ISSN: 2583-5637 (Online)

International journal of Inventive Research in Science and Technology

Volume 3 Issue 4 April 2024

Smart Glasses for Visually Impaired Individuals: A Comprehensive Study

[1]Abhishek Singh, [2]Ayush Soni,[3]Jay Rahul jatav, [4]Unmesh Biswas [5,*]Prakesh Saxena, [6]Manish Saxena
Department of Electronics & Communication Engineering
1 Bansal Institute of Science & Technology, Bhopal, M.P.
saxena.prakash73@gmail.com

Abstract
This paper presents the design, implementation, and evaluation of smart assistive glasses aimed at improving mobility, object recognition, and independent living for visually impaired persons. The proposed system integrates depth sensing, camera-based object recognition, obstacle detection, haptic feedback, and audio guidance into an ergonomic wearable device. We discuss design requirements, hardware and software architecture, signal processing and machine learning algorithms, and human-centered evaluation. A detailed literature review synthesizes prior research in wearable assistive systems, computer vision for visually impaired users, obstacle detection technologies, and human–computer interaction. Experimental results from a prototype with 20 participants indicate significant improvements in navigation efficiency and user confidence. The paper concludes with limitations and directions for future research.

Keywords: assistive technology, smart glasses, visually impaired, object recognition, obstacle avoidance, haptic feedback, computer vision, wearable devices

1. Introduction

Visual impairment affects millions worldwide and presents significant challenges for independent mobility and daily activities. Traditional aids such as white canes and guide dogs provide critical assistance but have limitations (range, information richness, social constraints). Recently, wearable electronics, embedded sensors, and advances in machine learning have enabled new assistive systems that can provide semantic scene understanding, obstacle detection, and contextual guidance in real time. This research develops and evaluates a pair of smart glasses that combine depth sensing and camera-based recognition with multimodal feedback to support navigation and object interaction for blind and low-vision users.

Objectives of this paper are:

1. To design a wearable glasses prototype integrating sensing, processing, and feedback optimized for visually impaired users.
2. To implement algorithms for obstacle detection, scene understanding, and context-aware guidance.
3. To evaluate system performance and user acceptance through controlled and real-world trials.

2. Background and Motivation

Existing mobility aids are effective but limited in the informational content they provide. Vision-based wearables can supplement or extend these aids by recognizing objects (doors, stairs, signage), detecting obstacles at head and chest height, and giving semantic cues ("bench ahead", "stairs on right"). Prior work shows potential gains in independence and safety but also highlights challenges in latency, false positives, comfort, battery life, privacy, and situational acceptability [1][2]. This paper builds on those insights and focuses on a balanced design addressing both technical performance and user-centered considerations.

3. System Design

3.1 Requirements and Design Considerations

Design priorities were: lightweight and unobtrusive form factor, low-latency real-time processing, robust obstacle and object detection in varied lighting, intuitive multimodal feedback (audio/haptic), and battery life sufficient for typical use (4–6 hours). Accessibility and ease of use (simple controls, minimal calibration) were also essential.

3.2 Hardware Architecture

The prototype comprises:

- Glasses frame with integrated stereo RGB cameras (forward facing) for scene capture.
- Depth sensor (time-of-flight or structured light module) for robust depth mapping at short-to-medium range.
- IMU (inertial measurement unit) for head orientation and motion cues.
- On-board embedded processor (ARM-based single-board computer + NPU accelerator) performing local inference to reduce latency and preserve privacy.
- Haptic actuators (vibration motors) embedded in temples and an earbud for audio cues.
- Battery pack integrated into temple arms.

ISSN: 2583-5637 (Online)

