International Journal of Engineering and Applied Computing
Vol. 1, pp. 39-49, August 2026

RESEARCH ARTICLE

Deep Learning-Based Assistive Navigation System for Visually Impaired Pedestrians Using YOLOv5

Deepika Yadav[1] . Hemant Yadav[2] . Pooja Yadav[1]

[1]Department of Computer Science and Information Technology, Mahatma Jyotiba Phule Rohilkhand University, Bareilly, Uttar Pradesh, India
E-mail: deepika.csit@mjprudor.ac.in, pooja.yadav@mjpru.ac.in
[2]Department of Computer Science and Engineering, Future University, Bareilly, Uttar Pradesh, India
E-mail: drhemantyadav22@gmail.com
Corresponding author: pooja.yadav@mjpru.ac.in

Received: 08/03/2026 | **Revised:** 23/07/2026 | **Accepted:** 25/08/2026

Abstract
In this study, an assistive navigation system is proposed for visually impaired pedestrians by combining object detection system You Only Look Once version 5 (YOLOv5), camera, Light Detection and Ranging (LiDAR) sensor, and voice feedback system based on Google Text-to-Speech (gTTS). The proposed system detects the objects around it, estimates the relative position of each object and delivers immediate audio guidance to facilitate safe navigation in urban areas. Using the annotated object detection data, the model was trained and evaluated and achieved a precision of 94.82%, recall of 69.82%, mean Average Precision at 0.5 of 91.57%, and mean Average Precision at 0.5–0.95 of 91.03%. The proposed framework integrates real-time visual perception with distance sensing and speech-based guidance, offering a continuous sense of the environment and enhancing mobility, safety, and user independence, unlike conventional assistive devices.

Keywords
Object Detection, Assistive Technology, Blind Navigation, YOLO Algorithm, Voice Integration, Google Text-to-Speech, Pedestrian Safety, Deep Learning

1. INTRODUCTION
Approximately 2.2 billion people worldwide experience vision impairment, of whom nearly one billion have preventable or treatable vision loss (World Health Organization, 2019). Visually impaired pedestrians face significant challenges while navigating complex urban environments, particularly when identifying obstacles and unfamiliar destinations (Dongre, 2020). Several assistive navigation systems have been developed to improve the mobility and independence of visually impaired individuals. However, these systems are frequently confined to specific pathways or nearby destinations (Chen et al., 2023). This study proposes a camera-based object detection framework integrated with a LiDAR sensor to detect surrounding obstacles and provide accurate navigation assistance for visually impaired pedestrians (Patel et al., 2024). Recent advances in deep learning (DL), particularly the YOLOv5 object detector, have enabled accurate real-time object detection suitable for assistive navigation systems (Srikanteswara et al., 2021). This technology endows assistive devices with the capability to translate visual data into speech, helping the visually impaired to navigate complex urban environments safely and independently.

Existing assistive navigation systems primarily rely on either camera-based object detection or ultrasonic sensing, often lacking accurate distance estimation and efficient real time feedback. Recent studies have explored multimodal sensing and advanced object detectors such as YOLOv8 and YOLOv11; however, these approaches typically require high computational resources or expensive hardware. Motivated by these limitations, this work proposes a lightweight assistive navigation framework integrating YOLOv5-based object detection, LiDAR-based distance estimation, and Google Text-to-Speech feedback to provide real-time obstacle awareness on resource-constrained edge devices. The main contributions of this paper are as follows:

1. Developed a real-time assistive navigation framework designed to support visually impaired pedestrians in safely navigating their surroundings.
2. Integrated object detection, obstacle distance estimation, and voice-based guidance into a unified system for comprehensive environmental awareness.
3. Evaluated the proposed framework using a custom annotated dataset comprising 1,250 images.
4. Achieved high object detection accuracy while maintaining real-time processing performance, demonstrating the system's suitability for practical applications.
5. Enhanced pedestrian safety and situational awareness by providing timely and intuitive audio-based navigation assistance.

The remainder of this paper is organized into six sections. Section 2 introduces the background and key concepts underlying the proposed system. Section 3 details the methodology, including the system architecture and implementation procedures. Section 4 presents the experimental results and evaluates the performance of the proposed approach. Section 5 discusses the obtained results, highlighting the effectiveness, limitations, and practical implications of the system. Finally, Section 6 summarizes the main findings and identifies possible directions for future research.

2. BACKGROUND

Any assistive device intended to assist the visually impaired in navigating urban environments must incorporate object detection and speech integration. Current advancements in the field are reviewed in the sections that follow:

2.1 Assistive Technologies for Blind Pedestrians

Conventional navigation devices like white canes and guide dogs offer only very limited environmental awareness, and don't detect moving or far away objects (Saravanan et al., 2022). Advances in assistive technology have been made, and computer vision and DL techniques have been combined to enhance obstacle detection and navigation, particularly in recent years (Chen et al., 2023). Camera based systems are able to identify pedestrian, vehicle and traffic signs in real-time, but are prone to perform poorly when lighting conditions are not optimal and cannot accurately estimate the distance of the objects (Tavakoli, 2024).

