iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79789

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Real-Time Object Detection and Navigation Assistance for the Visually Impaired

Mr. B. Vamshi[1], B. Ashritha[2], K. Hithasree[3], S. Sanjana[4], E. Sumanjali[5]
[1] Assistant Professor, [2, 3, 4, 5] Student, Dept of Computer Science and Engineering,G. Narayanamma Institute of Technology and Science, Hyderabad, India – 500090

Abstract: Visually impaired individuals were encountering serious issues in achieving safe and independent mobility both indoors and outdoors. All the assistive technologies available in the current state of the art, including white canes, guide dogs, smart canes, and advanced AI-based systems, were providing only partial assistance due to their respective limitations such as poor obstacle detection, high cost, low adaptability, and improper usage in continuous real-time applications. An assistive computer vision-based obstacle detection and navigation system for visually impaired individuals utilizing deep learning is proposed to address these problems. It employed an SSD-MobileNet model for accurate obstacle detection and identification of commonly found objects such as chairs, beds, humans, and vehicles. Audio and haptic feedback will be employed in real-time to assist in safe mobility. The proposed solution will be economical, adaptable, and efficient for real-time applications. The experimental outcomes revealed that the system improved the dependability of the navigation, user safety, and independence in terms of mobility of the visually impaired towards a better quality of life.
Keywords— Machine Learning, Deep Learning, Computer Vision, SSD-MobileNet, Obstacle Detection, Object Recognition, Real-Time Feedback.

I. INTRODUCTION

The factor of mobility and orientation is another factor that plays an important role in ensuring that visually impaired persons are independent, but it also poses a challenge to visually impaired persons. In the past, assistive devices such as the white cane and guide have to a certain extent been effective in ensuring that visually impaired persons are able to overcome the challenges that they face in relation to mobility and orientation. However, there are a number of disadvantages that are associated with traditional assistive devices, such as the obstacles that traditional assistive devices are able to clear, traditional assistive devices' inability to clear obstacles that are above and distant from them, and the fact that traditional assistive devices are expensive and require training to enable one to use traditional assistive devices effectively. As a result, visually impaired persons still face a number of challenges in relation to their safety and confidence while moving around an unfamiliar environment.

Recent developments in Artificial Intelligence (AI) and Computer Vision have opened up new avenues for the development of intelligent assistive systems to address these drawbacks. However, many of the existing systems are still not very useful for practical implementation. They are often expensive, bulky, or computationally complex, resulting in lower portability and accessibility. Furthermore, some of the existing systems face problems in accurate detection in indoor or dynamic environments, while others produce delayed or excessive feedback, which can be overwhelming to the users.

In an effort to increase the safety and movement of visually impaired persons, a new real-time assistive navigation system isproposed in this paper. The proposed system ensures the continuous recording of the environment through real-time video streaming from the camera of a smartphone. The proposed system makes effective and accurate use of the Single Shot MultiBox Detector with the MobileNet model, as it is previously trained with the COCO dataset. The approximate distance of the user from the obstacles is also calculated through the object distance computation unit. The proposed system provides clear and simple voice feedback to the user based on the information. The proposed system is practical and simple to be used in real-life situations. The proposed system is focused on the low cost, speed, and quick response time required for the safe and confident movement of visually impaired persons.

II. LITERATURE SURVEY

Using an RGB-D camera, the work [1] suggests a vision-assisted indoor navigation system with semantic mapping and localization. The method uses a Kalman Filter for improved positioning and SVD map alignment. The user is further assisted during navigation by speech and haptic interfaces, as well as real-time obstacle detection.

A wearable tool that uses gyros and ultrasonic sensors to detect obstacles is suggested by a study [2]. Flight time and threshold filters are used to calculate the distance. When there are obstacles close by, the user receives audio alerts.

This research work [3] proposes an infrastructure-enabled navigation module by incorporating the integration of RGB-D images, IMU data, as well as GPS data. There also seems to be a multi-modal CNN that is involved in the understanding and execution of semantic segmentation. There also seem to be some previous map and infrastructure advancements in outdoor navigation.

The comparative analysis in [4] is about the development of a smart white cane using ultrasonic sensors along with GPS and GSM technology. The algorithms of distance thresholds are used to identify obstacles, which then trigger alerts. The cane is also designed to assist the user with vibrations and a buzzer.

