

Context-Aware Navigation Assistant for the Visually Impaired

Vinaya Kulkarni¹, Sanjesh Pawale², Kiran Yesugade¹,
Anushka Patil¹, Aishwarya Nalawade¹, Riya Kad¹, Hiteshi Kukreja¹

¹ Department of Computer Engineering, Bharati Vidyapeeth's College of Engineering
for Women, Pune, Maharashtra, India

² Vishwakarma University, Pune, Maharashtra, India

Abstract

Visual impairment poses significant challenges to independent mobility by limiting environmental perception and obstacle awareness. Although recent advancements in computer vision have enabled intelligent assistive navigation systems, many existing approaches rely on cloud connectivity, provide excessive auditory feedback, or lack contextual understanding, thereby reducing their effectiveness in real-world environments. This paper proposes a context-aware vision-based navigation framework for assisting visually impaired individuals through intelligent scene understanding and real-time auditory guidance. Unlike conventional object detection systems that announce all detected objects, the proposed framework prioritizes navigation-relevant information by considering spatial proximity, positional relevance, and contextual significance. A hierarchical filtering strategy suppresses redundant detections and delivers concise audio feedback, reducing cognitive load while improving situational awareness. The entirely on-device architecture ensures low-latency operation, enhanced reliability, and user privacy without requiring internet connectivity. The proposed framework demonstrates the potential of context-aware environmental reasoning to provide practical, scalable, and efficient assistive navigation, contributing toward safer and more independent mobility for visually impaired individuals.

Keywords: Assistive Navigation, Visual Impairment, Computer Vision, Real-Time Object Detection, Human-Centered Artificial Intelligence, Environmental Perception, Context-Aware Navigation.

I. INTRODUCTION

Visual impairment significantly limits independent mobility by restricting the ability to perceive surrounding obstacles and environmental changes. Traditional aids such as white canes and guide dogs provide limited awareness, making navigation in unfamiliar or dynamic environments challenging. Recent

advances in computer vision have enabled assistive systems capable of detecting objects and providing auditory feedback. However, many existing solutions rely on cloud connectivity, lack contextual understanding, and generate excessive information, increasing cognitive load for users. Additionally, these systems often treat all detected objects with equal importance, resulting in redundant or irrelevant notifications that may distract users rather than assist them in making timely navigation decisions.

To address these limitations, this paper proposes a context-aware offline navigation framework that enhances mobility assistance through intelligent scene understanding. Unlike conventional systems that report all detected objects, the proposed approach prioritizes navigation-relevant information based on spatial proximity, relative positioning, and contextual importance. A hierarchical filtering mechanism reduces redundant feedback and delivers concise audio guidance, improving situational awareness. Furthermore, the framework incorporates real-time processing capabilities that enable continuous environmental monitoring and adaptive response to dynamic changes in the surroundings.

The system operates entirely on-device, ensuring low latency, reliability, and privacy without requiring internet connectivity. Its modular architecture allows for future integration of additional sensing modalities and advanced reasoning techniques without altering the core navigation strategy. The proposed framework demonstrates an efficient and scalable approach for providing practical assistive navigation, contributing to safer and more independent mobility for visually impaired individuals.

II. RELATED WORK

Recent advancements in artificial intelligence and computer vision have significantly improved the development of assistive navigation systems for visually impaired individuals [9], [12]. Deep learning-based object detection, scene understanding, multimodal perception, and intelligent audio

guidance have enabled the design of systems capable of recognizing environmental obstacles and providing navigation assistance in real time [1], [3], [5]. Despite these developments, existing solutions continue to face challenges related to contextual understanding, computational complexity, cloud dependency, and information overload [10], [16]. This section reviews the existing literature by categorizing previous studies according to their primary research focus.

