IFERP
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 12, Issue 04, April 2025

Comprehensive Survey of Assistive Technologies and Applications Designed to Enhance the Mobility and Independence of Visually Impaired Individuals

[1] Amit Kulkarni, [2] Atharva Pawar, [3] Darshan Gawade, [4] Abhishek Bhakare, [5] Dr. Surendra Mahajan

[1] [2] [3] [4] [5] Department of Information Technology, PVGCOET & GKPIM, Pune, India
Emails ID: [1] kulkarniap123@gmail.com, [2] atharvapawar1883@gmail.com, [3] darshansgawade25@gmail.com, [4] bhakare.abhi9@gmail.com, [5] sam_comp@pvgcoet.ac.in

Abstract— **This paper presents a comprehensive survey of assistive technologies & applications developed to support visually impaired individuals, emphasizing their needs, preferences, and the effectiveness of current solutions. The survey highlights various assistive technologies designed for activities such as navigation, object identification, and environmental awareness. Findings from the analysis suggest that while advancements in sensory substitution devices, navigation aids, and image recognition applications have been made, significant gaps remain in addressing specific challenges faced by visually impaired people, particularly in outdoor environments. The survey also identifies areas for improvement in the accessibility and user-friendliness of these solutions, as well as opportunities for future research to bridge these gaps, ultimately aiming to enhance the autonomy and quality of life for visually impaired individuals.**

Index Terms— visual impaired people, blind assistance, computer vision, mobile application, sight.

I. INTRODUCTION

Visual impairment affected over 2.2 billion individuals worldwide as of 2023 [1], significantly impacting their independence in daily activities such as reading, navigation, shopping, and social interactions. These limitations are more common among older adults, who often experience vision loss due to conditions such as age-related macular degeneration, cataracts, and diabetic retinopathy [2]. Advances in assistive and inclusive technologies offer visually impaired people (VIP) greater independence and social integration.

However, a gap remains between the needs of visually impaired users and the functionalities of many assistive solutions currently available, leading many users to avoid or minimize their use. While various assistive technologies now support tasks like currency differentiation and clothing selection, challenges persist in areas such as real-time navigation and spatial awareness, especially in outdoor contexts. This paper explores the current landscape of assistive technologies for VIP, with a focus on identifying the limitations of existing solutions and uncovering insights to guide the design of more effective assistive tools.

II. LITERATURE SURVEY

A survey was conducted with 20 blind individuals living in an urban area in Malaysia to understand the needs and challenges faced by blind people in outdoor navigation, providing insights for the development of a practical navigation system [3]. The survey explored four primary areas: occupation details, medical history related to vision, past experiences with outdoor navigation, and desired features for an ideal navigation system. The data revealed that half of the participants held low-wage jobs, while the remainder relied on government support, with most participants having an income of RM350 or less. Regarding vision, 25% of respondents had very poor vision, while the rest were completely blind, with most cases stemming from early childhood. In terms of navigation habits, 65% went outdoors more than twice a week but generally avoided unfamiliar areas, primarily relying on traditional white canes, and facing significant challenges in road-crossing and obstacle detection.

The survey results also highlighted specific preferences for an ideal navigation aid. Participants expressed a strong desire for a portable device capable of obstacle detection, location and direction guidance, and road-crossing assistance, while also wishing to retain the cane as a symbol of blindness. They preferred a solution that integrates with their cane, preserving both independence and identity. Additional suggestions from participants included increased government support, enhanced public transportation accessibility, and raising awareness to improve societal support for the blind community. Overall, the survey underscores the demand for an affordable, practical outdoor navigation system that addresses the unique needs of blind users, with key features like portability, obstacle detection, and road-crossing guidance.

An investigation into the navigational experiences and strategies of 14 legally blind individuals in the Bangalore metropolitan area of India highlights the challenges faced in a low-to-middle-income country (LMIC) setting [4]. Participants in the study reported facing numerous difficulties in the urban environment, including uneven terrain, debris, and heavy traffic. They also encountered a

ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 12, Issue 04, April 2025

lack of pedestrian crossings and infrastructure that could support safe navigation for the blind. To cope, participants often relied on techniques like following the edge of paths and roads, known as the "shoreline technique," and occasionally sought assistance from bus drivers or passersby. However, they sometimes felt disempowered when others were unresponsive to their requests for help. Major safety concerns included crossing busy highways without designated crossings and navigating public transportation, where difficulties in knowing when to disembark and a lack of accessibility features on older buses added to the complexity. Environmental factors like heavy rain and darkness further complicated their navigation, leading some participants to adopt group strategies such as interlocking arms.