2.2 DL-Based Object Detection

DL has significantly improved object detection by enabling automatic feature extraction through convolutional neural networks. Compared with traditional handcrafted feature-based methods, DL models provide higher detection accuracy and faster inference. Two-stage detectors, such as Faster R-CNN, achieve high detection accuracy but require greater computational resources (Alrowais et al., 2026). In contrast, single-stage detectors, including the YOLO family, perform localization and classification simultaneously, making them suitable for real-time assistive navigation.

The most notable benefit of DL is the ability to automatically extract features from images using convolutional neural networks, which significantly enhances object detection. In comparison to the handcrafted feature-based conventional methods, DL models achieve better detection accuracy and faster inferences. High detection accuracy can be done by using two-stage detectors like Faster R-CNN, but with more computational resources (Alrowais et al., 2026). Compared to this, single-stage detectors, like the YOLO family, can be used for real time assistive navigation as they are able to localize and classify at once. The evolution of object detection has come a long way over the years and is typically divided into two different phases. The first stage is for traditional object detection techniques using handcrafted features and the second stage is for deep learning-based techniques that have taken over the field since 2014. These methods are shown to have developed through the years in Fig. 1.

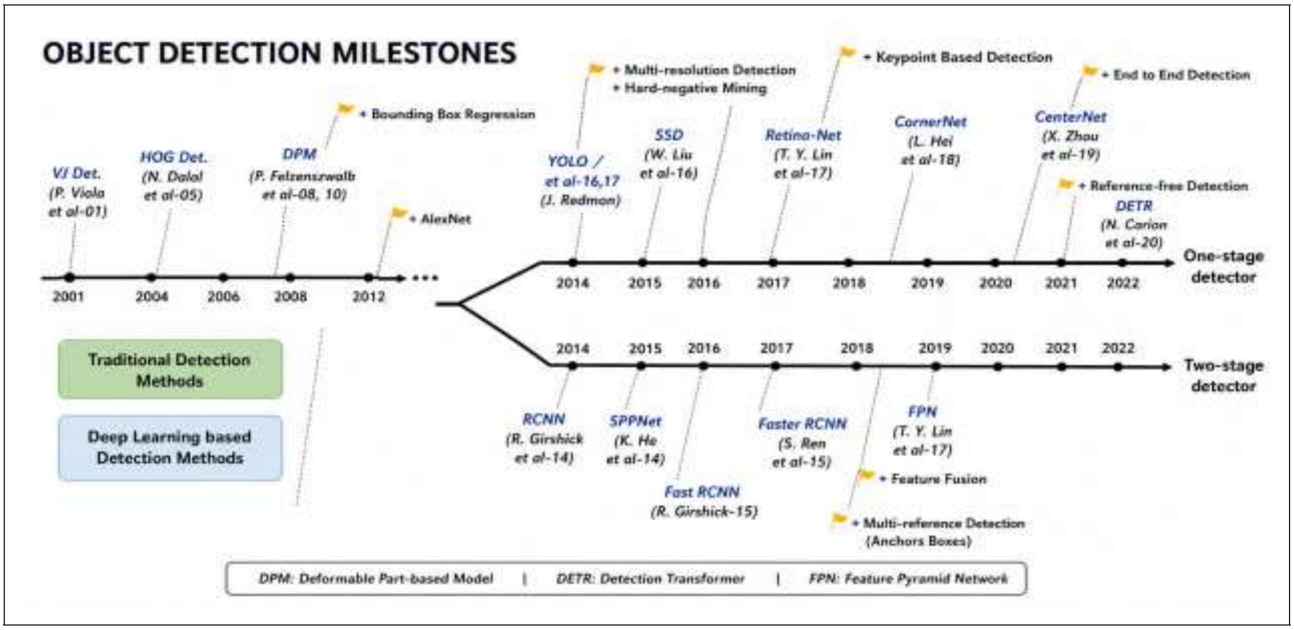

Fig. 1: Evolution of object detection algorithms from traditional methods to modern DL-based detectors (Zou et al., 2023)

2.3 Evolution of YOLO Architectures

From YOLOv1 to YOLOv11, the YOLO family has grown in leaps and bounds, enhancing detection accuracy and efficiency. YOLOv5 adopted the modular design, and advanced anchor augmentation techniques include automatic anchor optimization and Mosaic augmentation. YOLOv8 and YOLOv11 are some recent versions that further enhance feature extraction and detection of small objects. However, YOLOv5 is still popular for embedded and real-time applications due to its good trade-off of the accuracy and inference speed.

2.4 Voice Feedback for Pedestrian Navigation

Providing timely and clear audio feedback is crucial for pedestrian safety. Most of the time, assistive devices are based on the use of Text-To-Speech (TTS) engines or pre-recorded audio files in order to provide a voice guidance (Tavakoli, 2024). One example is the *LiCa* Blind Assistant that uses gTTS to convert text to speech (Saravanan et al., 2022). The researchers also suggest adding haptic feedback and providing multi-language options to make it more accessible and user-friendly (Lee et al., 2022). Recent research has shown that using multimodal feedback with both speech and haptic feedback can enhance confidence and navigation while enhancing user performance. However, speech feedback is still the most viable option given the need for little extra equipment and because it offers instant context around obstacles nearby. Table 1 presents a summary of recent studies on object detection and assistive navigation systems.