A. Deep Learning-Based Object Detection

Object detection has become the foundation of modern assistive navigation systems due to the remarkable performance of convolutional neural networks and transformer-based architectures [1], [6], [7]. Yang *et al.* [1] proposed YOLO-RDFA, an enhanced YOLOv8 framework incorporating feature enhancement and attention mechanisms to improve detection accuracy under complex imaging conditions. Similarly, Wazirali and Alkhamash [6] introduced a DETR-based object detection framework capable of improving recognition performance through transformer-based feature extraction. Ashiq *et al.* [7] presented a CNN-based object recognition and tracking system that enables continuous obstacle identification for visually impaired users. While these approaches achieve high detection accuracy, they primarily emphasize object recognition performance without addressing navigation-oriented prioritization of detected objects.

B. Intelligent Navigation and Assistive Guidance

Several researchers have explored intelligent navigation systems that convert visual information into auditory guidance for visually impaired individuals [2], [3], [13], [14], [20]. Yang and Sanie [2] proposed a sight-to-sound human-machine interface that transforms detected objects into audio cues for navigation. Masud *et al.* [3] developed a smart assistive framework integrating object detection and obstacle classification for real-time guidance. Thakur and Pandey [13] demonstrated the effectiveness of object detection-based navigation using deep learning techniques, whereas Nijhawan *et al.* [14] improved navigation accuracy through optimal confidence-score fusion. Jiang *et al.* [20] further explored visual-to-audio recognition systems for real-time auditory assistance. Although these systems provide meaningful assistance, they generally announce multiple detected objects simultaneously, increasing cognitive load during navigation.

C. Distance Estimation and Environmental Perception

Accurate distance estimation is essential for effective obstacle avoidance [4], [11]. Leong and Ramasamy [4] investigated monocular vision-based obstacle detection combined with distance estimation techniques for real-time navigation. Bibi *et al.* [11] proposed an automatic annotation framework incorporating depth estimation to improve environmental understanding. Although these approaches successfully

estimate obstacle distance, many require additional computational resources or specialized depth estimation models, making real-time deployment on lightweight devices more challenging.

D. Multimodal and Context-Aware Assistive Systems

Recent studies have focused on combining multiple sensing modalities to improve situational awareness [5], [8], [10], [19]. Ikram *et al.* [5] developed a multimodal situational knowledge framework that integrates environmental understanding for intelligent alert generation. Elmannai and Elleithy [10] proposed a data fusion framework combining heterogeneous sensory information to improve navigation reliability. Bahrami *et al.* [8] introduced an efficient pathfinding strategy for visually impaired individuals by integrating environmental perception with navigation planning. Zvorișteanu *et al.* [19] explored sensory substitution techniques using 3D audio feedback. Although multimodal systems improve robustness, they frequently depend on multiple sensors, increasing system complexity, cost, and computational requirements.

E. Offline and Portable Navigation Systems

The growing demand for portable assistive technologies has motivated the development of offline navigation frameworks [16], [17], [18]. Sebban *et al.* [16] presented an offline mobile assistance system that performs visual recognition without cloud dependency, thereby improving user privacy and reducing communication latency. Earlier studies by Tapu *et al.* [17] and Mohane and Gode [18] investigated smartphone-based portable navigation systems capable of detecting nearby obstacles in real time. These approaches demonstrate the feasibility of offline assistive technologies; however, they often provide exhaustive descriptions of detected scenes rather than selectively communicating navigation-critical information.

F. Research Gap

The literature indicates substantial progress in object detection, obstacle recognition, distance estimation, multimodal perception, and assistive navigation [1]–[20]. Nevertheless, several important research gaps remain. First, most existing systems prioritize detection accuracy rather than contextual relevance during navigation. Second, many frameworks communicate all detected objects regardless of their navigational significance, resulting in excessive auditory feedback and increased cognitive burden. Third, several approaches rely on cloud-based processing or computationally intensive architectures that limit deployment on portable devices. Finally, relatively few studies integrate contextual scene understanding with intelligent navigation decision-making while maintaining an entirely offline architecture.