The research work [5], Arduino-based assisting goggles with obstacle detection using YOLO and OpenCV, have been developed. The assisting goggles detect obstacles that are then converted into either voice or acoustic signals. The design has been made wearable and more portable.

The paper [6] discusses the development of a smart assisting navigator using Raspberry Pi and deep learning-based image processing algorithms. The CNN models are able to detect objects and perform facial recognition in real-time processing. The internet of things has enabled the sending of notifications and assistance in navigation.

The paper [7] introduces a supporting navigation system that incorporates GPS location tracking and ultrasonic sensors. Obstacle detection and user location determination have been enhanced through methods supporting fusion calculations. Real-time wireless connectivity facilitates surveillability and safe movement.

The comparative study [8] suggests an intelligent outdoor navigation system that combines CNN-based vision, Google Maps, and the GPS system. The feature extraction of semantic scenes is used for classifying obstacles.

III. PROPOSED SYSTEM

The system aimsto make a positive impact in the lives of visually impaired persons by enabling them to move independently and safely while being aware of their surroundings at all times. The proposed system has the intention of achieving this by using computer vision technology to identify objects in their surroundings and provide them with feedback. The system is developing a system that is real-time, cost-effective, portable, and user-friendly for use in real-life situations. Real-time video feedback from the surroundings can be obtained by using a camera module that provides video feedback to a mini computer. Video frames can be processed in real-time to provide accurate and timely feedback to visually impaired persons even in adverse situations.

With respect to the object detection and classification objective of the proposed system, it uses a pre-trained Single Shot MultiBox Detector SSD model with a MobileNet backbone due to its computational efficiency and detection accuracy. Accordingly, the proposed system utilizes the COCO dataset for training the SSD model and thus enables it to effectively detect different types of objects that are usually encountered during day-to-day life, such as pedestrians, cars, doors, stairs, furniture, and street-level obstacles. To enable the proposed system to be reliable in terms of detection accuracy, non-maximum suppression and confidence filtering are utilized.

In order to increase the efficiency of the spatial perception, there is also a module for the estimation of the distances, and it is included in the system to carry out the numerical calculations according to the data obtained from the bounding box. The distance may be referred to as near, medium, and far distances according to the size and position of the objects. The information obtained from the objects and their distance is converted to small amounts of feedback by using the text-to-speech technology, which provides feedback to the system in the form of sound. This reduces the workload and helps in taking decisions quickly without the risk of collision. Additionally, the modularity provides the chance to add other features to the system, such as providing vibration feedback and updates to the GPS and deep learning model, to develop an efficient assistive technology system to increase the quality of life for visually impaired persons.

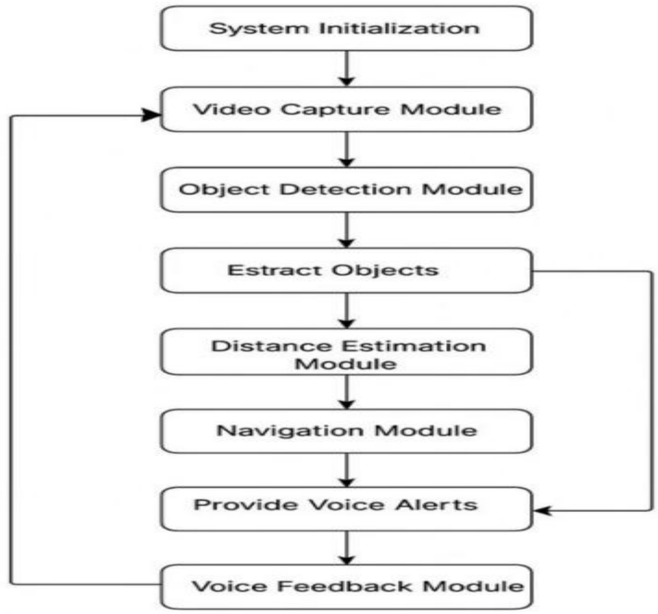

Fig. 1.Data Flow of the System

IV. IMPLEMENTATION

The proposed assistive mobility and navigation system is expected to be designed and developed in such a way that it is able to provide accurate and real-time navigation assistance using computer vision, deep learning, and audio feedback technology. The main objectives of the proposed system are to be real-time responsive, computationally efficient, cost-effective, and user-friendly, so that it can be effectively utilized for assisting navigation in indoor and outdoor environments. The proposed system is designed and developed on a Python-based platform because of its flexibility and strong support for image processing and machine learning libraries. TensorFlow is used for the execution of deep learning algorithms, and OpenCV is used for video capture, processing, and real-time execution. The smartphone camera is the hardware module used for image and video capture. It is continuously capturing live video frames to track the environment.