Table 1: Summary of recent object detection and assistive navigation studies

Reference	Paper Title	Application	Method	Performance	Major Contribution
(Sheela et al., 2024)	Traffic Sign Categorization Using YOLO Algorithm	Autonomous Vehicles	YOLOv8	Real-time detection	Validated high-speed traffic sign detection for autonomous driving
(Mathivanan et al., 2024)	Computer Vision Empowered Assistive Technology for Blind People	Assistive Navigation	YOLOv9	Audio + Haptic Feedback	Low-cost navigation assistance for visually impaired users
(Gao, 2024)	YOLO-Based Object Detection for Industrial Cleaning Robots	Industrial Robotics	YOLOv4 / YOLOv8	Real-time inference	Network pruning and quantization improve efficiency

(Gui et al., 2024)	Remote Sensing Image Object Detection on Edge Platforms	Remote Sensing	YOLOv3	High accuracy	Efficient edge-based remote sensing image detection
(Al Amin et al., 2024)	FPGA-Based Real-Time Object Detection for ADAS	ADAS	YOLOv3-Tiny	15 FPS, 99% Accuracy	Low-power FPGA implementation for intelligent transportation
Abdulhaq et al., 2025)	Real-Time Object Detection on Embedded Platforms	Embedded Vision	SSD, YOLO, Mobile Net	Comparative Study	Lightweight detectors perform better on embedded devices
Tian et al., 2024)	Comparative Analysis of DL Detectors	Embedded Systems	Faster R-CNN, YOLO, SSD	Speed vs Accuracy	SSD MobileNet-V2 and Tiny-YOLO achieve the best trade-off
(Saiveena & Praveen, 2024)	Hybrid DETR + YOLOv8 for Autonomous Vehicles	Autonomous Driving	DETR + YOLOv8	Reduced latency	Hybrid architecture improves tiny object detection.
Li and Feng, 2025)	DBC-YOLO: Tiny Object Detection for Drones	Drone Surveillance	DBC-YOLO	104 FPS	Excellent tiny-object detection with edge computing support.
(Kumar et al., 2026)	YOLO-Based Fire Detection: Systematic Review	Fire Detection	YOLO Variants	Review	Summarizes performance and computational trade-offs.
(Ganapathi et al., 2025)	FLARE:Lightweight Thermal Image Detection	Thermal Imaging	FLARE	0.72% mAP Improvement	Ultra-lightweight detector outperforming YOLOv8n
(Rahman et al., 2025)	YOLO + UNET Fusion for Road Segmentation	Autonomous Driving	YOLO + U Net	Real-time segmentation	Improves road segmentation in autonomous driving
(Sun et al., 2023)	Visible-Infrared Image Fusion for Object Detection	UAV Detection	Optimized YOLOv2	97.43% Accuracy	Multi-modal fusion improves object detection performance
(Shabbir et al., 2025)	LicensePlate Recognition on Raspberry Pi	Security	YOLOv8 + Easy- OCR	Low Latency	Efficient edge-based automatic license plate recognition
(Wang et al., 2025)	Micro-LED Defect Inspection	Industrial Inspection	Pruned YOLO + CORA	F1-score = 0.9436	Reduces parameters by 51% while improving defect detection

3. METHODOLOGY

3.1 System Overview

The proposed framework seeks to enhance the mobility and safety of visually impaired pedestrians in terms of navigation by merging the functionalities of computer vision, distance measurement and speech-based assistance into a single framework. The architecture consists of four functional modules: (i) image acquisition using a camera, (ii) real-time object detection using the YOLOv5) model, (iii) distance estimation using a LiDAR sensor, and (iv) audio guidance through gTTS. The overall architecture of the proposed assistive navigation system is shown in Fig. 2. Initially, the camera continuously captures video frames from the surrounding environment. Each frame is processed by the trained YOLOv5 model to detect and classify nearby objects. The LiDAR sensor simultaneously measures the distance between the user and the detected objects. Finally, the detected object labels and distance information are converted into speech using the gTTS engine and delivered to the user through earphones or speakers.

3.2 Hardware Configuration

The proposed prototype consists of a Raspberry Pi 3 Model B+ equipped with a Broadcom BCM2837B0 Cortex-A53 64bit processor operating at 1.4 GHz and 1 GB LPDDR2 SDRAM. A USB camera is employed for real-time image acquisition, while a RPLIDAR A1M8 sensor is integrated for obstacle distance estimation. The system operates on Raspberry Pi OS and communicates with external audio devices through Bluetooth or wired headphones. The experimental evaluation was also performed on a desktop computer equipped with an Intel Core i7 processor, 16 GB RAM, and NVIDIA GPU for model training. Python 3.10 together with the PyTorch DL framework was used for implementation.