To address these limitations, this work proposes a context-

aware offline navigation framework that combines real-time scene understanding with hierarchical object prioritization and intelligent auditory guidance. Instead of reporting every detected object, the proposed framework selectively communicates only navigation-relevant information by considering spatial proximity, positional relevance, and

contextual importance. This strategy aims to reduce cognitive overload while providing efficient, reliable, and real-time navigation assistance for visually impaired individuals.

Table 1: Literature Survey Summary

Authors & Year	Paper Title	Methodology Used	Key Findings	Future Scope
Yang et al. (2024)	YOLO-RDFA: Object Detection in RD Imagery with Improved YOLOv8 Based on Feature Enhancement and Attention Mechanisms	Improved YOLOv8 with feature enhancement and attention mechanisms	Improved detection accuracy and robustness in complex scenes	Extend the framework for real-time assistive navigation applications
Yang & Saniie (2020)	Sight-to-Sound Human-Machine Interface for Guiding and Navigating Visually Impaired People	Vision-to-audio interface using object recognition	Enabled auditory scene interpretation for navigation	Improve contextual reasoning and personalized guidance
Masud et al. (2022)	Smart Assistive System for Visually Impaired People: Obstruction Avoidance Through Object Detection and Classification	Deep learning-based object detection and obstacle classification	Enhanced real-time obstacle avoidance	Incorporate adaptive object prioritization and scene understanding
Leong & Ramasamy (2023)	Obstacle Detection and Distance Estimation for Visually Impaired People	Monocular vision with distance estimation	Improved obstacle localization for navigation	Increase distance estimation accuracy under varying environments
Ikram et al. (2025)	Multimodality-Based Situational Knowledge for Obstacle Detection and Alert Generation	Multimodal data fusion and situational reasoning	Improved obstacle awareness using multiple information sources	Develop lightweight multimodal systems for portable devices
Wazirali &	Enhancing Object Detection in Assistive	Transformer-based object	Improved detection	Optimize computati

Alkhamas h (2016)	Technology for the Visually Impaired: A DETR-Based Approach	detection (DETR)	performance using transformer architecture	onal efficiency for embedded platforms
Ashiq et al. (2022)	CNN-Based Object Recognition and Tracking System to Assist Visually Impaired People	CNN-based object detection and tracking	Accurate real-time object recognition and tracking	Improve tracking robustness in crowded dynamic environments
Bahrami et al. (2025)	Efficient Real-Time Pathfinding for Visually Impaired Individuals	AI-based path planning and navigation	Efficient route planning with obstacle avoidance	Integrate dynamic scene understanding and semantic mapping
Sahu & Tiwari (2021)	Assistive Devices Analysis for Visually Impaired Persons: A Review	Comprehensive literature review	Summarized existing assistive technologies and challenges	Develop intelligent, affordable, and user-friendly navigation systems
Elmannai & Elleithy (2025)	A Highly Accurate and Reliable Data Fusion Framework for Guiding the Visually Impaired	Multi-sensor data fusion	Improved navigation reliability through sensor integration	Reduce hardware complexity and computational cost
Bibi et al. (2025)	An Auto-Annotation Approach for Object Detection and Depth-Based Distance Estimation	Auto-annotation with monocular depth estimation	Enhanced object localization and distance prediction	Improve real-time deployment on low-resource devices
Zafar et al. (2022)	Assistive Devices Analysis for Visually Impaired Persons: A Review on Taxonomy	Taxonomy-based literature review	Categorized assistive technologies and research trends	Explore AI-driven context-aware navigation systems
Thakur & Pandey (2025)	Navigation Using Object Detection for Visually Impaired Individuals	YOLO-based object detection	Demonstrated effective obstacle recognition for	Integrate contextual decision-making and

