NAVISIGHT: A Deep Learning and Voice-Assisted System for Intelligent Indoor Navigation of the Visually Impaired

Mrs. V.Vijaya Lakshmi [1], Karapakula Balaji Sai Vrushadree[2], K.Eekshitha[3], M.Hemavathi[4], T.Gurubunny[5]

[1] *Assistant* Professor, Department of *Electronics* & *Communication* Engineering , *Annamacharya Institute Of Technology &Sciences*, Tirupati, Andhra Pradesh, India, vijji.siri1992@gmail.com

[2] U.G Scholar, Department of Electronics & Communication Engineering , Annamacharya Institute Of Technology & Sciences, Tirupati, Andhra Pradesh, India, kbsv2004@gmail.com

[3] U.G Scholar, Department of Electronics & Communication Engineering , Annamacharya Institute Of Technology & Sciences, Tirupati, Andhra Pradesh, India,koppalaeekshitha@gmail.com

[4] U.G Scholar, Department of Electronics & Communication Engineering , Annamacharya Institute Of Technology & Sciences, Tirupati, Andhra Pradesh, India, hemameduri320@gmail.com

[5] U.G Scholar, Department of Electronics & Communication Engineering , Annamacharya Institute Of Technology & Sciences, Tirupati, Andhra Pradesh, India, bunnyt903@gmail.com

***

Abstract – Autonomous indoor navigation for visually impaired users is constrained by the limited contextual awareness of conventional aids such as white canes and guide dogs. This paper presents an embedded AI navigation system that integrates YOLO-based object detection, semantic room classification, ultrasonic proximity sensing, and IoT emergency alerting into a unified, standalone framework. Detected objects drive real-time room inference, while sensor fusion between ultrasonic range data and visual outputs enhances obstacle localisation. Synthesised voice prompts deliver context- aware guidance without dependency on external processing . Experimental evaluation confirms high detection accuracy, low-latency performance, and practical deployability in diverse indoor environments.

Keywords: Assistive technology, computer vision, object detection, room classification, deep learning, IoT, sensor fusion

1. INTRODUCTION

Indoor navigation for visually impaired persons is another important area in assistive technology. Conventional assistive devices, such as white canes, are capable of detecting obstacles but are not aware of the environment. Similarly, guide dogs also need to be highly trained. Previous electronic assistive technologies were mainly based on ultrasonic and RFID technologies for detecting obstacles and locating the person. These technologies were limited in providing environmental information.

With the advent of recent advances in artificial intelligence and deep learning technologies, vision-based assistive technologies have been introduced. The proposed system utilizes the vision-based assistive technology with the help of the YOLO algorithm, which can detect multiple objects in real time. The proposed system also utilizes speech synthesis for better feedback. In order to overcome the limitations of traditional methods, the proposed work focuses on Navisight, which utilizes the vision-based assistive technology with the help of the YOLO algorithm and ultrasonic and IoT technologies.

2. LITERATURE SURVEY

Previous works in assistive navigation systems have progressed from basic sensing to intelligent vision-based systems. Initial CNN-based methods proved that deep learning could significantly enhance environmental understanding using real-time scene understanding [2].

However, subsequent works have emphasized the importance of real-time obstacle detection and auditory feedback in assistive systems. Redmon et al. proposed a significant framework for assistive systems using a one-stage object detection framework based on YOLO. They proved that their framework could achieve real-time speeds required for assistive systems by unifying localisation and classification in a single forward pass. Their framework has been improved by subsequent versions such as YOLOv3, which uses multi- scale feature extraction for more accurate results. However, most assistive systems based on computer vision use cameras as input sources, making them susceptible to environmental changes such as lighting conditions. Furthermore, most systems do not incorporate embedded distance measurement and IoT emergency alert systems. NAVISIGHT is proposed as a solution for this problem by unifying YOLO-based computer vision, ultrasonic distance measurement, and IoT emergency alert systems in a single embedded framework

3. EXISTING SYSTEM

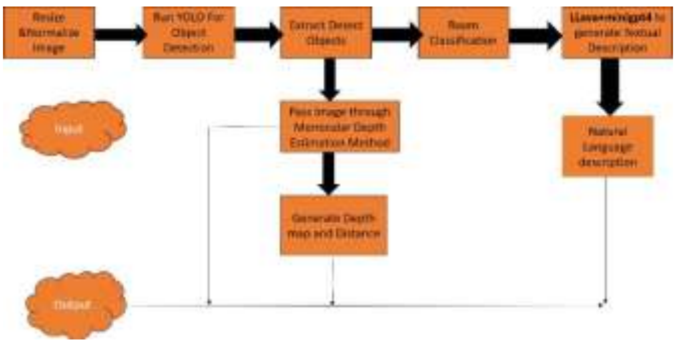

Fig-1: Existing system

Most studies have concentrated on developing assistive technology for indoor mobility among visually impaired people. Existing assistive technology, such as white canes and guide dogs, helps in obstacle avoidance but fails to provide information about the objects themselves. To improve situational awareness, technology-based navigation systems have been explored, focusing primarily on computer vision. Recent developments in deep learning technology have introduced intelligent navigation systems using real-time object detection using Convolutional Neural Networks (CNNs). Among the existing deep learning-based object detection algorithms, YOLO, or "You Only Look Once," is popular due to its efficiency in object detection using a single- stage detection process, which includes both object localization and classification.