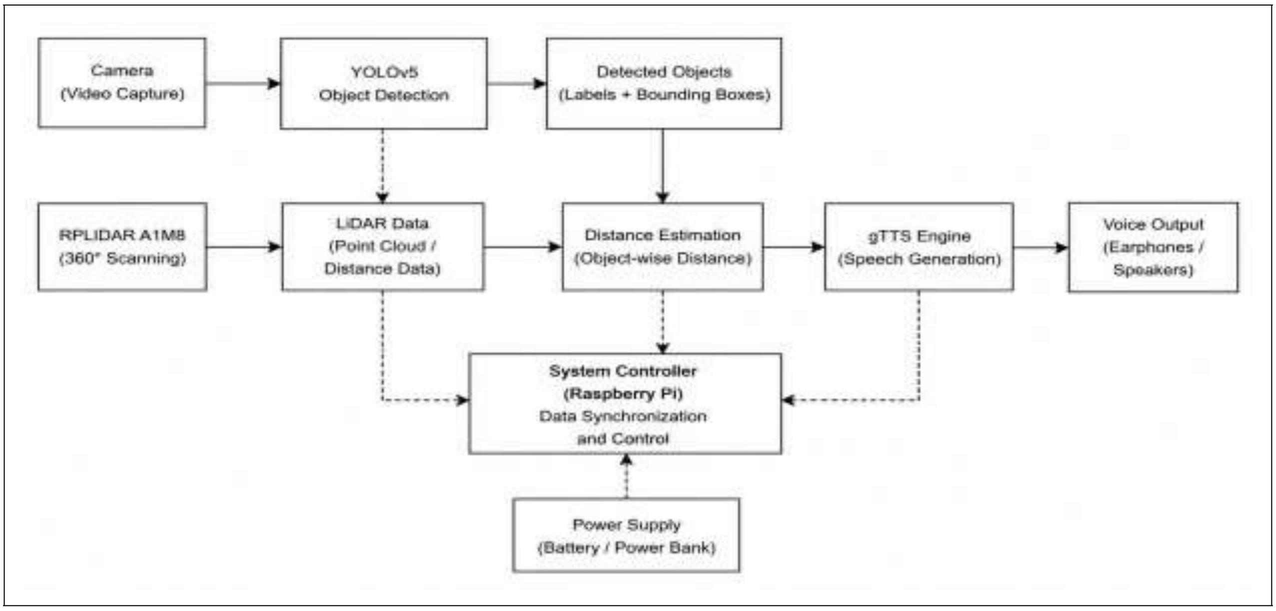

Fig. 2: Architecture of the proposed assistive navigation system for the blind pedestrians

3.3 Dataset Preparation

For the object detection model, a custom annotated dataset is used with 1,250 images that were sourced from two places: from publicly available google images, and from real images taken with a camera under various conditions. The data set consists of several classes of objects that are frequently encountered in pedestrian navigation, including pedestrians, vehicles, bicycles, motorcycles, traffic sign and other road-side obstacles. Pictures were manually annotated with bounding boxes in the YOLO annotation format. The data was split into training, validation, and testing sets in a ratio of 80:10:10. To achieve generalization and mitigate overfitting, the images were subjected to data augmentation methods such as horizontal flipping, random scaling, brightness adjustment, and random cropping. Before training, all of the input images were resized at a resolution of 640x640 pixels.

3.4 YOLOv5 Object Detection Model

YOLOv5 was chosen for the proposed assistive navigation system due to its superior detection accuracy, swift inference speed, and compatibility with embedded and real-time applications. YOLOv5 is distinguished from the traditional two stage object detector in that it detects objects and identifies them in the same neural network, thereby saving computing power and processing speed. The network architecture consists of three components: Backbone for extracting discriminative image features, Neck for fusing multi-scale feature information to enhance detection over objects of different sizes, and Detection Head for class prediction and bounding-box coordinates for the detected objects. The integrated architecture provides efficient and accurate object detection, making YOLOv5 an ideal algorithm for assistive real-time navigation for walking visually impaired people. During training, the model minimizes a composite loss function consisting of localization loss, confidence loss, and classification loss.

$$L = Lbox + Lobj + Lcls \tag{1}$$

where $Lbox$ represents the bounding-box regression loss, $Lobj$ denotes objectness loss, and $Lcls$ corresponds to classification loss.

The model was trained using the Adam optimizer with an initial learning rate of 1×10−3, batch size of 16, image resolution of 640×640, and 100 training epochs. The confidence threshold was fixed at 0.25, while the Intersection over Union (IoU) threshold was set to 0.50 during inference.

3.5 Distance Estimation Using LiDAR

While YOLOv5 identifies surrounding objects, it does not provide accurate distance measurements. Therefore, a LiDAR sensor is incorporated to estimate the distance between the user and detected obstacles. The real-time obstacle distance estimation was implemented using RPLIDAR A1M8 ($Slamtec$). The sensor is capable of a 360° scanning range, an effective detection distance of up to 12 m, a scanning frequency of 5.5 Hz, and communicates via a USB/serial interface, making it suitable for real-time assistive navigation applications. For each detected object, the corresponding LiDAR measurements are synchronized with the camera frame. The nearest obstacle distance is calculated as:

$$D = min(d1, d2, \dots, dn) \tag{2}$$

where di represents the measured distance of the i^{th} LiDAR point associated with the detected object.

Based on predefined safety thresholds, obstacles are categorized as near, moderate, or far to prioritize warning messages.