The study's findings underscore how infrastructure limitations and environmental challenges in low-to-middle-income country (LMIC) contexts affect the mobility and safety of blind individuals. These insights reveal potential opportunities for designing more suitable assistive technologies and services that address the specific challenges within these settings. By engaging with local stakeholders and developing affordable, practical solutions, it may be possible to enhance the independence and safety of blind users in these environments.

A survey conducted with 49 people with visual impairments (PVI) in Ireland explored specific needs and challenges related to outdoor navigation [5]. The study examined different aspects of the navigation process, including journey preparation, street crossing, and environmental information access. Key insights from the results include the fact that 42.9% of participants rely on sighted guides, which suggests that current assistive technologies are not fully meeting their needs for independent mobility. Additionally, 67.3% of respondents reported difficulties in accessing timely information about temporary changes to roads or pavements, a lack that affects their confidence, safety, and independence.

Participants emphasized the need for specific environmental details before and during journeys, such as information about pavements, public transport, and street crossings. Knowing whether intersections have audio-enabled traffic lights was especially important to participants. Additionally, obstacles like uneven pavements, poles, and overhanging branches were reported as frequent hazards, with over 63% of people with visual impairments (PVI) reporting injuries from these obstacles, underscoring the need for enhanced obstacle detection in assistive systems. To address these issues, the authors suggest a four-phase solution to improve outdoor navigation, which includes improved journey planning, real-time awareness, obstacle detection, and user feedback. These findings highlight crucial gaps in existing navigation technologies and emphasize the need for a more user-centered approach to support the mobility and independence of people with visual

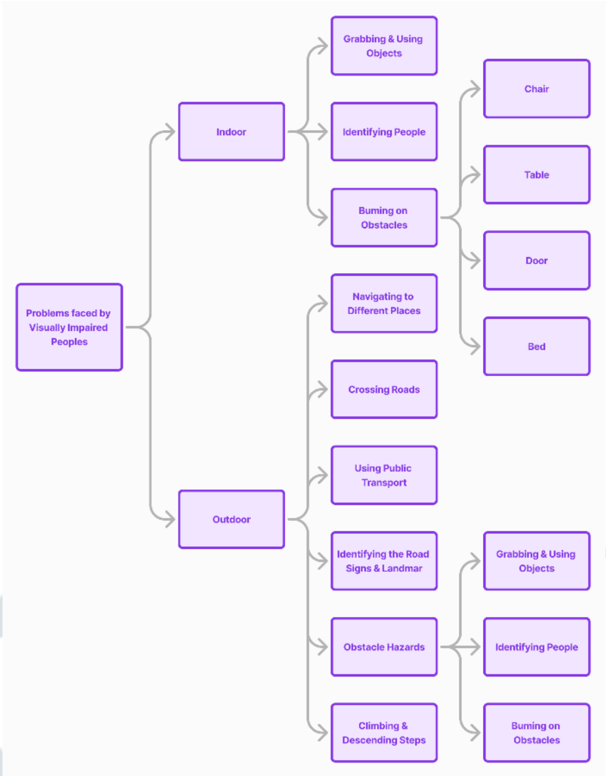

Figure 1. Various problems faced by Visually Impaired People (VIP) in day-to-day life.

impairments (PVI).

A smartphone-based mobility assistant application has been developed to enhance spatial and cognitive awareness in navigation for visually impaired and blind (VIB) individuals [6]. Using a Samsung Galaxy A80 with a 3D depth camera, the system generates depth maps and detects obstacles. It leverages the ARCore Depth Lab API to calculate distances from the depth map and alerts users through audio and vibration feedback about nearby obstacles, with a threshold of 1.6 meters identified as optimal for obstacle detection. In addition to obstacle detection, the application includes an object recognition module that employs a custom TensorFlow Lite model to recognize over 90 everyday object classes relevant to visually impaired and blind (VIB) users, such as walking canes, bags, and signage. A unique feature allows users to interact with detected objects via audio and vibration feedback.