The captured video stream is then processed frame by frame in real-time. The frames are then exposed to various preprocessing tasks such as resizing the frame to the required input resolution, pixel normalization to ensure the robustness of the model, and finally converting the frame to a tensor format that is compatible with the object detection model. Various preprocessing tasks are extremely important in ensuring that the frames are standardized and that the model is able to function properly despite the environmental changes, such as lighting conditions. OpenCV assists in the successful execution of the tasks with zero latency.

Fig. 2. COCO Dataset

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Object detection is performed by employing a pre-trained Single Shot MultiBox Detector (SSD) with a MobileNet model. The rationale behind choosing this model is that this model is a lightweight model and can be applied for real-time applications, especially in scenarios where the computational power is less. The SSD MobileNet model is trained on the COCO dataset, as depicted in Fig 2, which includes a vast number of objects that are normally encountered, such as pedestrians, cars, furniture, and other objects that can be considered as obstacles during navigation. The SSD model has the capability to detect and classify multiple objects in a single frame.

For the purpose of enhancing the reliability of the detection result, non-maximum suppression (NMS) is applied as a post-processing method. NMS suppresses the unnecessary bounding boxes by choosing the detection with the highest confidence score for each object. The MobileNet model applies depthwise separable convolutions, which have been able to decrease the computational complexity to a large extent while maintaining an acceptable level of accuracy. This enables the system to function correctly even when it is always running.

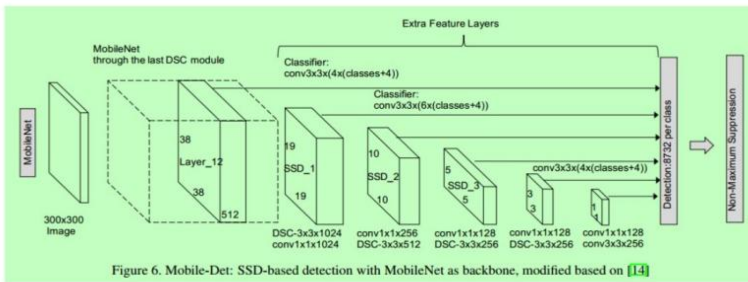

Fig. 3. Architecture of SSD with MobileNet

In order to improve the spatial awareness of this system, a distance calculation module is integrated into this system, and the size of the bounding boxes is taken into account. A highly simplified version of the distance calculation module is integrated into this proposed system, in which the size of the detected objects on the video screen is increased if the object is closer to the camera and the size of the detected objects on the video screen is reduced if the object is far from the camera. This system includes mathematical calculations by making use of the NumPy library to calculate the distance between the user and the detected objects. Although the distance calculation is not highly accurate, there is enough information in this calculation to make a proper decision. The calculated distances are then categorized into three groups: near, medium, and far, depending on some threshold values. The categorization enables the system to provide feedback that is more comprehensible and enables the system to concentrate more on the near objects that are of immediate concern. The near objects provide instant notifications, medium-distance objects assist the user in staying alert about the surroundings, and far objects are of less concern to avoid disrupting the user. The approach of prioritization prevents the depletion of the user's mental resources and ensures that only relevant and up-to-date information is provided.

The distance estimation and object detection module is highly integrated for effective decision-making. After the object has been detected and its distance category determined, the system moves on to make a decision on whether to give an audio alert. The selective alerting system is highly effective in resource management and ensures that the system does not give frequent alerts, making the real-time process even smoother. Audio alerts are the primary interaction interface between the system and the user. The information about the identified object and distance is conveyed to the user in the form of an audio notification by using a text-to-speech library, i.e., 'pyttsx3.' This text-to-speech library is utilized for a specific purpose based on the following reasons:There is no scarcity of latency, and offline support is available too. Along with this, support for Python applications is available too. The audio notification is very short, conveying information about the name of the identified object and distance.

The system has a continuous processing loop that involves the capture of frames, detection and distance calculations, as well as audio feedback. Basic timing control mechanisms are also included to prevent repetitive alerts unless there are significant changes in the position and proximity of objects.