			navigation	adaptive guidance
Nijhawan et al. (2019)	Real-Time Object Detection for Visually Impaired with Optimal Combination of Scores	Multi-score confidence optimization	Improved detection reliability and precision	Enhance navigation intelligence using contextual information
Foysal et al. (2023)	Advancing AI-Based Assistive Systems for Visually Impaired People	Multi-class object detection with currency recognition	Extended AI assistance beyond navigation	Develop comprehensive daily-life assistive platforms
Sebban et al. (2025)	SeeAround: An Offline Mobile Live Support System for the Visually Impaired	Offline mobile computer vision system	Demonstrated reliable offline assistive navigation	Improve contextual scene understanding and obstacle prioritization
Tapu et al. (2013)	A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People	Smartphone vision-based obstacle detection	Established feasibility of portable navigation assistance	Improve accuracy using modern deep learning techniques
Mohane & Gode (2016)	Object Recognition for Blind People Using Portable Camera	Portable camera-based object recognition	Enabled basic environmental awareness	Incorporate real-time object prioritization and speech guidance
Zvorișteanu et al. (2021)	Sensory Substitution for the Visually Impaired: Sound of Vision System	3D audio sensory substitution	Improved spatial perception through immersive audio	Reduce cognitive load through intelligent audio filtering
Jiang et al.	Let Blind People See: Real-Time Visual Recognition with Results Converted to	Real-time visual recognition with 3D audio	Demonstrated effective visual-to-audio	Improve contextual interpretation and personalization

	3D Audio	conversion	translation	ed navigation assistance
--	----------	------------	-------------	--------------------------------

III. PROPOSED SYSTEM

The proposed navigation system is designed to provide real-time environmental awareness and navigation assistance for visually impaired individuals through intelligent object detection and audio feedback. The system processes live video captured from a monocular camera and identifies navigation-relevant obstacles using deep learning-based object detection models. Unlike conventional assistive systems that announce every detected object, the proposed framework prioritizes only the nearest and most relevant obstacles before generating concise audio instructions. This approach minimizes unnecessary audio notifications and enables users to concentrate on immediate navigation hazards.



Fig. 1: System Architecture

The overall workflow of the proposed system is illustrated in Fig. 1. Initially, the camera continuously captures the surrounding environment. The captured frames are processed using two object detection models, followed by duplicate removal using Non-Maximum Suppression (NMS). The detected objects are then analyzed to estimate their relative distance from the user. Based on the estimated distance and spatial location, a context-aware prioritization module selects the most navigation-relevant object(s). Finally, the selected information is converted into directional Hindi audio guidance for the user.

A. System Architecture

The The proposed system consists of six functional modules: Image Acquisition, Object Detection, Distance Estimation, Context-Aware Object Prioritization, Direction Identification, and Audio Guidance Generation. These modules operate sequentially to process the captured visual information and provide meaningful navigation assistance.

The image acquisition module continuously captures live

video frames using a monocular RGB camera. The object detection module identifies general objects and staircases using separate deep learning models. The detected objects are refined through Non-Maximum Suppression to eliminate duplicate detections. The distance estimation module approximates the relative distance of each detected object using monocular vision. The context-aware prioritization module then filters unnecessary objects and selects only those that are most relevant for safe navigation. Finally, the direction identification and audio guidance modules generate concise voice instructions indicating the relative position of nearby obstacles.

B. Image Acquisition Module

The Image Acquisition Module continuously captures real-time video frames using a monocular RGB camera. Each frame is resized before further processing to ensure efficient computation while preserving sufficient visual information for object detection. The acquired frames serve as the primary input to the navigation framework.

C. Object Detection Module

The Object Detection Module is responsible for recognizing obstacles present in the captured environment. A pre-trained YOLOv8 model is employed for detecting general objects such as pedestrians, vehicles, chairs, bottles, bags, and other commonly encountered obstacles. Since staircases are not consistently detected by the general model, a separately trained staircase detection model is incorporated to improve recognition accuracy. The outputs obtained from both models are merged and refined using Non-Maximum Suppression (NMS), ensuring that duplicate detections are removed while preserving the highest-confidence predictions.