3.6 Voice Feedback Module

The detected object class and corresponding distance are converted into natural speech using the gTTS engine. The generated speech message informs the user about both the object type and its approximate distance. For example, (a) "Person detected two meters ahead", (b) "Vehicle approaching from the left", and (c) "Bicycle detected nearby." Only high-priority obstacles within the predefined safety range are announced to reduce unnecessary audio interruptions and improve user experience.

3.7 Performance Evaluation

The proposed model was evaluated using standard object detection metrics, including Precision, Recall, and mean Average Precision (mAP).

(a) Precision is calculated as:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3}$$

(b) Recall is computed as

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively. The mean Average Precision (mAP) at an Intersection over Union threshold of 0.5 is calculated as:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{5}$$

where AP_i represents the average precision for the i^{th} object class and N denote the total number of object classes.

The real-time capability of the proposed system was further evaluated using inference speed measured in frames per second (FPS), defined as:

$$FPS = \frac{Number\ of\ processed\ Frames}{Execution\ Time} \tag{6}$$

The proposed system achieved real-time performance while maintaining high object detection accuracy, demonstrating its suitability for assistive navigation applications.

4. RESULTS

This proposed system was evaluated in real-world urban environments, including busy streets and intersections. Fig. 3 shows the sample annotated image from the training dataset used for YOLOv5 object detection. Key performance metrics included detection accuracy, processing speed, and the user experience with voice feedback. Table 2 summarizes the quantitative performance of the proposed YOLOv5-based assistive navigation system. While Fig. 4 shows the precision-recall graph of proposed YOLOv5 model of the different detected object classes. The model retrieved high precision (94.82%) and mAP@0.5 (91.57%) to ensure real-time pedestrian navigation without losing accuracy in object detection. Moreover, the system achieved a frame rate of about 30 Frames Per Second (FPS), which allowed for the real-time processing of video streams without significant delays. The results show that the proposed scheme can achieve a good balance between detection accuracy and computational complexity, which is suitable for assistive navigation applications. YOLO still struggling with smaller objects compared to its closest competitors. For categories like "bottle," "sheep," and "tv/monitor," YOLO scores 8-10% lower than R-CNN or other state-of-the-art methods. However, YOLO performs better in categories like "cat" and "train," demonstrating its strengths in certain areas (Redmon et al., 2016). Table 3 encapsulates the accuracy of the suggested and comparative techniques, as described by the author. The feature set map of the suggested and related techniques is presented in Fig. 5. The accuracy of the method will be fulfilled by the feature set mapped to the model. The R-CNN +Yolo deliberately enhanced the accuracy of the object detection using the pre-defined shapes. The shape of objects is trained with different types of samples in different dimensions (Alagarsamy et al., 2023).

Fig. 3: Sample annotated image from the training dataset used for YOLOv5 object detection.

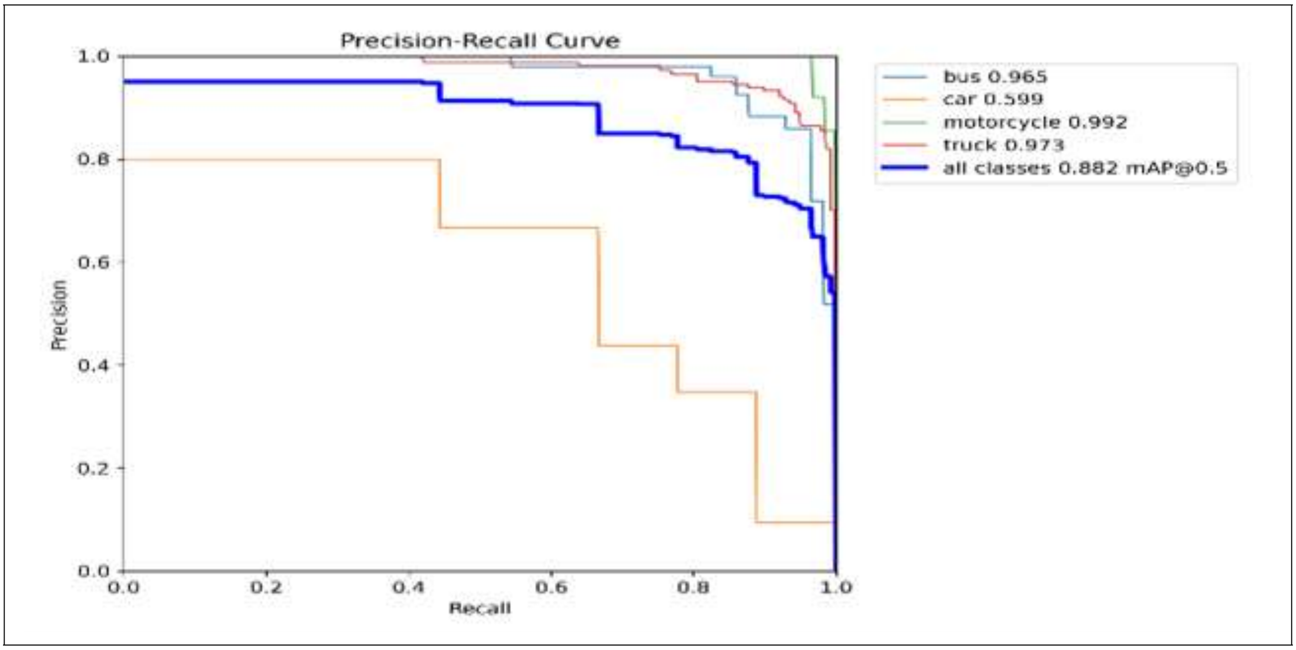

Fig. 4: Precision–Recall curves for the object classes detected by the proposed YOLOv5 model.