In general, the deployment of the system is based on a modular and scalable architecture. The system is cost-effective and portable because of the use of lightweight models and hardware. The different modules of the system, such as object detection, distance estimation, and voice feedback, can be developed separately. The integration of real-time perception and audio feedback into one system makes it an efficient and effective solution for improving the mobility, safety, and independence of the visually impaired.

The implementation of the system is in a way that it follows the philosophy of modularity and scalability. The use of lightweight models, in addition to the hardware, makes the system cost-effective and portable. The modularity of each part of the system makes it possible for the system to be modified in case there are any future developments, such as object detection, estimation of distance, and voice feedback. The use of real-time perception, as well as voice feedback, makes the system capable of providing a highly effective solution for the enhancement of mobility for the visually impaired.

V. RESULTS

The testing of the system was done both indoor and outdoor to check its efficiency in performing object detection, range calculation, and audio alerts. On the basis of the experiment, the SSD MobileNet model performed very well in object detection with an overall accuracy of 92%. The precision and recall of the model are 90% and 88%, respectively, which gives an F1 score of 89%. The Mean Average Precision is 87%, and it determines the efficiency of the system in detecting different classes of objects that are encountered on a daily basis.

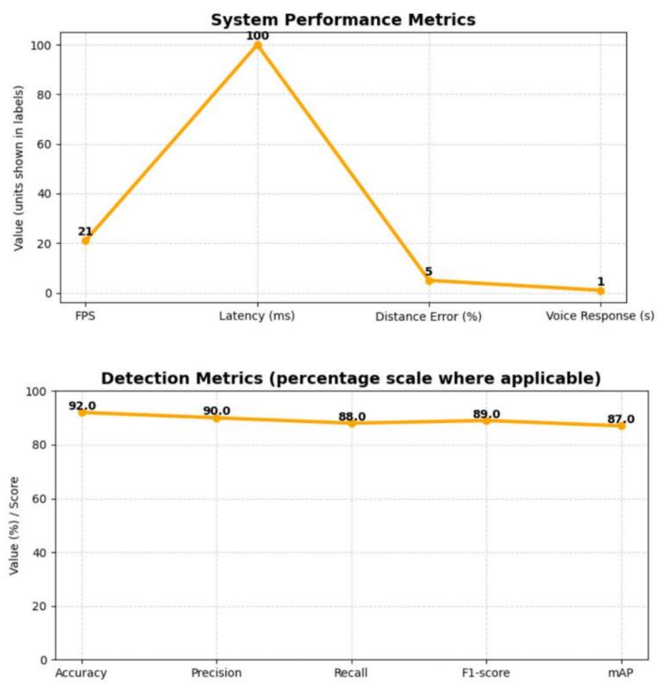

Fig. 4. Perfomance Measure graph of System

Regarding real-time performance, the average video streaming was conducted at a rate of 20-22 frames per second to allow smooth and continuous operation. The total latency of the entire system, from video capture to audio alert generation, was maintained below 100ms per frame to allow a quick response to the detected object. The accuracy in distance measurement maintained an average error rate below 5%.

Table I
Performance Measure of the Model

Accuracy	92%
Precision	90%
Recall	88%
F1 Score	89%
mAP	87%

The voice feedback feature consisted of informative and timely voice notifications about the type and distance to the object. The notifications could be easily discerned even in challenging scenarios involving multiple objects, and the factor of closer objects ensured that the viewer's attention was directed to the nearest one.