The interface incorporates accessibility features like voice commands and gestures for an inclusive experience. Usability testing with five blindfolded participants showed that the system was comfortable to use, with particular benefits from the audio alerts and object detection. Despite limitations, such as difficulty identifying specific object types, this smartphone-based solution offers a portable, flexible mobility assistant that combines obstacle avoidance and object recognition, with potential for future expansion as smartphone technology progresses.

A system is designed to assist visually impaired individuals by leveraging object detection and computer vision on a smartphone [7]. Its main functions include object detection, distance estimation, direction determination, and audio feedback. The system utilizes TensorFlow's object detection API [8] to classify and locate objects in the camera's view, with a focus on identifying potentially hazardous objects. It provides the user with the object's name and the coordinates of its bounding box, which are then used to estimate the distance of the object from the user based on the camera's focal length. This spatial information helps users navigate their surroundings more safely.

For direction determination, the camera's view is divided into three segments: left, centre, and right, to indicate the relative position of detected objects. The audio output module then combines details about the object's name, distance, and direction into a spoken message delivered through speakers or headphones, providing a clear understanding of nearby objects. Designed to be accessible and cost-effective, the system operates solely on a smartphone without requiring extra hardware. Future improvements could include text-to-speech, food detection, facial recognition, and currency recognition, further supporting the independence and mobility of visually impaired users.

SIGHT, a mobile application designed to assist visually impaired individuals with real-time object recognition and navigation [9]. Using advanced computer vision techniques, specifically YOLO's (You Only Look Once) [10] lightweight version YOLOv8n for object detection and MiDaS (Mixed Data Sampling) for depth estimation, SIGHT provides users with valuable spatial information. YOLOv8n enables fast and accurate object recognition, while MiDaS adds depth perception from single camera images, allowing users to gauge object distances and spatial relationships for safer navigation. Together, these models offer a powerful tool that not only identifies objects but also indicates their relative locations.

This application enhances accessibility through voice interaction, allowing users to issue voice commands like "What objects are in front of me?" and receive auditory responses, minimizing the need for tactile interaction. Extensive performance evaluations show that SIGHT performs effectively on various smartphones, including low-end devices, by using optimizations such as GPU (Graphics Processing Unit) acceleration and model quantization to ensure smooth, real-time operation. Compared to other solutions, SIGHT combines object detection, depth estimation, and an intuitive voice interface, offering visually impaired user greater mobility and independence. The authors highlight SIGHT's potential to make a meaningful impact on the daily lives of this underserved community.

Enhancing monocular depth estimation requires overcoming challenges related to dataset limitations, including lack of diversity, incomplete ground truth, and environment-specific biases. A strategy that combines multiple complementary datasets has been shown to improve generalization in depth estimation models [11]. Key advancements include the development of robust, scale- and shift-invariant loss functions that can handle varying depth representations across datasets, and a principled approach to dataset mixing through multi-objective optimization, as opposed to simple equal-part mixing.

Additionally, a novel dataset of stereo image pairs extracted from 3D movies has been introduced, offering dynamic scenes with dense relative depth information. Experimental results show that models trained using these techniques significantly outperform prior work on various test datasets, even when those datasets were not part of the training (zero-shot cross-dataset transfer). Key findings include the importance of pretraining the model encoder on large-scale auxiliary tasks, the value of the 3D movies dataset, and the superiority of principled dataset mixing over naive strategies. The authors conclude that their work advances monocular depth estimation, highlighting the importance of leveraging diverse training data, and make their models publicly available for further research.

The researchers Ranjana Jadhav, Shreyas Khare developed a real-time object detection system aimed at assisting blind and visually impaired individuals [12]. The system utilizes a mobile camera to capture live images, which are processed by a deep learning model on a laptop to detect and identify objects in the environment. It is based on the SSD (Single Shot MultiBox Detector) architecture [13], which uses a convolutional neural network (CNN) to extract features from input images. The MobileNet model [14] was chosen as the backbone due to its efficiency and low computational requirements. Anchor boxes of different aspect ratios are used to detect objects of various shapes and sizes within each grid cell.