Table 2: Performance summary of the proposed Yolov5-based assistive navigation system

Performance Metric	Value
Precision	94.82%
Recall	69.82%
mAP@0.5	91.57%
mAP@0.5:0.95	91.03%
Inference Speed	30 FPS
Image Resolution	640×640 pixels
Training Epochs	100
Batch Size	16
Confidence Threshold	0.25
IoU Threshold	0.50

Table 3: Accuracy Comparative Study

S. No.	State of the Art Methods	mAP @ .5 (%)
1	CNN	90.23
2	RCNN	91.26
3	YOLO	89.36
4	Edge-based Detection	88.78
5	Hypernet	84.45
6	YOLOv5	91.57

5. DISCUSSION

This section focuses on the discussion based on usability challenges. During user testing, participants indicated that rapid or excessive voice notifications could become overwhelming, especially in busy urban environments. To enhance user experience, it is essential to prioritize critical alerts, such as warnings about vehicles at intersections, over non-essential notifications. Taiwo et al. (2024) emphasized the importance of customizable voice settings that allow users to control the type and frequency of alerts, a recommendation that received strong support from their findings.

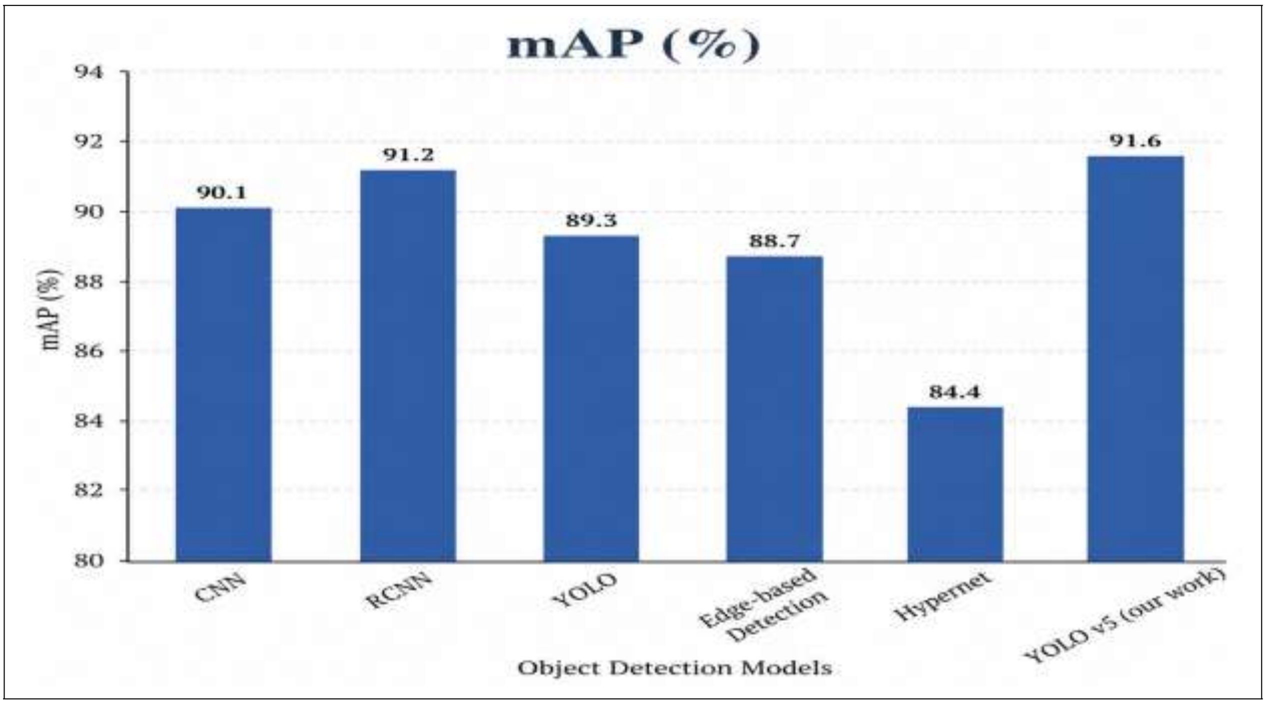

Fig. 5: Comparative performance analysis of the proposed YOLOv5-based framework and existing object detection methods.

6. CONCLUSION

This paper introduced a deep learning based assistive navigation system for visually impaired pedestrians and combined the YOLOv5 object detection model, an RPLIDAR A1M8 sensor for distance measurement to obstacles, and Google Text-to-Speech (gTTS) for real-time voice guidance. The proposed framework integrates visual object recognition with precise distance estimation for timely and context-aware navigation support in complex urban scenes. Experimental results showed that the system was able to operate in real time at around 30 frames per second and achieved a precision of 94.82, a recall of 69.82, an mAP@0.5 of 91.57 and an mAP@0.5:0.95 of 91.03. The outcomes demonstrate the feasibility of the proposed system to improve the environmental awareness, mobility and independence of the users. Future research will emphasize enhancing the detection capability under difficult weather and lighting conditions, adding multilingual and personalized voice assistance, and thorough field testing to evaluate the viability and reliability of the proposed framework.