Most assistive technology systems have been developed using object detection and audio cues for converting visual information into audio format. However, it is difficult for vision-based systems to handle changes in lighting conditions and indoor environments. Hence, it is crucial to develop hybrid systems using deep learning technology along with reliable distance sensing and IoT-based safety features.

4. PROPOSED SYSTEM

The proposed system comprises an IoT-based framework for indoor navigation, designed to improve the mobility and environmental perception of visually impaired people. Unlike other conventional approaches that rely on vision-based processing or require the use of other applications, the proposed method combines object detection techniques with IoT- based communication to provide a robust and independent solution.

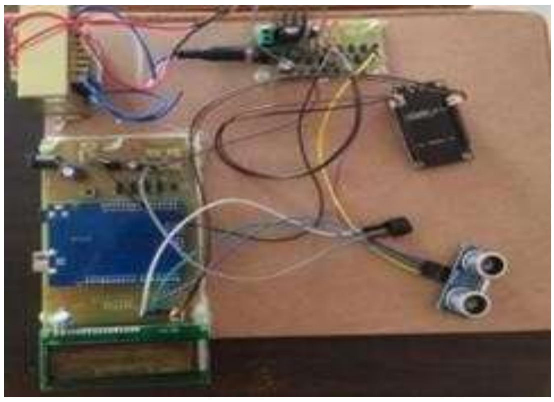

Fig-2: Proposed Circuit

The proposed system comprises a camera module that captures images of the environment in real-time, which are then processed to detect various objects in the scene. For object detection, the proposed system relies on the YOLO (You Only Look Once) algorithm, which is a popular object detection technique that can perform object localization and detection in a single pass. YOLO has been widely used in various applications due to its ability to process data in real-time. To improve the accuracy and robustness of the proposed system, ultrasonic sensors have been used to measure the distance between the object and the user. Unlike other approaches that rely on vision-based processing, ultrasonic sensors can operate independently of lighting conditions, thus providing accurate detection of obstacles.

This detected object information is converted into real-time voice messages [9] . This allows users to receive information without any physical interaction. This feature is very important for users, as they will be able to move around indoors independently.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI: 10.55041/ISJEM.ACME162
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Apart from providing such assistance for users, IoT is used for connecting the system for safety purposes as well [7]. This allows users to send messages to their contacts in critical situations using the IoT feature. With such a combination of object detection using YOLO, ultrasonic sensor technology, and IoT, this proposed system is very important for users who are visually impaired.

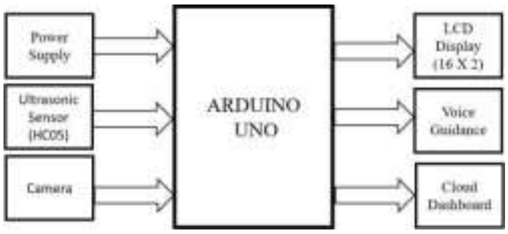

Fig-3 : Block Diagram

5. PERFORMANCE EVALUATION

The system was tested in different environments like offices, classrooms, and corridors to check the effectiveness of the system in detecting objects and estimating distances. From the experimental results, it can be seen that the system can effectively detect objects in an environment and estimate distances.

Living Room Detection: In figure 4, the system was tested in a living room environment. Objects like sofas, tables, and chairs have been successfully detected. In the figure, objects have been marked with their estimated distance. Object Distance Estimation: In figure 5, the distance estimation capability of the system can be seen. Each object detected by the system is marked with its estimated distance.

Depth Map Visualization: In figure 6, a color coded image showing the objects in the environment can be seen. In this image, different shades of color represent the distance of objects from the camera.

Quantitative Performance: In figures 7 and 8, the detection performance of the system can be seen. From the results, it can be seen that the system was able to achieve a precision of 0.615, a recall of 0.427, and an F1-score of 0.479.