Once objects are detected, the system calculates the approximate distance to them based on the size of the bounding box and converts this information into audio feedback using a text-to-speech library. The system was tested on common household objects such as cups, remotes, and chairs, achieving high accuracy (96 - 99%) in object detection and reasonably accurate distance estimation. This real-time system offers valuable assistance by providing visually impaired users with information about their surroundings. The researchers plan to further improve accessibility by integrating the system into a single mobile application for easier use.

The development of a smart assistant tool, named "Guardian," aims to assist visually impaired individuals in their daily lives [15]. The tool consists of a pair of glasses with an embedded camera that is integrated with a mobile application. The camera captures images of text, objects, and people, which are then processed and recognized using various machine learning models.

The first feature of the tool is text detection and recognition. It uses Tesseract OCR to detect and extract text from images, which is then converted to speech and played back to the user through headphones. The second feature is face and emotion recognition. It uses a convolutional neural network to detect faces and classify the person's emotional state, which is then conveyed to the user through audio output. The third feature is object detection, which employs the YOLOv5 algorithm to identify common everyday objects and provide the user with voice instructions about their location and distance. The authors report that the various models used in the tool have achieved high accuracy, with Tesseract OCR reaching 97%, the emotion recognition model achieving 97%, and the object detection model attaining 92.05% accuracy. The authors conclude that the Guardian tool can significantly benefit a large population of visually impaired individuals by enabling them to perform daily tasks more independently.

A mobile-based navigation system has been developed for visually impaired individuals, utilizing image detection and voice translation technologies [16]. The primary objective of the system is to provide real-time navigation information and familiarize users with their surroundings, with the aim of improving their overall quality of life, independence, and confidence while walking outside.

The system was implemented on an Android platform and uses the device's camera to capture images of the user's environment. Machine learning algorithms are then used to detect obstacles and identify key landmarks, such as doors and stairs. The system also includes a voice translation feature to make the navigation instructions more accessible to users. The researchers conducted testing with blind and visually impaired individuals, who provided positive feedback on the application's performance. The researchers recommend future work to explore different approaches in this field of study and suggest using the technology to create an alternate educational platform for blind people with restricted access to their environment. They also suggest improving the system's accuracy and exploring its potential for use with other mobile operating systems.

An innovative AI-assisted navigation system has been introduced to empower visually impaired individuals [17]. The proposed system leverages the capabilities of the SSD MobileNet-V2 FPNLite object detection model to provide real-time voice-based guidance for indoor navigation. By focusing on speed, accuracy, and resource efficiency, the system aims to enhance the safety and independence of users with visual impairments.

The core of the system is the SSD MobileNet-V2 FPNLite model, which combines the advantages of Faster-RCNN and YOLO to enable rapid and precise detection of objects. The seamless integration of this AI-powered object detection with an Android mobile application allows for intuitive voice-based feedback, informing users about their surroundings. The project's future roadmap includes further

Table I: Summary Table for Assistive Applications

Application	Description
Google Maps	A real-time GPS based app helps visually impaired users navigate with voice-guided directions, explore routes, and find nearby landmarks for improved mobility
Be My Eyes	A mobile app that connects people who are blind or have low vision with sighted volunteers and company representatives for visual assistance
Lazarillo	It provides audio-guided navigation, helping visually impaired users explore their surroundings, locate nearby places, and navigate routes independently.
Blindsquare	It offers audio navigation for visually impaired users, providing detailed information on points of interest and intersections for safe outdoor and indoor travel.
Seeing AI	It narrates the world for visually impaired users by describing people, expressions, objects, text, currency and surroundings through the device's camera.
A Smartphone-Based Mobility Assistant App	A mobility assistant leveraging depth imaging to detect obstacles and recognize objects using Google ARCore, depth camera and COCO SSD MobileNet v2 model. It supports voice commands and gestures.
Android Based Object Detection System for Visually Impaired	An app that detects objects using TensorFlow's Object Detection API and provides audio feedback. It also calculates object distance using OpenCV library.
SIGHT App	SIGHT offers real-time navigation assistance, objects & surrounding information by utilizing YOLOv8n, Midas.
Real-Time Object Detection system for Visually Challenged	A mobile app that uses an external laptop for object detection, provides real time audio feedback. It calculates distance of object based on bounding box coordinates.
Guardian App	A smart assistant tool that recognizes text, objects, and people's faces and emotions, and providing audio output. It uses an external camera mounted on glasses, delegating tasks from mobile device.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 12, Issue 04, April 2025