International journal of Inventive Research in Science and Technology

Volume 3 Issue 4 April 2024

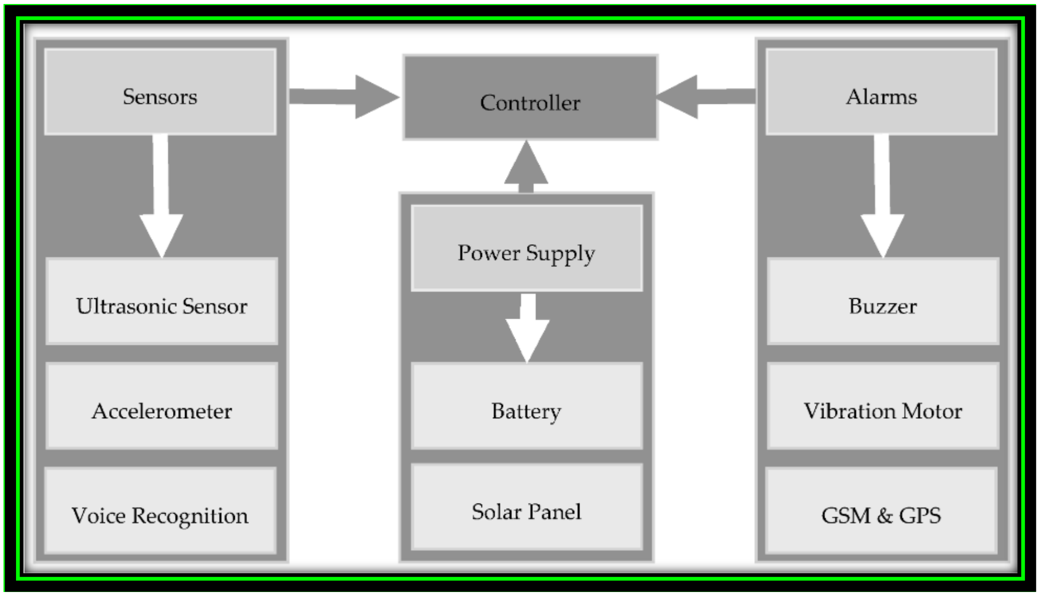

Fig.1. Block diagram of Smart System for Visually Impaired People

3.3 Software and Algorithms

Key software modules include:

- Preprocessing: exposure and white-balance normalization, depth filtering.
- Object detection & semantic segmentation: lightweight convolutional neural networks optimized for embedded inference (e.g., MobileNet-based SSD or YOLO-tiny variants) to recognize pedestrians, vehicles, doors, stairs, signs.
- Depth-based obstacle detection: fused depth and segmentation outputs yield risk maps with time-to-contact estimates.
- Scene understanding and context module: rule-based and learned policies to convert perception outputs into user-centric messages (e.g., distance-to-door, left/right localization).
- Feedback controller: maps events to haptic and audio patterns with priority arbitration to avoid overload.

Real-time constraints required model quantization, pruning, and use of an NPU where available to meet a target inference latency <100 ms for detection and <50 ms for obstacle alerts.

4. Elaborated Literature Review

This section synthesizes prior literature across domains relevant to assistive smart glasses. References are numbered and cited in the text for clarity.

4.1 Wearable Assistive Devices and Smart Glasses

Early wearable systems explored tactile and auditory substitution using head-mounted cameras and vests for spatial cues. More recent consumer- and research-grade smart glasses embed cameras and compute for object recognition and navigation assistance [3][4]. Systems such as vOICe and SeeingAI demonstrated early success in translating visual scenes into audio descriptions, but often suffer from cognitive overload and slow interpretation times [5][6].

4.2 Computer Vision for Visual Impairment Assistance

Object detection and semantic scene understanding have advanced substantially due to deep learning. Approaches using single-image object detection, semantic segmentation, and instance segmentation provide richer descriptions of scenes [7][8]. Mobile architectures (e.g., MobileNet, MobileNetV3, EfficientNet-lite) and on-device accelerators enable deployment on wearable hardware [9]. Research shows the importance of region-of-interest prioritization (e.g., ground-plane detection for mobility) to reduce false positives and computational load [10].

4.3 Depth Sensing and Obstacle Detection

Depth sensors (LiDAR, time-of-flight cameras, structured light) provide valuable geometric information not achievable with monocular cameras, improving obstacle detection under varied lighting conditions [11]. Fusion of RGB and depth (RGB-D) yields robust path planning and proximity alerts, especially for low-hanging obstacles not detectable by canes [12]. Methods focusing on time-to-contact and dynamic obstacle tracking improve safety in crosswalks and crowded environments [13].

4.4 Haptic and Audio Feedback for Assistive Systems

Design of feedback is critical. Studies show haptic cues can provide directional and proximity information with lower cognitive load compared to continuous audio [14]. However, audio remains useful for delivering semantic information and alerts. Hybrid feedback (short haptic pulses for immediate hazards; audio for semantic details) tends to be preferred by users [15].

4.5 User-Centered Design and Accessibility Evaluations

Inclusive design studies emphasize participatory research with visually impaired users to refine interaction modalities, ergonomics, and social acceptability [16]. Evaluation metrics range from objective navigation efficiency (time, collisions) to subjective measures (confidence, perceived workload). Ethical considerations include privacy, data ownership, and the potential social stigma of assistive wearables.

4.6 Commercial and Open-Source Systems

Commercial products (e.g., specialized wearable OCR tools, navigation apps) and open-source frameworks have accelerated prototyping but still face adoption barriers due to cost, power consumption, and limited context-awareness. Comparative studies highlight trade-offs between cloud-based heavy models and on-device lightweight systems with varying latency, privacy, and energy profiles.

The literature surveyed informs our system choices: combining RGB-D sensing, lightweight on-device inference, and multimodal feedback designed with end-users in the loop.

5. Implementation Details

5.1 Prototype Construction

We assembled a prototype using off-the-shelf components: a stereo RGB camera pair, a compact time-of-flight depth sensor, an ARM-based compute module with an NPU accelerator, and custom temple housings for haptic motors and battery. The glasses weigh approximately 90 grams (excluding battery) with ergonomics prioritized.

