

Design and Development of a Virtual Eye for Visually Impaired People in Indoor and Outdoor Applications

A. Selwin Mich Priyadharson*, M. Lasya Sree, Padachala Chandu, and Vella Yogeswar Reddy

Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

*Corresponding author: selwinmich@gmail.com

Abstract. Despite recent progress, people who are blind or visually impaired still face navigation difficulties, even in familiar surroundings. The current state-of-the-art approach to this problem is multi-sensor assistive technology, but additional sensors increase hardware cost and system complexity. This work proposes a camera-based navigation system that relies on vision alone and requires no additional sensors. The processing unit is the NVIDIA Jetson Nano, which can run real-time computer vision and deep learning algorithms. A single camera continuously captures image frames from the surrounding environment, and the processing module applies You Only Look Once (YOLO) object detection to the frames to recognize persons and objects such as vehicles, walls, furniture, and other obstacles. Audio instructions are delivered to the user's earphones as warnings or walking guidance. Qualitative tests showed that the system detected common objects in real time both indoors and outdoors, although detection performance dropped in dark conditions; detection accuracy and latency were not measured quantitatively. The single-camera design simplifies the system while still allowing users to understand their position relative to their surroundings.

Keywords. Computer vision; YOLO object detection; visually impaired assistance; NVIDIA Jetson Nano Developer Kit

I. INTRODUCTION

For most people, moving from one place to another is a normal and easy task. They can see everything around them and assess any obstacles, which makes navigation simple. Visually impaired people cannot see their surroundings, so they rely on touch to sense and identify obstacles. Navigating a familiar environment becomes easier over time because they gradually learn every part of their surroundings, but moving to a new place or adapting to a different environment is much more difficult. Every environment is unpredictable: a chair placed in the path or a moving vehicle can quickly become a danger, and visually impaired people often depend on others for their daily needs. Traditional mobility aids, such as white canes, provide physical support but detect objects only on close contact, so the user cannot anticipate what lies ahead.

In the last few years, researchers have developed electronic travel aids that use ultrasonic sensors, light detection and ranging (LiDAR), and wearable sensing modules. These systems improve awareness but also make the devices more complicated: more sensors mean more wiring, testing, and power use, and the technology itself can become difficult to manage. Research on assistive navigation for people with visual impairments has advanced quickly. Wang and colleagues [1] presented the YOLO-OD algorithm, a vision-based solution to obstacle detection for navigation assistance. Ben Atitallah and colleagues [2] created an embedded obstacle detection scheme, although hardware complexity remained an issue.

Boobalan and colleagues [3] built a system based on YOLOv7 that recognizes objects and provides voice guidance, which shows a high level of integration between machine perception and the human response. Around the same time, Choe and colleagues [4] assessed how well deep learning detectors run on NVIDIA Jetson devices. The work of Sarvajcz and colleagues [5], in which a Jetson Nano enabled real-time pedestrian and sign detection in road environments, brought further attention to edge-AI-based navigation. Obayya and colleagues [6] extended this concept by combining recurrent neural networks with YOLO for indoor assistance, which improved contextual understanding but increased computational requirements. He and Saha [7] focused on the challenges of outdoor navigation and tested various YOLO architectures under practical conditions. Pratap and colleagues [8] evaluated the latency of real-time detection algorithms for indoor navigation. Localization and object detection techniques have also been combined; a typical example is the fusion of visual simultaneous localization and mapping (SLAM) with YOLO developed by Javadian Farzaneh and Mahvash Mohammadi [9]. Islam and colleagues [10] developed a deep learning-based system that detects obstacles and describes the surroundings to assist visually impaired people.