enhancements, such as the integration of Pedestrian Dead Reckoning (PDR) technology for improved positioning and multi-sensor fusion, as well as the expansion of capabilities to include features like multilingual voice output, two-way communication, and advanced position and angle detection. By continuously improving the system and addressing the unique challenges faced by visually impaired individuals, the researchers aim to revolutionize the lives of those with visual disabilities, fostering greater independence, safety, and overall quality of life.

III. COMPARISON OF ASSISTIVE APPLICATIONS

In this section, the paper compares various assistive applications designed for visually impaired individuals to identify their functionalities, strengths, and limitations. The Table 1 summarizes the list of applications and their descriptions, providing an overview of their primary purposes. Additionally, a feature-based comparison in Table 2 highlights key attributes such as obstacle detection, voice guidance, offline functionality, and user interface design. This analysis reveals gaps in existing solutions and provides insights into the areas where current assistive applications could be improved, highlighting the need for more comprehensive and user-centric approaches.

IV. MATHEMATICAL MODEL

In this section, a mathematical model is described to design an effective assistive application for visually impaired people (VIP). This model formalizes the processes of real-time object detection, depth estimation, and navigation assistance, enabling a structured understanding of how these components interact and contribute to the overall functionality of the system.

A. Environment Representation

The user's environment is represented as a three-dimensional (3D) space $\mathbb{E}$, which contains objects or obstacles that may impact the user's navigation. Mathematically $\mathbb{E}$ is defined as:

$$\mathbb{E} = \{(x, y, z, o) \mid x, y, z \in \mathbb{R}, o \in O_{types}\}$$

where:

- $\mathbb{E}$: The set of all possible objects or obstacles in the environment.
- x, y, z: Spatial coordinates of an object in a Cartesian coordinate system, measured in meters. Each coordinate belongs to $\mathbb{R}$, the set of all real numbers.
- o: The type or class of the object, belonging to a predefined set O_{types} (e.g., "chair," "table," "wall").

Table II: Feature based Comparison of Assistive Applications

	Google Maps	Be My Eyes	Lazarillo	Blind Square	Seeing AI	Smart phone-Based Mobility Assistant App	Android Based Object Detection System for Visually Impaired	SIGHT App	Guardian App
Platform	Android & IOS	Android & IOS	Android & IOS	IOS	Android & IOS	Android	Android	Android	Android
Input Source	GPS	Human Guide	GPS	GPS	Mobile Camera	Mobile Camera	Mobile Camera	Camera & Voice Commands	Images through glass camera
Hardware Component	Mobile	Mobile	Mobile, GPS, other location tech	Mobile, GPS, other location tech	Mobile	Mobile	Mobile	Mobile Camera	Mobile, Glasses with Camera Embedded
Coverage Area	Global Coverage	Depend on user need	Wide Area	Wide Area	Global Coverage	Indoor / Outdoor (1.6m restricted)	Wide Area	Dependent on user's needs	Indoor / Outdoor
Object Detection	No	Yes	No	No	Yes	Yes	Yes	Objects	Yes
Obstacle Detection	No	No	No	No	No	Yes	No	Yes	Yes
Feedback	Audio Feedback	Audio Feedback by Guide	Auditory Feedback, Tactile Feedback	Auditory Feedback, Tactile Feedback	Yes	Yes	Yes	Auditory Feedback, Vibration	Only for objects and facials
Navigation	User Directed	User Directed	GPS Based	GPS Based	No	User Directed	No	Obstacle Avoidance	Yes
Text Reading	No	Yes	No	No	Yes	No	No	No	Yes
Cost	Free	Free	Free	Low	Free	Free	Free	Free	High

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 12, Issue 04, April 2025

B. Input Data Representation

The chest-mounted camera continuously captures frames in real time:

$$I(t) = \{f_1, f_2, \dots, f_n\}$$

where:

- $I(t)$: Sequence of frames captured at time t.
- f_i: A single frame of dimension $H \times W$ pixels. Each pixel is combination of RGB (Red, Green, Blue) values.