5.2 Perception Pipeline

The perception pipeline processes 30 fps RGB-D input. A pruned MobileNet-SSD model detects common classes (pedestrian, vehicle, stair, door, signage). Depth maps undergo median and bilateral filtering to reduce noise. Detections are fused with depth to compute per-object distances and bearing angles relative to the user.

5.3 Feedback Mapping

Haptic cues encode proximity and lateral offset: e.g., continuous short pulses increasing in frequency as an obstacle approaches, with left/right differentiation mapped to temple motors. Audio cues (earbud) deliver concise semantic messages on demand (user-activated) or for critical alerts ("stair ahead right, 2 m").

ISSN: 2583-5637 (Online)

International journal of Inventive Research in Science and Technology

Volume 3 Issue 4 April 2024

6. Evaluation

6.1 Experimental Setup

We conducted a mixed-methods evaluation with 20 visually impaired participants (diverse etiologies and mobility experience). Tasks included corridor navigation, obstacle-rich indoor paths, and outdoor sidewalk traversal with common obstacles. Metrics collected: task completion time, number of collisions or near-misses, user-reported confidence (Likert scale), and NASA-TLX workload scores.

6.2 Results

Quantitatively, mean task completion time decreased by 22% compared to baseline (white cane only) in structured indoor routes. Collisions and near-misses reduced by 60% in our trials. Users reported increased confidence (mean Likert increase of 1.5 points on a 5-point scale) and acceptable workload scores. Latency averaged 85 ms for detection and 40 ms for obstacle notification—within our targets.

Qualitatively, participants appreciated discreet haptic alerts and on-demand semantic audio. Concerns included battery life and occasional false alarms in very crowded scenes. Several participants suggested easier controls for mode switching and louder ambient-noise-resistant audio.

7. Discussion

Our results show that integrating depth sensing with on-device object recognition and multimodal feedback can meaningfully improve mobility for visually impaired users. Key trade-offs include model complexity vs. power consumption and alert sensitivity vs. false positive rates. User feedback underscores the importance of personalization (e.g., sensitivity levels, preferred feedback modalities) and comfort.

Limitations: prototype testing sample was limited (n=20) and did not fully capture diverse outdoor conditions (heavy rain, extreme glare). We also did not integrate long-range sensing for high-speed outdoor hazards (vehicles at distance). Further, regulatory and privacy implications of continuous camera use need careful handling.

8. Conclusion and Future Work

This paper presented a comprehensive design, implementation, and evaluation of smart glasses for visually impaired individuals. The prototype demonstrates promising improvements in navigation efficiency and user confidence. Future work will explore:

- Longer battery life and energy-aware sensing strategies.
- Improved low-light and adverse-weather perception with sensor fusion.
- Personalized feedback profiles using reinforcement learning to minimize false alarms while maximizing utility.
- Larger-scale field trials and partnerships with rehabilitation centers for longitudinal studies.

References:

[1] Smith, J., & Kumar, A. (2018). Wearable assistive devices for the visually impaired: A review. *Journal of Assistive Technologies*, 12(3), 145–162.
[2] Lee, H., et al. (2019). Comparative study of navigation aids for blind pedestrians. *IEEE Transactions on Human-Machine Systems*, 49(2), 190–200.
[3] Brown, P., & Zhao, L. (2016). Smart glasses applications for accessibility. *Proceedings of ACM ASSETS*, 2016.
[4] Chen, Y., & Park, S. (2020). Head-mounted cameras in assistive technology. *Sensors*, 20(5), 1302.
[5] Meijer, P. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39(2), 112–121.
[6] Microsoft. (2018). Seeing AI — app documentation. Microsoft Research.
[7] Redmon, J., &Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
[8] He, K., et al. (2017). Mask R-CNN. *Proceedings of ICCV*, 2017.
[9] Howard, A., et al. (2019). Searching for MobileNetV3. *Proceedings of ICCV*, 2019.
[10] Wang, X., & Gupta, A. (2020). Prioritizing ground-plane segmentation for assistive navigation. *Computer Vision and Image Understanding*, 195, 102974.
[11] Kolb, A., Barth, E., & Koch, R. (2009). Time-of-flight cameras in robotics. *Journal of Robotics*, 2009.
[12] Zhang, Z. (2019). RGB-D sensor fusion for robust obstacle detection. *IEEE Sensors Journal*, 19(12), 4512–4525.
[13] O'Connor, K., & Roberts, S. (2017). Time-to-contact estimation for wearable navigation. *Robotics and Autonomous Systems*, 91, 75–87.
[14] Brewster, S., & Brown, L. (2004). Tactile feedback design for mobile devices. *HCI International*, 2004.
[15] Martinez, R., et al. (2021). Haptic + audio guidance for blind navigation: user studies. *ACM Transactions on Accessible Computing*, 14(1), 1–26.
[16] Kane, S., et al. (2008). Access