D. Distance Estimation Module

The Distance Estimation Module estimates the relative distance between the user and detected objects using monocular vision. The bounding box height of each detected object is utilized along with predefined object dimensions to approximate its distance from the camera. Objects located beyond the predefined navigation range are excluded from further processing, ensuring that only nearby obstacles are considered during navigation.

E. Context-Aware Object Prioritization Module

The Context-Aware Object Prioritization Module forms the core of the proposed system. Instead of announcing every detected object, the module intelligently selects only those obstacles that are most relevant to the user's immediate movement. Initially, the nearest detected object is identified using the estimated distance. Additional objects are considered only when they lie within a similar spatial range and may simultaneously affect the user's navigation. Preference is given to objects located directly ahead of the user, followed by objects positioned on the left or right. This prioritization strategy reduces redundant audio feedback while improving navigation efficiency.

F. Direction Identification Module

The Direction Identification Module determines the relative position of the selected obstacle by analyzing the center coordinates of its bounding box. The image frame is divided into three navigation regions representing the left, center, and right directions. Based on the region in which the object is located, appropriate directional information is generated for the user.

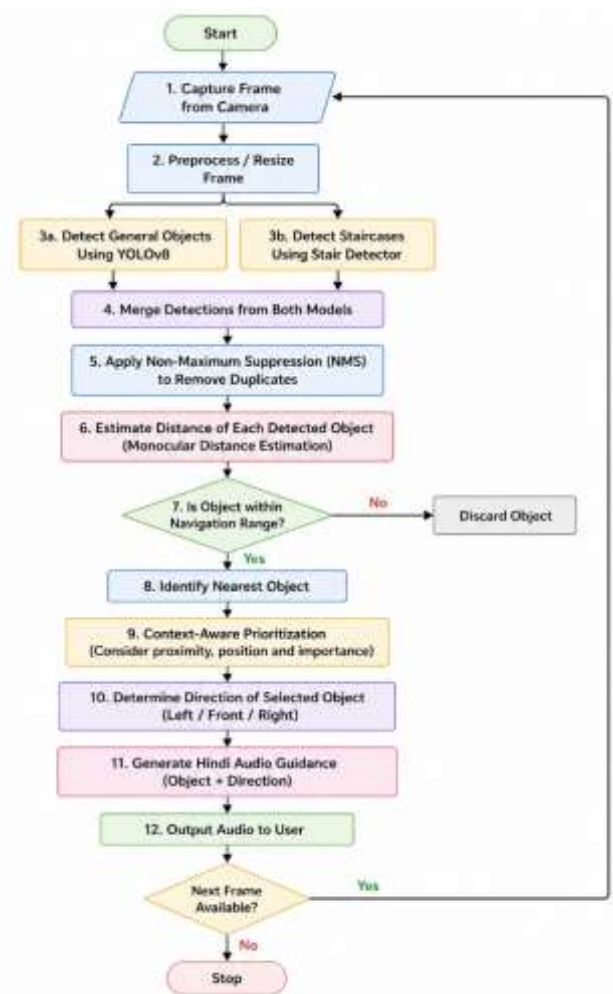
G. Audio Guidance Module

The Audio Guidance Module converts the selected navigation information into concise Hindi speech instructions. Only the obstacle type and its corresponding direction are communicated, enabling users to react quickly without being overwhelmed by excessive auditory information. Since the complete framework operates offline, the generated speech is available without requiring an active internet connection, ensuring reliable real-time navigation assistance.

IV. ALGORITHM

The proposed navigation algorithm processes live video frames to identify navigation-relevant obstacles and generate concise audio guidance for visually impaired users. The algorithm combines general object detection with staircase detection, followed by distance estimation, context-aware prioritization, and direction-based speech generation. The overall workflow of the proposed algorithm is presented below.