People rely mainly on their eyes to understand their surroundings, and for blind people the camera can take the place of the eyes. Instead of combining multiple sensors, the proposed system uses only a camera. An NVIDIA Jetson Nano serves as the main processing unit and performs real-time

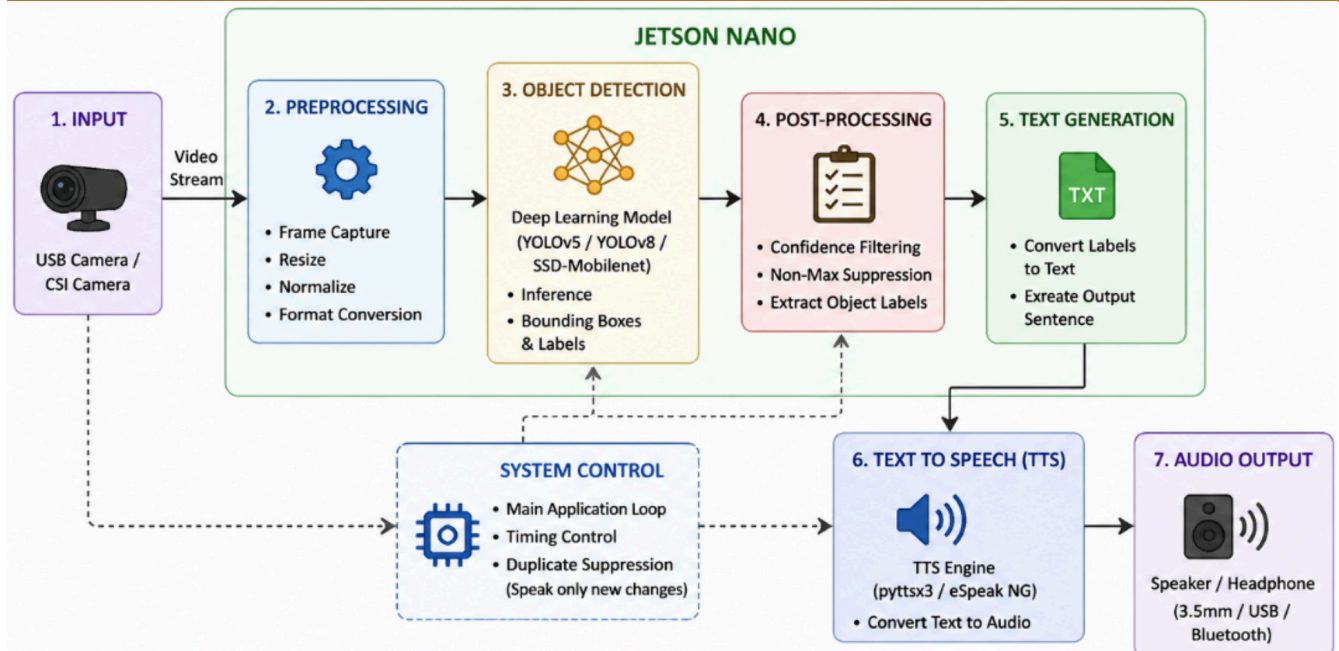


Fig. 1. Architecture of the virtual eye.

computer vision and deep learning operations. A camera continuously observes the environment, and each frame is analyzed by YOLO-based object detection. Objects or persons detected ahead are announced by voice, so the user can hear the warning and respond quickly. The system does not replace human decision-making; it simply provides awareness. Removing the sensors was intentional, because it makes the design cheaper, more affordable, simpler, and less dependent on hardware. Changing lighting, moving objects, and crowded places can make detection harder, but advances in embedded artificial intelligence (AI) make such solutions increasingly practical.

The aim of this work is therefore to design a simple, low-cost assistive navigation system that uses only a camera to perceive the environment and gives real-time audio instructions that help users move around safely and easily. Section II describes the proposed architecture, Section III the methodology, and Section IV the results and discussion. Section V concludes the paper.

II. PROPOSED ARCHITECTURE

The procedure is sequential, and each stage follows directly from the previous one. The architecture of the virtual eye is shown in Fig. 1.

A. Image Capture

The camera captures live images of the environment in which the user moves, which makes it the most critical sensory element of the system. It acts as the eyes of the entire system and supplies all the information the system requires about the environment. In this step, the camera records digital images or video frames of the environment at defined intervals.