C. Object Detection and Classification:

For each frame f_i, the system identifies objects and their bounding boxes:

$$O(t) = \{(o_i, b_i, c_i) \mid o_i \in O_{types},$$
$$bi\ is\ the\ bounding\ box,\ ci\ is\ the\ confidence\ score\}$$

where:

- o_i: Detected object class.
- $b_i = (x_{min}, y_{min}, x_{max}, y_{max})$: Coordinates of the bounding box.
- c_i: Confidence score, with $c_i \in [0, 1]$.

D. Depth Estimation

Depth estimation generates a depth map for each frame f_i, representing the depth of each pixel:

$$M_d = depth(f_i)$$

where:

- $M_d(x, y)$: Depth at pixel location (x, y).
- $depth$: Depth estimation function applied to frame f_i.

For each detected object o_i with bounding box b_i, the depth is calculated as the average depth of the pixels within the bounding box:

$$d_i = \frac{1}{N_i} \sum_{(x,y)\in b_i} M_d(x, y)$$

where:

d_i: Depth of the object o_i.

N_i: Total number of pixels in the bounding box b_i, i.e., $N_i = (x_{max} - x_{min}) * (y_{max} - y_{min})$

$(x, y) \in b_i$: Pixels within the bounding box of the object.

E. Direction Estimation

To determine the relative direction of objects, the camera frame is divided into three horizontal segments:

$$R(t) = \{(o_i, r_i) \mid r_i \in \{L, C, R\}\}$$

The direction r_i is calculated as:

$$f = \begin{cases} L, & if\ x_{min} < \frac{w}{3} \\ C, & if\ \frac{w}{3} \le x_{min} \le \frac{2w}{3} \\ R, & if\ x_{min} > \frac{2w}{3} \end{cases}$$

F. Path Assitance

Auditory feedback is generated to inform users about obstacles and their locations:

$$A(t) = \{(o_i, d_i, r_i)\}$$

For example:

- Obstacle detected: Chair, 2.5 meters, on your left.

G. Fail-Safe Mechanism

A fail-safe mechanism monitors the system's status $S(t)$:

