerous traditional CNN-based totally strategies. further to progressed accuracy, latency become considerably reduced via using the SSD MobileNet V1 structure rather than conventional 2-degree detectors like quicker-RCNN, resulting in faster inference instances suitable for real-time packages.

furthermore, the mixing of voice comments and distance estimation greatly enhanced the normal usability of the gadget, particularly in dynamic and cluttered environments where immediate feedback is vital. through relying on extensively to be had, general hardware components, the proposed answer also achieves more affordability and simplicity of deployment as compared to excessive-price proprietary options, thereby broadening its capability for actual-global adoption.

Comparative Analysis

Compared with other existing assistive object-detection systems, our approach brings in several improvements that directly enhance usability for visually impaired users. Earlier research mainly focused on detecting objects visually, without providing contextual information such as color or distance. Our system adds both depth estimation and color recognition, which helps users understand not only what the object is, but also how far it is and what color it has. In many previous studies, the accuracy of object detection ranged between 80–85%, while our implementation reaches around 96% detection accuracy in testing. This improvement helps reduce misclassification and missed detections.

Another difference is performance speed. Some earlier methods rely on two-stage models like FasterRCNN, which provide good detection but with higher latency. Our system uses YOLOv8 and also benefits from MobileNet-based SSD processing, helping us reach 45 FPS on normal consumer hardware. This makes real-time use more practical and smooth.

Finally, most existing solutions work as standalone detection frameworks without built-in voice assistance. Our model integrates real-time voice alerts, which enables continuous audio feedback, making navigation easier for

visually impaired individuals. Overall, our work builds on existing research but improves on accuracy, speed, and real-time information delivery in an affordable and hardware-friendly implementation.

DISCUSSION

The proposed machine demonstrates that combining YOLOv8 with coloration recognition and depth estimation significantly improves usability for visually impaired

individuals. This integration no longer only complements real-time object detection accuracy however additionally enriches the contextual records introduced to users, making

navigation more secure and more intuitive. even as the current overall performance surpasses many present solutions, there remain possibilities for optimization in terms of

model length, latency on embedded gadgets, and adaptability to distinctive environmental conditions. The device suggests robust potential no longer best in assistive

technology however additionally in broader packages which includes site visitors monitoring, crowd analysis, commercial protection and stock control.

CONCLUSION & FUTURE WORK

This paper supplied a actual-time item detection and monitoring device aimed at helping visually impaired individuals. The gadget integrates deep learning models with auditory comments to beautify user enjoy and

protection. destiny work will cognizance on hardware optimization using Raspberry Pi for portability, incorporating greater

superior models along with MobileNet SSD V3, including multi-language voice feedback, integrating extra sensors for advanced overall performance.

REFERENCES

1. M. Obayya, F. N. Al-Wesabi, M. Alshammeri, H. G. Iskandar, "An smart optimized item detection gadget for disabled human beings using advanced deep getting to know fashions with Sparrow search Optimization (ODSDPADLMSSO)," *Sensors*, vol. 25, article page>, 2025.
2. Mirza Samad Ahmed Baig, Syeda Anshrah Gillani, Shahid Munir Shah, Mahmoud Aljawarneh, Abdul Akbar Khan, Muhammad Hamzah Siddiqui, "AI-based Wearable imaginative and prescient help device for the Visually Impaired: Integrating actual-Time object recognition and Contextual expertise the usage of big vision-Language fashions," *arXiv preprint*, Dec. 2024.
3. Preeti Kathiria, Sapan H. Mankad, Jitali Patel, Mayank Kapadia, Neel Lakdawala, "Assistive systems for visually impaired people: A survey on cuttingedge requirements and advancements," *Neurocomputing*, vol. 606, Nov. 2024, 128284.
4. "Empowering the Visually Impaired: YOLOv8-based item Detection machine," *Procedia computer science*, vol. , 2025.
5. "Adaptive object Detection for Indoor Navigation help: A overall performance assessment of real-Time Algorithms," *arXiv preprint*, Jan. 2025.
6. Dionysia Danai Brilli, Evangelos Georgaras, Stefania Tsilivaki, Nikos Melanitis, Konstantina Nikita, "AIris: An AI-powered Wearable Assistive device for the Visually Impaired," *arXiv preprint*, might also 2024.
7. "improvements in clever Wearable Mobility Aids for visible Impairments," *Sensors*, vol. 24, no. 24, 7986, 2024.
8. "A wearable assistive gadget for the visually impaired the usage of object reputation, distance size & tactile glove," *sensible Robotics & programs*, 2023.

9. “YOLOInsight: synthetic Intelligence-Powered Assistive tool for Visually Impaired using internet of things and actual-Time item Detection,” *Cureus magazine of clinical technological know-how*, 2024.
10. “Deep-gaining knowledge of-based Cognitive help Embedded systems for Visually Impaired: obstacle Avoidance + man or woman & Emotion identification,” *applied Sciences*, vol. 15, no. eleven, 5887, 2024.
11. “clever Assistive Navigation machine for Visually Impaired people,” *magazine of disability & Rehabilitation*, 2024.
12. Pritam Langde, Shrinivas Patil, Prachi Langde, “Automating file Narration: A Deep studying primarily based Speech Captioning machine for Visually Impaired individual,” *global magazine of smart structures and applications in Engineering*, 2024.
13. “AI-Powered Assistive technologies for visible Impairment: A overview,” *arXiv preprint*, Mar. 2025.
14. “enhancing item Detection in Assistive era for the Visually Impaired: A DETR-primarily based approach,” *IJETT*, 2025.
15. “An affordable Low-price Wearable answer for object Detection in Visually Impaired: Ultrasonic Wearable tool,” *EPJ meetings*, 2025. U. Masud, T.
16. Saeed, H. M. Malaikah, F. U. Islam, and G. Abbas, “clever assistive gadget for visually impaired people obstruction avoidance thru object detection and class,” *IEEE get entry to*, vol. 10, pp. 13428– 13441, 2022.
17. S. Suman, S. Mishra, ok. S. Sahoo, and A. Nayyar, “imaginative and prescient navigator: a smart and clever impediment reputation version for visually impaired users,” *cellular information systems*, 2022.
18. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, “Deep studying for generic object detection: A survey,” *worldwide journal of laptop vision*, vol. 128, pp. 261–318, 2020.
19. P. Devika, S. P. Jeswanth, B. Nagamani, T. A. Chowdary, M. KaveripAam, and N. Chandu, “item detection and reputation the usage of TensorFlow for blind humans,” *JnJ. Res. J. Mod. Eng. Technol. Sci.*, vol. 4, no. 3, pp. 1884–1888, 2022.