A Logitech C270 HD webcam, connected through a USB interface, captures frames from the user's surroundings in real time. The camera supports HD video at a maximum frame rate of 30 fps, which is suitable for computer vision applications. The webcam is connected to the NVIDIA Jetson Nano Developer Kit, and video frames are captured

continuously and passed to the object detection model. The camera thus serves as the visual sensor of the navigation system.

B. Preprocessing

The Jetson Nano is shown in Fig. 2. It receives the captured frames, and as an embedded platform it can perform the parallel computations that deep learning tasks require. The NVIDIA Jetson Nano Developer Kit is an embedded computer designed for edge computing and AI workloads. It has a quad-core ARM Cortex-A57 CPU and a 128-core NVIDIA Maxwell GPU that delivers 472 GFLOPS for AI processing, together with 4 GB of LPDDR4 RAM, which makes it well suited to developing and running deep learning architectures, particularly object detection networks. The system can communicate through Gigabit Ethernet, USB, HDMI output, and MIPI CSI-2 camera inputs for processing live images. The board operates efficiently at 5 to 10 W.



Fig. 2. NVIDIA Jetson Nano Developer Kit.

C. Object Detection

The YOLO algorithm, integrated with OpenCV, analyzes every frame. In this pilot work, persons and the following objects are detected: chairs, cars, doors, walls, and general obstacles. Detection runs on each frame in real time.

D. Audio Feedback

The detected information is converted into speech with text-to-speech technology. Example outputs include "Person

ahead,” “Obstacle on the left,” and “Vehicle approaching.”

III. METHODOLOGY

The block diagram of the virtual eye is shown in Fig. 3. The proposed virtual eye was evaluated in real time in both indoor and outdoor settings. The goal of the evaluation was to determine how well the system could find objects, give timely audio feedback, and help users find their way.

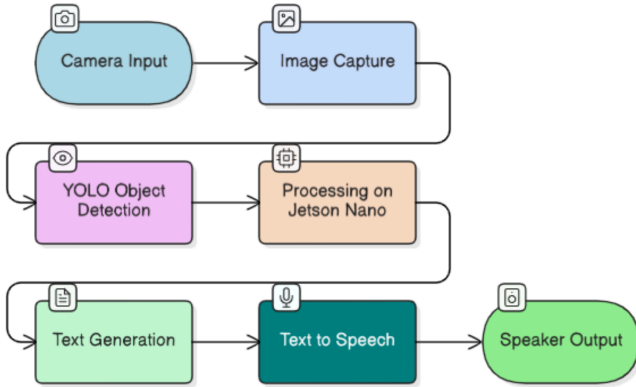


Fig. 3. Block diagram of the virtual eye.

The NVIDIA Jetson Nano was used as the embedded processing platform for the experiments, with a single monocular camera connected to it. The YOLO object detection model was run locally so that inference could take place in real time without the cloud. The test environments included indoor corridors and rooms, outdoor walkways, moderately crowded areas, and scenarios with static and dynamic objects. The camera was positioned to mimic a user's natural viewing direction, and earphones delivered the audio output to simulate real-world use by people with visual impairments. Table 1 lists the setup specifications.

TABLE 1. SPECIFICATIONS OF THE PROPOSED VIRTUAL EYE SETUP

Parameter	Specification
Processing unit	NVIDIA Jetson Nano
Input device	Logitech C270 HD webcam (USB)
Detection model	YOLO (real-time object detection)
Framework	OpenCV
Feedback method	Text-to-speech audio output
Processing mode	Edge computing (offline)

A. Object Detection Performance

The system detected typical navigation objects along the user's path, such as people, chairs, cars, doors, walls, and other obstacles. These objects were chosen because they are common navigation challenges for blind individuals. The main test criteria were object detection, responsiveness, and consistent performance during operation, since timely detection is critical for giving users accurate information about the objects around them.

In the authors' observations, the YOLO object detection model processed the captured frames in real time and detected multiple objects in the same frame (Figs. 5 and 6); detection accuracy and frame rates were not measured. The quality of real-time detection strongly affects how well a navigation system operates.