Fig. 2: Flow of System



Algorithm 1: Context-Aware Navigation Algorithm.

Input: Live video frame (F)

Output: Hindi audio guidance

- Step 1: Capture a video frame (F) from the camera.
- Step 2: Resize the frame for real-time processing.
- Step 3: Detect general objects using the YOLOv8 object detection model.
- Step 4: Detect staircases using the specialized staircase detection model.
- Step 5: Merge detections obtained from both models.
- Step 6: Apply Non-Maximum Suppression (NMS) to eliminate duplicate detections.
- Step 7: Estimate the relative distance of each detected object.
- Step 8: Discard objects located beyond the predefined navigation range.
- Step 9: Identify the nearest detected object.
- Step 10: Select additional objects only if they lie within the defined distance margin of the nearest object.
- Step 11: Determine the relative position (Left, Front, or Right) of each selected object.
- Step 12: Prioritize the selected objects according to navigation relevance.
- Step 13: Generate concise Hindi audio instructions.

Step 14: Repeat the above steps for each captured frame until the navigation session is terminated.

V. CONCLUSION

This paper presented an offline vision-based navigation assistance system designed to improve the independent mobility of visually impaired individuals through real-time environmental perception and intelligent audio guidance. The proposed framework integrates a general object detection model with a specialized staircase detection model to accurately identify navigation-relevant obstacles. A context-aware object prioritization strategy was introduced to filter unnecessary detections and communicate only the nearest and most significant obstacles, thereby reducing cognitive overload and improving the effectiveness of audio feedback. Furthermore, the system estimates the relative position of detected objects and generates concise Hindi voice instructions to assist users in navigating safely within their surroundings. Since the entire framework operates offline, it eliminates dependence on internet connectivity while ensuring reliable and low-latency performance. The proposed system provides a practical and scalable solution for real-time assistive navigation. Future work will focus on improving distance estimation accuracy, integrating depth-sensing techniques, enhancing dynamic obstacle recognition, and expanding multilingual voice support to further improve navigation efficiency and user experience.