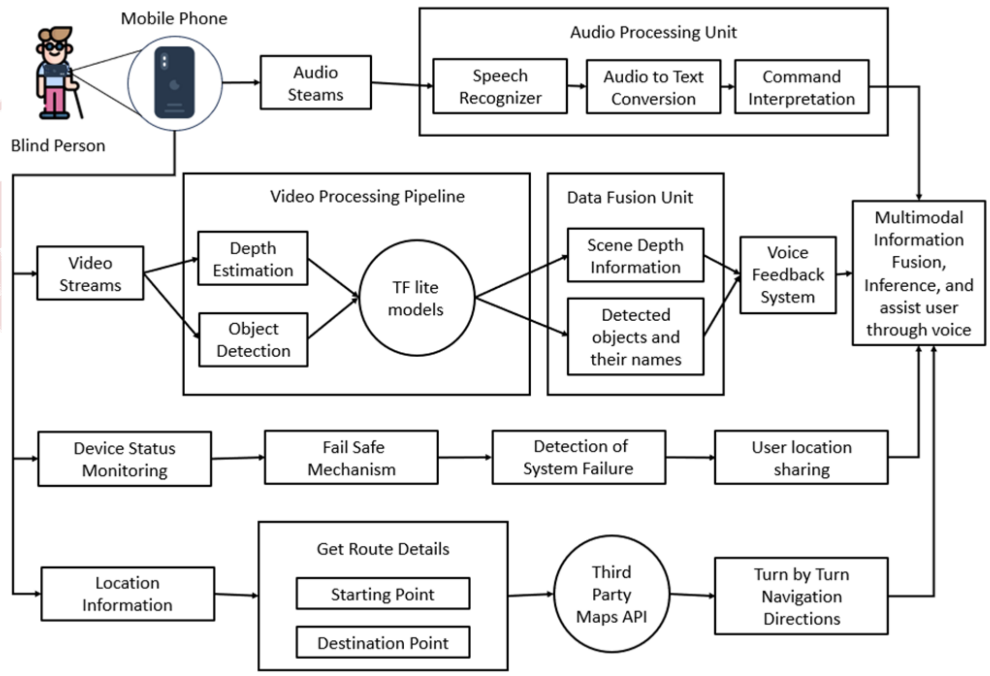

Figure 2. Application Design for Visually Impaired People (VIP)

$$S(t) = \{B(t), G(t), C(t)\}$$

where:

- $B(t)$: Battery level (%)
- $G(t)$: GPS Signal Strength
- $C(t)$: CPU Load (%)

Emergency actions are triggered if any component of $S(t)$ exceeds its corresponding threshold.

V. PROPOSED RESULT

Various applications leverage advanced models and technologies, such as MobileNet SSD, YOLO, MiDaS, and TensorFlow APIs, to assist visually impaired people (VIP) through object detection and spatial awareness. Some solutions also incorporate external hardware, such as depth cameras, to enhance the user experience. Common features of these applications include object identification, distance estimation, people recognition, gesture interpretation, currency counting, and navigation assistance. However, despite these advancements, the adoption of these applications remains limited due to their fragmented nature and the lack of a comprehensive solution that caters to the diverse needs of the visually impaired community in both indoor and outdoor environments.

To address this gap, this paper proposes the development of a comprehensive mobile-based application that integrates multiple essential assistive features into a single platform, designed to assist visually impaired individuals in both indoor and outdoor settings. The overall design of this application is illustrated in Figure 2.

The proposed application requires the user to mount the mobile device at chest level using a pouch or belt. Leveraging the mobile device's built-in video, audio, and location inputs, the application operates entirely through voice commands, ensuring independence for visually impaired users after initial setup, including installation and permissions.

The application's audio processing unit utilizes audio streams and a speech recognizer to interpret voice commands. Simultaneously, video frames captured by the device's camera are processed using a video processing pipeline powered by machine learning models, particularly deep neural networks. To overcome the computational and size limitations of TensorFlow for mobile devices, LiteRT models (formerly TensorFlow Lite) [18] are employed. These optimized on-device machine learning models analyze video frames to identify objects, their directions, and their distances, providing real-time auditory feedback to the user.

Real-time location data, obtained via GPS, enables turn-by-turn navigation assistance. By using third-party mapping

APIs such as Google Maps Platform [19] or Mapbox [20], the application suggests routes from the user's start location to their destination. Considering the potential challenges of running resource-intensive tasks on mobile devices, such as battery depletion or system unresponsiveness, the application incorporates a fail-safe mechanism. This mechanism periodically monitors the device's status and, in the event of failure, shares the user's location with pre-registered family members, ensuring timely assistance and enhancing user safety.

By integrating advanced machine learning models, optimized on-device processing, real-time location tracking, and fail-safe mechanisms, the proposed application offers a robust, user-centric solution to assist visually impaired people (VIP). Its ability to seamlessly operate through voice commands and provide comprehensive assistance for both indoor and outdoor navigation ensures accessibility and independence for its users. This unified approach aims to bridge the gaps in existing solutions, offering a reliable, scalable, and inclusive platform to significantly enhance the mobility and quality of life for the visually impaired community.

VI. CONCLUSION

This survey provides valuable insights into the current state of assistive technologies for visually impaired individuals, emphasizing both the areas where these solutions effectively address user needs and where they exhibit limitations. The analysis of responses and preferences from visually impaired users highlights essential requirements for effective assistive devices, including real-time feedback, intuitive interfaces, and reliable obstacle detection. The findings also emphasize the significance of fostering collaboration among researchers, developers, and end-users to ensure that assistive technologies are better aligned with real-world needs and environments. This paper recommends adopting a targeted approach in the development of future assistive technologies to address existing gaps and enhance the overall user experience for visually impaired individuals. By contributing to the knowledge base in this domain, this research aims to support the creation of more inclusive and accessible technological advancements that promote independence and well-being within the visually impaired community.