CREDIT AUTHOR STATEMENT

1. **Deepika Yadav:** Conceptualization, Methodology, Visualization
2. **Hemant Yadav:** Validation, Formal analysis, Writing – Review & Editing
3. **Pooja Yadav:** Investigation, Supervision, Writing – Review & Editing

REFERENCES

World Health Organization. (2019). World report on vision: Executive summary. Accessed: 2026-07-15. https://www.who.int/publications/i/item/world-report-on-vision

Dongre A. (2020). Visual recognition-based system to assist blind persons. PhD thesis, National College of Ireland, Dublin. https://norma.ncirl.ie/4143/

Chen Y., Shen J., & Sawada H. (2023). A wearable assistive system for the visually impaired using object detection, distance measurement and tactile presentation. Intelligence & Robotics, Vol. 3, Issue 3, pp. 420–435. https://www.oaepublish.com/articles/ir.2023.24

Patel I., Kulkarni M., & Mehendale N. (2024). Review of sensor-driven assistive device technologies for enhancing navigation for the visually impaired. Multimedia Tools and Applications, Vol. 83, Issue 17, pp. 52171–52195. https://doi.org/10.1007/s11042-023-17552-7

Srikanteswara R., Reddy M. C., Himateja M., & Kumar K. M. (2021). Object detection and voice guidance for the visually impaired using a smart app. Recent Advances in Artificial Intelligence and Data Engineering: Select Proceedings of AIDE 2020, pp. 133–144. Springer. https://doi.org/10.1007/978-981-16-3342-3_11

Saravanan R., Nakkeeran R., & Baskar K. (2022). An assistive device based on recent deep learning approach for object detection integrated with audio guidance for visually impaired people. International Journal of Analytical and Experimental Modal Analysis.

Tavakoli Y. A. (2024). Seeing with sound: Object detection and localization by YOLOv8 and audio feedback for blind individuals. https://thesis.unipd.it/handle/20.500.12608/66485

Alrowais F., Almofarreh M., & Marzouk R. (2026). Enhanced pedestrian walkway object detection using deep learning and pelican optimization algorithm for assisting disabled persons. 16, Scientific Reports. https://doi.org/10.1038/s41598-025-32129-0

Zou Z., Chen K., Shi Z., Guo Y., & Ye J. (2023). Object detection in 20 years: A survey. Proceedings of the IEEE, Vol. 111, Issue 3, pp. 257–276. https://doi.org/10.1109/JPROC.2023.3238524

Lee J., Cha K.-A., & Lee M. (2024). Multi-modal system for walking safety for the visually impaired: Multi-object detection and natural language generation. Applied Sciences, Vol. 14, Issue 17, p. 7643. https://doi.org/10.3390/app14177643

Sheela K. S., Kingsly R. J., Bhalaji C.K., Selvalakshmi C.B., Radhika A. (2024). Traffic sign categorization using YOLO algorithm: Leveraging real-time object detection for improved road safety. 9th International Conference on Communication and Electronics Systems, pp. 1547–1553. https://doi.org/10.1109/ICCES63552.2024.10860011

Mathivanan P., Abishek N., et al. (2024). Computer vision empowered assistive technology for blind people. IEEE International Conference on Emerging Research in Computational Science, pp. 1–7. 10.1109/ICERCS63125.2024.10895124

Gao W. (2024). YOLO-based gripping method for industrial robots. International Journal of Computer Applications in Technology, Vol. 75, Issue 1, pp. 48–57. https://doi.org/10.1504/ijcat.2024.144660

Gui S., Song S., Qin R., & Tang Y. (2024). Remote sensing object detection in the deep learning era—a review. Remote Sensing, Vol. 16, Issue 2, p. 327. https://doi.org/10.3390/rs16020327

Al Amin R., Hasan M., Wiese V., & Obermaisser, R. (2024). FPGA-based real-time object detection and classification system using YOLO for edge computing. IEEE Access, Vol. 12, pp. 73268–73278. https://doi.org/10.1109/ACCESS.2024.3404623

Abdulhaq K., & Ahmed A. A. (2025). Real-time object detection and recognition in embedded systems using open-source computer vision frameworks. International Journal of Electrical Engineering and Sustainability, pp. 103–118. https://doi.org/10.65998/ijees.v3i1.113

Tian J., Jin Q., Wang Y., Yang J., Zhang S., & Sun D. (2024). Performance analysis of deep learning-based object detection algorithms on COCO benchmark: A comparative study. Journal of Engineering and Applied Science, Vol. 71, Issue 1, p. 76. https://doi.org/10.1186/s44147-024-00411-z

Saiveena K., & Praveen P. (2025). A hybrid deep learning approach for vehicle detection and image segmentation in autonomous driving. 2025 8th International Conference on Computing Methodologies and Communication (ICCMC), pp. 1289–1294. 10.1109/ICCMC65190.2025.11140811