B. Audio Feedback Evaluation

The audio feedback was evaluated on the clarity and speed of the instructions produced. In the authors' observations, the audio alerts followed detections closely enough for the user to respond to changes in the environment, although the latency was not measured. The messages were kept brief to limit the listening load on the user.

C. Real-Time Navigation Assessment

The virtual eye setup is shown in Fig. 4. Navigation tests were carried out to analyze the behavior of the system while the user was moving. As obstacles appeared in front of the camera, real-time frame analysis triggered alerts. The main observations were real-time detection of obstacles while walking, detection of moving entities such as people and cars, real-time operation without Internet connectivity, and correct operation under normal lighting conditions, with lower detection performance in dark conditions. These observations suggest that combining visual perception with an embedded system can assist navigation without additional sensor inputs.

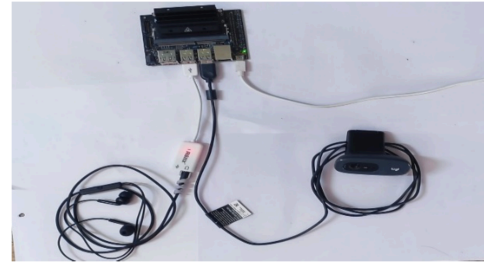


Fig. 4. Virtual eye setup with the Jetson Nano.

D. Advantages of the Camera-Only Approach

Removing the sensors made a considerable difference, and the system was easier to use than anticipated. The advantages include lower hardware cost, simpler wiring with fewer adjustments, portability and light weight, better maintainability, and adaptability across settings.

Sensor-based systems calculate distances directly, but their functionality may be limited in dense or cluttered environments. Vision systems, in contrast, are aware of their surroundings: a camera not only observes its environment but also identifies objects. A vision system can also be scaled, since software upgrades improve performance without altering the hardware. The effectiveness of a camera system depends on lighting, however, and detection performance dropped in dark conditions.

IV. RESULTS AND DISCUSSION

This paper has introduced an assistive navigation system that gives visually impaired people real-time environmental awareness and audio guidance. The system uses a monocular camera connected to the NVIDIA Jetson Nano Developer Kit, which acts as the embedded processing unit. The captured visual data are processed by a deep learning-based YOLO object detection algorithm together with the OpenCV computer vision framework, and the combination of both technologies allows efficient identification and classification of objects in the immediate environment. A single monocular camera offers a simple sensing process with reduced hardware

complexity while still allowing useful perception of the environment. In contrast to existing assistive devices, which depend heavily on many sensors, the camera-based system detected objects, including persons, furniture, vehicles, and other obstacles. Detections by the camera are shown in Figs. 5 and 6.

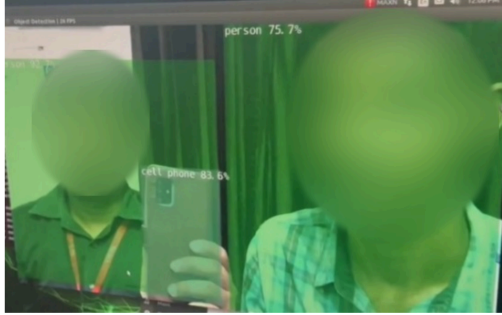


Fig. 5. Detection of persons and objects (example 1; faces blurred for privacy).

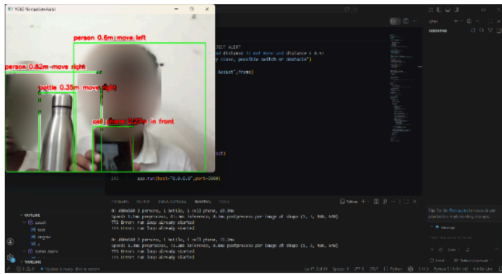


Fig. 6. Detection of persons and objects (example 2; faces blurred for privacy).

The experiments were carried out under different conditions, both indoors (chair, bottle, and television) and outdoors (bus and train). In these qualitative tests, the proposed solution recognized common navigation objects and announced them to the user through audio notifications; detection performance dropped in dark conditions.