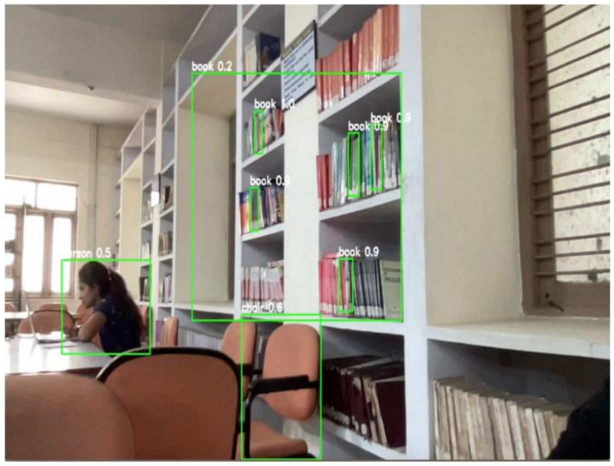

Fig. 5. Object Detection and Distance Estimation

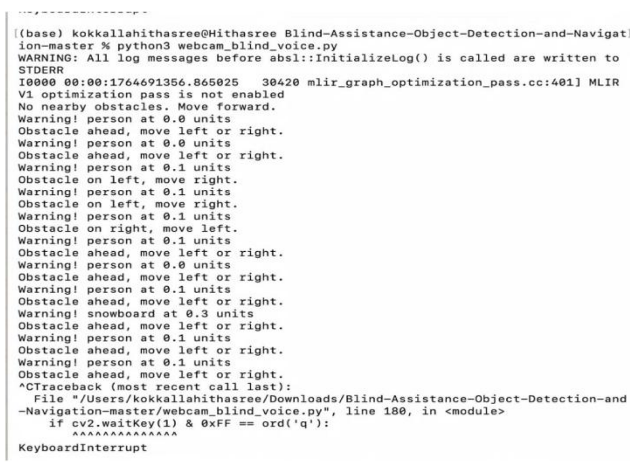

Fig. 6.Alerts and Navigation

The results obtained from the experiment prove the effective blend of accuracy, processing rate, and usability that the system provides. The effective use of accurate object recognition, accurate distance measurement, and sound feedback proves the efficacy of the results the system produces concerning the awareness that the environment provides.

VI. CONCLUSION

The proposed system is able to perform object detection, real-time distance estimation, and audio feedback. This would be made possible by the combination of computer vision and deep learning using an SSD MobileNet model that has been trained on visualizations of the COCO dataset to ensure the correct identification of the objects surrounding the system and real-time proximity information. The experimental results demonstrate the effectiveness of the system with high accuracy in the detection, low latency, and correct distance estimation, thus suitable for practical applications. It is a lightweight, cost-effective, and user-friendly framework that can be applied both indoors and outdoors. The real-time audio feedback enables the user to get clear and actionable information without being cognitively overloaded. In conclusion, it can be said that the system is a correct example of the combination of technology innovation and usability, thus providing an efficient system for environmental monitoring.

REFERENCES

[1] B. Li et al., "Vision-Based Mobile Indoor Assistive Navigation Aid for Blind People," in IEEE Transactions on Mobile Computing, vol. 18, no. 3, pp. 702-714, 1 March 2019.

[2] F. Shaikh, M. A. Meghani, V. Kuvar and S. Pappu, "Wearable Navigation and Assistive System for Visually Impaired," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2018.

[3] H. Son, A. Vora, G. Pandey and J. Weiland, "Infrastructure Enabled Guided Navigation for Visually Impaired," in IEEE Transactions on Intelligent Transportation Systems, vol. 26, no. 5, pp. 6764-6777, May 2025.

[4] I. N and V. Loganathan, "Comprehensive Assistive Mobility System for the Visually Impaired: The Smart Blind Stick with Ultrasonic Sensor Solution," 2024 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2024.

[5] J. R. B. Guevara et al., "Mobility Assistive Technology with Artificial Intelligence: Indoor Navigation Assistance for the Visually Impaired Using Arduino-Based Assistive Goggles," 2023 IEEE 15th International Conference.

[6] M. Kulkarni, M. Chitale, S. Chitpur, A. Chivate, P. Chopade and S. Deshmukh, "Smart Assistive Navigator for the Blind using Image Processing," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023.

[7] N. Al Mahmud, T. Hossain, R. V. Biswas, M. Foysal Ahammad Joy, M. Abu Raihan and D. Banik, "Assistive System and GPS Tracker for the Visually Impaired," 2023 26th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, 2023.

[8] O. Gamal, S. Thakkar and H. Roth, "Towards Intelligent Assistive System for Visually Impaired People: Outdoor Navigation System," 2020 24th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 2020.

[9] Salvador Mar-nez-Cruz 1 , Luis A. Morales-Hernández 1 , (Member, IEEE), Gerardo I. Pérez- Soto 2 , Juan P. Benitez-Rangel 1 , And Karla A Camarillo-Gómez 3 “An Outdoor NavigaRon Assistance System for Visually Impaired People in Public TransportaRon”. 26

[10] Z. Yang et al., "SeeWay: Vision-Language AssisRve NavigaRon for the Visually Impaired,"