Li Y., & Feng L. (2026). Global-local collaborative learning: A dynamic and attentive framework for tiny object detection in aerial imagery. IEEE International Conference on Embedded Systems, Mobile Communication and Computing (EMC²), pp. 193–198. 10.1109/EMC68537.2026.11441937

Kumar K., Sharma M., & Sharma R. (2026). A review of YOLO-based realtime fire detection models and applications. IEEE International Conference on Signal Processing and Electronics Design, pp. 96–101. 10.1109/ICSPED65817.2026.11448339

Ganapathi I. I., Abdelhafez F. O., Velayudhan D., Habtie M. A., Karki H., Al Awadhi K. Y., Dias J., & Werghi N. (2025). Flare: Flare analysis and regression-based estimation. IEEE International Conference on Advances in Data-Driven Analytics and Intelligent Systems, pp. 1–6. 10.1109/ADACIS65663.2025.11437182

Rahman W., Roy S., & Islam A. (2025). A comprehensive model for advanced road scene understanding: YOLO-CNN fusion for accurate road segmentation and object detection in varied conditions. PhD thesis, Brac University.

Sun Y., Meng Y., Wang Q., Tang M., Shen T., & Wang Q. (2023). Visible and infrared image fusion for object detection: A survey. International Conference on Image, Vision and Intelligent Systems, pp. 236–248. Springer. https://doi.org/10.1007/978-981-97-0855-0_24

Shabbir N., Ahmed R., Raza M. W., Zeb A., Elahi H., & Waheed H. (2025). Comparative performance and resource utilization analysis of OCR models for number plate recognition on Raspberry Pi 4. IEEE International Conference on Communication Technologies, pp. 1–6. 10.1109/ComTech65062.2025.11034455

Wang, J., Xie, Z., & Chen, M. (2025). An intelligent framework for micro-LED defect detection with deep learning. IEEE Sensors Journal, vol. 26, Issue 4, pp. 5749-5756. 10.1109/JSEN.2025.3645818

Redmon J., Divvala S., Girshick R., & Farhadi A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788. 10.1109/CVPR.2016.91

Alagarsamy, S., Rajkumar, T. D., Syamala, K., Niharika, C. S., Rani, D. U., & Balaji, K. (2023). A real-time object detection method for visually impaired using machine learning. 2023 International Conference on Computer Communication and Informatics (ICCCI), pp. 1–6. 10.1109/ICCCI56745.2023.10128388

Taiwo R., Bello I. T., Abdulai S. F., Yussif A.-M., Salami B. A., Saka A., & Zayed T. (2024). Generative AI in the construction industry: A state-of-the-art analysis. https://arxiv.org/pdf/2402.09939

AUTHOR'S BIOGRAPHY

Deepika Yadav is currently pursuing a Ph.D. in Computer Science at M.J.P. Rohilkhand University, Bareilly, Uttar Pradesh, India, which she commenced in October 2024. She has over two years of teaching experience in Computer Science, with expertise in programming, core computing subjects, and mentoring students in academic projects. She has also qualified the UGC NET-JRF examination. Her research interests encompass Artificial Intelligence, Deep Learning, Medical Image Analysis, and AI-driven Healthcare Applications. Her research focuses on developing intelligent, automated, and efficient solutions to real-world healthcare challenges by leveraging advanced computational techniques and artificial intelligence. Her work aims to contribute to the development of innovative AI-based approaches that can enhance healthcare decision-making, diagnosis, and patient care.

Hemant Yadav holds a B.Tech. in Computer Science & Information Technology, an M.Tech., an MBA, and a Ph.D. He currently serves as Group Director at Future University, Bareilly, India. With more than 25 years of experience in teaching, research, academic leadership, and professional training, he has held several key academic and administrative positions, including Head of the Department of Computer Science, Assistant Director, Deputy Director, and Director. Dr. Yadav has made significant contributions to research in diverse areas, including Artificial Intelligence, Machine Learning, Data Mining, the Internet of Things, Healthcare Technologies, and Information Technology. He has authored more than 30 research papers published in reputed journals and presented at National and International Conferences. In addition to his research contributions, he has actively participated in and led faculty development, executive development, and academic training programmes. His professional profile reflects a strong combination of academic leadership, research expertise, institutional development, and capacity building, with a continued focus on advancing education, research, and technology-driven innovation.

Pooja Yadav is an Associate Professor in the Department of Computer Science and Information Technology, Faculty of Engineering and Technology, M.J.P. Rohilkhand University, Bareilly. She obtained her B. Tech degree in Computer Science and Engineering in the year 2002, from Purvanchal University. She received her M. Tech. degree in Computer Science and Engineering from Aligarh Muslim University, Aligarh, in the year 2005. She did her Ph.D. in the field of Computer Science from the Indian Institute of Technology Roorkee. She has 20 years of experience in academics. Her research interest includes Data Mining, Internet of Things, Machine Learning, and Deep Learning, especially in the Healthcare sector. She has published research papers in various refereed journals and conferences of international repute. She is an active reviewer of the reputed Journals. She is the author of 4 books and some book chapters. She is actively involved in Teaching Learning Process, along with departmental duties. She is a member of the Institution of Engineers (India), the Institute of Electrical and Electronics Engineers (IEEE), and the Computer Society of India.