V. CONCLUSION

This work developed a vision-based navigation system that recognizes objects in the surroundings and reports them to the user through a text-to-speech system. A USB camera was connected to the NVIDIA Jetson Nano Developer Kit to enable deep learning-based image processing. By combining the YOLO algorithm with computer vision methods, the system detects people, cars, doors, obstacles, and many other objects in the surroundings. The detected objects are converted into speech by the text-to-speech system, so the user can understand the situation by ear. Qualitative tests indicated that the device detected common objects in real time both indoors and outdoors, while detection performance dropped in dark conditions. Detection accuracy and latency were not measured quantitatively. Because the system requires no additional sensors, the camera-only design also reduces the size of the navigation device.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and preparation of this work and approved the final version of the manuscript.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

DECLARATION OF GENERATIVE AI USE

The authors declare that no generative artificial intelligence tools were used in the preparation of this manuscript. The authors take full responsibility for the content of the publication.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- [1] Wei Wang, Bin Jing, Xiaoru Yu, Yan Sun, Liping Yang, and Chunliang Wang, "YOLO-OD: Obstacle detection for visually impaired navigation assistance," *Sensors*, vol. 24, no. 23, Art. no. 7621, 2024, doi: 10.3390/s24237621.
- [2] Ahmed Ben Atitallah, Yahia Said, Mohamed Amin Ben Atitallah, Mohammed Albekairi, Khaled Kaaniche, Turki M. Alanazi, Sahbi Boubaker, and Mohamed Atri, "Embedded implementation of an obstacle detection system for blind and visually impaired persons' assistance navigation," *Comput. Electr. Eng.*, vol. 108, Art. no. 108714, 2023, doi: 10.1016/j.compeleceng.2023.108714.
- [3] P. Boobalan, S. Bhuvanikha, M. Sivapriya, and R. Sivakumar, "Object detection with voice guidance to assist visually impaired using YOLOv7," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 4, pp. 764-768, 2023, doi: 10.22214/ijraset.2023.50182.
- [4] Chungjae Choe, Minjae Choe, and Sungwook Jung, "Run your 3D object detector on NVIDIA Jetson platforms: A benchmark analysis," *Sensors*, vol. 23, no. 8, Art. no. 4005, 2023, doi: 10.3390/s23084005.
- [5] Kornel Sarvajez, Laszlo Ari, and Jozsef Menyhart, "AI on the road: NVIDIA Jetson Nano-powered computer vision-based system for real-time pedestrian and priority sign detection," *Appl. Sci.*, vol. 14, no. 4, Art. no. 1440, 2024, doi: 10.3390/app14041440.
- [6] Marwa Obayya, Fahd N. Al-Wesabi, Wafi Bedewi, and Menwa Alshammeri, "An intelligent framework for visually impaired people through indoor object detection-based assistive system using YOLO with recurrent neural networks," *Sci. Rep.*, vol. 15, Art. no. 43720, 2025, doi: 10.1038/s41598-025-27603-8.
- [7] Chenhao He and Pramit Saha, "Investigating YOLO models towards outdoor obstacle detection for visually impaired people," *arXiv:2312.07571*, 2023, doi: 10.48550/arXiv.2312.07571.
- [8] Abhinav Pratap, Sushant Kumar, and Suchinton Chakravarty, "Adaptive object detection for indoor navigation assistance: A performance evaluation of real-time algorithms," *arXiv:2501.18444*, 2025, doi: 10.48550/arXiv.2501.18444.
- [9] Mohammad Javadian Farzaneh and Hossein Mahvash Mohammadi, "Implementation of a blind navigation method in outdoors/indoors areas," *arXiv:2212.12185*, 2022, doi: 10.48550/arXiv.2212.12185.
- [10] Raihan Bin Islam, Samiha Akhter, Faria Iqbal, Md. Saif Ur Rahman, and Riasat Khan, "Deep learning based object detection and surrounding environment description for visually impaired people," *Heliyon*, vol. 9, no. 6, Art. no. e16924, 2023, doi: 10.1016/j.heliyon.2023.e16924.