VI. REFERENCES

- [1] M. Yang, M. Dong, C. Li, and F. Nie, "YOLO-RDFA: Object detection in RD imagery with improved YOLOv8 based on feature enhancement and attention mechanisms," *IEEE Access*, vol. 12, pp. 158226–158238, 2024, doi: 10.1109/ACCESS.2024.3485499.
- [2] G. Yang and J. Saniie, "Sight-to-sound human-machine interface for guiding and navigating visually impaired people," *IEEE Access*, vol. 8, pp. 26612–26622, 2020, doi: 10.1109/ACCESS.2020.2970979.
- [3] U. Masud, T. Saeed, H. M. Malaikah, F. ul Islam, and G. Abbas, "Smart assistive system for visually impaired people: Obstruction avoidance through object detection and classification," *IEEE Access*, vol. 10, pp. 13428–13440, 2022, doi: 10.1109/ACCESS.2022.3146320.
- [4] X. Leong and R. K. Ramasamy, "Obstacle detection and distance estimation for visually impaired people," *IEEE Access*, vol. 11, pp. 136609–136629, 2023, doi: 10.1109/ACCESS.2023.3338154.
- [5] S. Ikram, I. S. Bajwa, S. Gyawali, A. Ikram, and N. Alsubaie, "Multimodality-based situational knowledge for obstacle detection and alert generation to enhance navigation assistive systems," *IEEE Access*, vol. 13, pp. 1–15, 2025, doi: 10.1109/ACCESS.2025.3367421.
- [6] R. Wazirali and E. Alkhamash, "Enhancing object detection in assistive technology for the visually impaired: A DETR-based approach," *IEEE Access*, vol. 4, pp. 1–10, 2016, doi: 10.1109/ACCESS.2016.2586308.
- [7] F. Ashiq, M. Asif, M. Bin Ahmad, S. Zafar, K. Masood, T. Mahmood, M. T. Mahmood, and I. H. Lee, "CNN-based object recognition and tracking system to assist visually impaired people," *IEEE Access*, vol. 10, pp. 14819–14834, 2022, doi: 10.1109/ACCESS.2022.3148036.
- [8] A. Bahrami, M. Yazdani, and A. Fotouhi, "Efficient real-time pathfinding for visually impaired individuals," *IEEE Access*, vol. 13, pp. 1–12, 2025, doi: 10.1109/ACCESS.2025.3351987.
- [9] P. K. Sahu and R. Tiwari, "Assistive devices analysis for visually impaired persons: A review," *Int. J. Sci. Technol. Res.*, vol. 9, 2021.
- [10] W. M. Elmannai and K. M. Elleithy, "A highly accurate and reliable data fusion framework for guiding the visually impaired," *IEEE Access*, vol. 13, pp. 1–15, 2025, doi: 10.1109/ACCESS.2025.3365129.
- [11] M. Bibi, M. Faseeh, A. N. Khan, A. Rizwan, Q. W. Khan, R. Ahmad, and D.-H. Kim, "An auto-annotation approach for object detection and depth-based distance estimation in security and surveillance systems," *IEEE Access*, vol. 13, pp. 1–15, 2025, doi: 10.1109/ACCESS.2025.3349782.
- [12] S. Zafar, M. Asif, M. Bin Ahmad, T. M. Ghazal, T. Faiz, M. Ahmad, and M. Adnan Khan, "Assistive devices analysis for visually impaired persons: A review on taxonomy," *IEEE Access*, vol. 10, pp. 1–12, 2022, doi: 10.1109/ACCESS.2022.3157890.
- [13] V. Thakur and S. Pandey, "Navigation using object detection for visually impaired individuals," in *Proc. INDIACom*, 12th Int. Conf. Computing for Sustainable Global Development, Apr. 2025, doi: 10.23919/INDIACom66777.2025.11115651.
- [14] S. S. Nijhawan, S. Bhardwaj, A. Kumar, and G. Nijhawan, "Real-time object detection for visually impaired with optimal combination of scores," in *Proc. INDIACom*, Int. Conf. Computing for Sustainable Global Development, New Delhi, India, Mar. 2019, pp. 215–219.
- [15] N. U. Foysal, T. J. Thamir, M. S. Uddin, and M. F. Islam, "Advancing AI-based assistive systems for visually impaired people: Multi-class object detection and currency classification," in *Proc. ICICT4SD*, 2023, doi: 10.1109/ICICT4SD59951.2023.10303628.
- [16] O. Sebban, A. Azough, and M. Lamrini, "SeeAround: An offline mobile live support system for the visually impaired," *Bull. Electr. Eng. Informatics*, vol. 14, no. 1, 2025, doi: 10.11591/eei.v14i1.7904.
- [17] R. Tapu, B. Mocanu, A. Bursuc, and T. Zaharia, "A smartphone-based obstacle detection and classification system for assisting visually impaired people," in *Proc.*

IEEE ICCV Workshops, 2013, pp. 444–451, doi: 10.1109/ICCVW.2013.65.

[18] V. Mohane and C. Gode, “Object recognition for blind people using portable camera,” in *Proc. World Conf. Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, Feb.–Mar. 2016.

[19] O. Zvorișteanu, S. Caraiman, R.-G. Lupu, N. A. Botezatu, and A. Burlacu, “Sensory substitution for the visually impaired: A study on the usability of the sound of vision system in outdoor environments,” *Electronics*, vol. 10, no. 14, p. 1619, 2021, doi: 10.3390/electronics10141619.

[20] R. Jiang, S. Qu, and Q. Lin, “Let blind people see: Real-time visual recognition with results converted to 3D audio.”