REFERENCES

[1] Blindness and Vision Impairment section of World Health Organization's Website, August 2023: https://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment

[2] Harvey P. T. (2003). Common eye diseases of elderly people: identifying and treating causes of vision loss. Gerontology, 49(1), 1–11. https://doi.org/10.1159/000066507

[3] Alwi, S.R., & Ahmad, M.N. (2013). Survey on outdoor navigation system needs for blind people. 2013 IEEE Student Conference on Research and Developement, 144-148. https://doi.org/10.1109/

SCORED.2013.7002560

[4] Nagraj, A., Kuber, R., Hamidi, F., & Prasad, R. S. G. (2021). Investigating the navigational habits of people who are blind in India. In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21) (Article 33, pp. 1–10). Association for Computing Machinery. https://doi.org/10.1145/3441852.3471203

[5] El-Taher, F. E.-Z., Miralles-Pechuán, L., Courtney, J., Millar, K., Smith, C., & McKeever, S. (2023). A survey on outdoor navigation applications for people with visual impairments. IEEE Access, 11, 14647–14666. https://doi.org/10.1109/ACCESS.2023.3244073

[6] See, A. R., Sasing, B. G., & Advincula, W. D. (2022). A smartphone-based mobility assistant using depth imaging for visually impaired and blind. Applied Sciences, 12(6), 2802. https://doi.org/10.3390/app12062802

[7] Badave, A., Jagtap, R., Kaovasia, R., Rahatwad, S., & Kulkarni, S. (2020). Android-based object detection system for visually impaired. 2020 International Conference on Industry 4.0 Technology (I4Tech), 34–38. https://doi.org/10.1109/I4Tech48345.2020.9102694

[8] Tensorflow Object Detection API Documentation: https://tensorflow-object-detection-api-tutorial.readthedocs.io/en/latest

[9] Acar, T., Solmaz, A., Cengiz, I., & Bozkir, A. S. (2024). From pixels to paths: SIGHT – A vision-based navigation aid for the visually impaired. 2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 1–8. https://doi.org/10.1109/HORA61326.2024.10550694

[10] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779–788. https://doi.org/10.1109/CVPR.2016.91

[11] Ranftl, R., Lasinger, K., Hafner, D., & Koltun, V. (2020). Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. IEEE Transactions on Pattern Analysis and Machine Intelligence, PP, 1–1. https://doi.org/10.1109/TPAMI.2020.3019967

[12] Vaidya, S., Shah, N., Shah, N., & Shankarmani, R. (2020). Real-time object detection for visually challenged people. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 311–316. https://doi.org/10.1109/ICICCS48265.2020.9121085

[13] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), Computer vision – ECCV 2016 (Lecture Notes in Computer Science, Vol. 9905). Springer, Cham. https://doi.org/10.1007/978-3-319-46448-0_2

[14] Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv. https://doi.org/10.48550/arXiv.1704.04861

[15] Amarasinghe, C. K., Pinto, R. D. S. P., & Sudusinghe, K. N. (2022). Guardian - Smart assistant tool for visually impaired people. 2022 4th International Conference on Advancements in Computing (ICAC), 429–434. https://doi.org/10.1109/ICAC57685.2022.10025262

[16] Costales, J. A., Vida, A. K. S., & Albino, M. G. (2023). Mobile-based navigation system for visually impaired persons using image detection with voice translation. 2023 3rd International Conference on Information Communication and Software Engineering (ICICSE), 32–36. https://doi.org/10.1109/ICICSE58435.2023.10211761

[17] Ajina, A., Lochan, R., Saha, M., Showghi, R. B. K., & Harini, S. (2024). Vision beyond sight: An AI-assisted navigation system in indoor environments for the visually impaired. 2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS), 1–6. https://doi.org/10.1109/ICETCS61022.2024.10543550

[18] Lite Runtime Section of Google AI Edge Documentation: https://ai.google.dev/edge/litert

[19] Google Maps Platform Documentation and API References: https://developers.google.com/maps/documentation

[20] Mapbox Documentation and API references: https://docs.mapbox.com