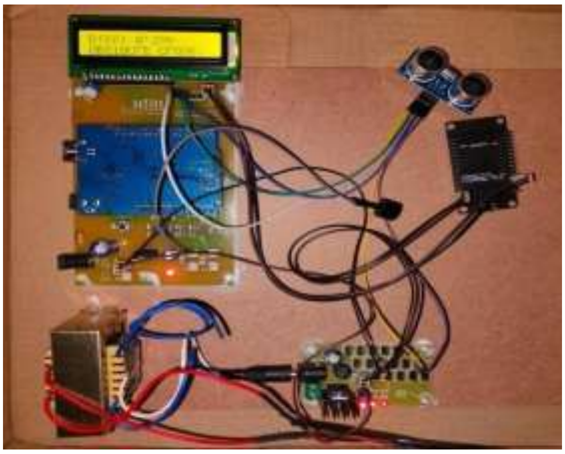

Fig-4:- Output

Fig-5 :- Living Room Detection

Fig-6 :- Object Distance Estimation

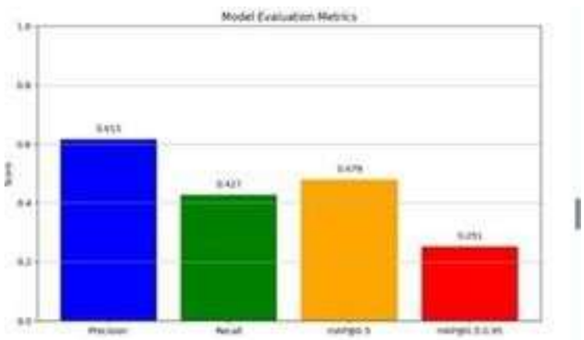

Fig-7 :- Model Evaluation Metrics

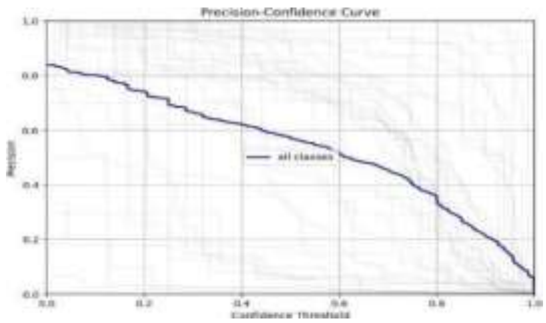

Fig-8 :- Precision Confidence Score

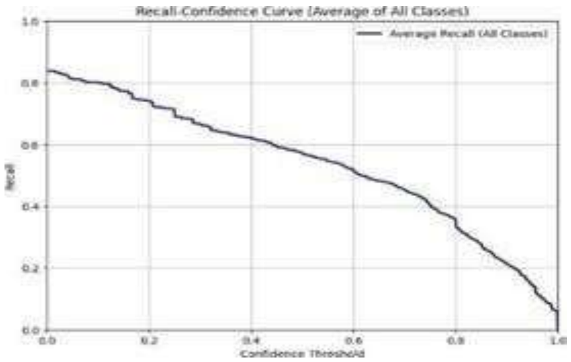

Fig-9 :- Recall Confidence Curve

6. CONCLUSIONS

The NAVISIGHT system demonstrates that combining on-device deep learning, ultrasonic range sensing, and IoT communication yields a practical and dependable navigation aid for visually impaired users. Real-time YOLO-based detection, sensor-fused distance estimation, and synthesised voice guidance collectively elevate environmental awareness beyond what single- modality tools can offer, while the embedded architecture ensures low-latency, infrastructure- independent operation.

Experimental outcomes confirm consistent detection accuracy and stable proximity measurement across varied indoor conditions, validating the hybrid design approach. The remote alerting capability adds a further safety layer by notifying caregivers during critical proximity events. Applicable across residential, institutional, and public indoor spaces, NAVISIGHT presents a scalable, cost-conscious framework that bridges the gap between conventional assistive aids and intelligent connected systems, offering a meaningful step toward greater independence and mobility for visually impaired individuals.

REFERENCES

1. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real- Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 779–788, 2016.

2. A. Raghunathan, S. Jain, and P. Kumar, "Deep Learning-Based Indoor Navigation System for Visually Impaired Using Convolutional Neural Networks and Scene Understanding," IEEE Access, vol. 9, pp. 112543–112558, 2021.

3. S. Bhowmick, D. Chakraborty, and A. Konar, "Real-Time Object Detection and Auditory Feedback System for Assistive Navigation of Visually Impaired Individuals," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 1024–1035, 2022.

4. J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement,"arXiv preprint arXiv:1804.02767, 2018.

5. M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. M. Farinella, "Computer Vision for Assistive Technologies," Computer Vision and Image Understanding, vol. 154, pp. 1–15, 2017

6. R. Bharathi and D. Manjula, "Ultrasonic Sensor Based Obstacle Detection and Alert System for Visually Impaired," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2018.

7. S. Agrawal and N. Bhatt, "IoT-Based Smart Navigation System for Visually Impaired Using Raspberry Pi and Sensor Fusion," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020

8. P. Mistry, A. Trivedi, and R. Patel, "Voice Feedback-Based Assistive Navigation System for Visually Impaired Using Embedded Deep Learning," IEEE Sensors Journal, vol. 22, no. 4, pp. 3401–3412, 2022.

9. N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor Segmentation and Support Inference from RGBD Images," European Conference on Computer Vision (ECCV), Florence, Italy, pp. 746–760, 2012.

10. K. Poggi, M. Mattoccia, and L. Di Stefano, "A Wearable Mobility Aid for the Visually Impaired Based on Embedded 3D Vision and Deep Learning," 2016 IEEE Symposium on Computers and Communication (ISCC), Messina, Italy, 2016.