International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

Blindaid: Assisting The Visually Impaired In Object Detection And Tracking Using Slam

Dr. Shyamala G[1], Dr Pallavi G B[2], Dr Latha N R[3], Shreesha H Shetty[4], Sushanth[5], Varun Bhatkal[6], V Kenny Philip[7]

[1,2,3] *Professor, Department of Computer Science and Engineering, B.M.S. College of Engineering, Bangalore*
[4,5,6,7] UG Students, Department of Computer Science and Engineering, B.M.S. College of Engineering, Bangalore
[1]shyamala.cse@bmsce.ac.in
[2]pallavi.cse@bmsce.ac.in
[3]latha.cse@bmsce.ac.in
[4]shreesha.cs21@bmsce.ac.in
[5]sushanth.cs21@bmsce.ac.in
[6]varunatmaram.cs21@bmsce.ac.in
[7]philip.cs21@bmsce.ac.in

Abstract—When a blind person loses something or forgets where they put it, it can be difficult and time-consuming to locate it. To address this problem, we created blindAid, an application that recognizes the object the user is using and recalls its most recent location. If the user needs the object and can't remember where he put it, the application will assist him in finding it by utilizing voice commands. This study proposes a new Internet of Things-based solution that combines advanced object recognition with Simultaneous Localization and Mapping (SLAM) technology.. We have made an IOT-model which can communicate with the application we have created. Voice recognition is enabled in the application to record user requests through voice commands and provide a speech output in response. The program can also operate in the background, eliminating the requirement for constant openness. This makes life easier for blind people and enhances the user experience.

Keywords: object detection, SLAM, IOT, machine learning, and audio feedback

I.INTRODUCTION

A visually impaired individual may occasionally find it difficult to locate basic items such as a water bottle, a cell phone, or other items. When he is by himself, needs water, and can't find the bottle, this could make him panic. Additionally, the user may find the activity to be extremely stressful and time-consuming. These difficulties could cause individuals to feel overly reliant on others, which could be uncomfortable and have an impact on their mental health. In order to overcome these obstacles, a dependable and trustworthy application and gadget must be developed, and advancements must be made.

In order to overcome these difficulties, this research suggests an Internet of Things-based system that combines object detection with Simultaneous Localization and Mapping (SLAM) technology. The technology is designed to recognize objects in real time, map the locations of objects in indoor areas, and swiftly guide people to the things they desire. Visually impaired people can more effectively explore and interact with their surroundings because to the system's combination of advanced item detection techniques and spatial mapping.

This method ensures accessibility and cost-effectiveness by utilizing low-cost hardware, such as Raspberry Pi and camera modules. Additionally, by providing customers with clear instructions, it leverages aural feedback to reduce the need for trial-and-error searches. The effort has a transforming impact on the lives of visually impaired individuals by improving their mobility, freedom, and quality of life. It also aligns with the broader goals of promoting inclusivity and accessibility.

II. SCOPE

By creating an Internet of Things-based assistive technology, this project seeks to lessen the challenges visually impaired persons face while trying to locate misplaced objects in indoor environments, such as homes and workplaces. The system utilizes state-of-the-art Simultaneous Localization and Mapping (SLAM) and object detection technologies to facilitate precise object recognition and real-time spatial mapping. Additionally, lowering the device's manufacturing costs to make it affordable is the goal. The project uses Raspberry pi 5 on which all the other components like camera, wifi-module to communicate with the application, and OCR(Optical Character Recognition) are integrated.

The application has a user-friendly interface, which is built keeping in mind the visually challenged person's perspective, making sure the user is capable of finding the object he/she is trying to find without other person's intervention. Additionally, the object detection model is trained well to detect most of the objects especially the objects that are needed in day-to-day chores.

The application can run in the background as well avoiding the need to keep the app open all the time. The app is designed in such that the user can activate it to accept voice commands by just calling a wake-up word which can be customized based on the user's need(like "Hey, Guider) so that there is no need for the user to open the app and click on a button as it can be challenging for a visually challenged person.

III. LITERATURE SURVEY

1. Machine Learning & Deep Learning-Based Object Detection

The research paper focuses on helping visually challenged people by building a smart cane which is embedded with ultrasonic sensors for object detection, a camera for image capturing, and a Raspberry Pi 3B model. It also uses YOLOv5 and OCR for text recognition. It offers real-time object detection with a 5-meter range and is primarily focused on outdoor mobility by comprehending road conditions. This enables the user to utilize a mobile application to travel on their own. Future advancements will include LiDAR for accurate depth sensing, SLAM-based interior navigation for seamless mobility inside buildings, AI-powered identification enhancements for enhanced accuracy, and crowd sourced hazard updates to keep the system up to date with evolving circumstances. These advancements will increase visually impaired consumers' dependability, accessibility, and freedom. [1].

This solution integrates IoT and AI technology with a smart cane to enhance mobility for those with vision impairments. It has speech broadcasting, Bluetooth, a camera, and ultrasonic sensors. It uses cloud-based processing, YOLOv5 for object detection, and OCR for text recognition to enable real-time obstacle detection, road condition recognition, and traffic sign identification. To ensure a smooth and independent travel experience, a mobile application also offers Bluetooth connectivity, one-click cane retrieval, navigation help, and assistance with public transportation.

This technology makes effective obstacle avoidance and environmental awareness possible by fusing computer vision, deep learning, and Internet of Things-based sensor modules. Future advances will include LiDAR for high-precision depth estimates, SLAM-based indoor navigation for greater autonomy, AI-driven recognition enhancements for better accuracy, and crowd sourced hazard updates to guarantee adaptability to real-world conditions [2].

A system is developed that consists of a smart cane and smart glasses that employ computer vision and machine learning to detect objects, detect and avoid potholes, recognize road signs, and assist with autonomous mobility. It handles real-time video data from a camera installed on a smart glass using OpenCV and TensorFlow Lite. It offers voice commands for precise lane detection, instant notifications when danger is detected, and the ability to recognize emotions and recognizable people.

Future developments will concentrate on lowering costs while locating and utilizing better, less expensive alternatives, optimizing hardware for higher performance, and boosting real-time processing efficiency in order to guarantee that such systems are accessible and affordable to a wide range of individuals. [5].

In order to reliably identify a wide range of items, this system uses 3D glasses in conjunction with deep learning-based object identification, facial recognition to recognize faces and emotions, and optical character recognition (OCR) to identify text. Ultrasonic sensors and a Raspberry Pi are used in the system. The system uses computer vision and deep learning models to provide real-time item detection, text recognition, and facial recognition. With the use of the program, which offers aural feedback, navigational direction, and alarms when needed, users may explore both indoors and outdoors on their own.

Future developments will concentrate on increasing text recognition accuracy, streamlining real-time communication to provide quicker feedback, and strengthening the device's robustness for long-term use. [11].

This system employs Mobile-Net and Raspberry Pi to achieve an accuracy of 83.3% in real-time item detection, and it integrates GPS/GSM for enhanced safety and position monitoring. It use object identification based on deep learning to recognize and classify objects and provide audible feedback to assist vision impaired people in navigating their surroundings. Additionally, a web-based program makes remote monitoring possible, enabling caregivers to track the user's location through real-time location updates and, for added security, snapshots.

Future improvements to this system will include deep learning models for increased accuracy and faster processing, bone-conduction audio feedback for improved accessibility and user experience, and hardware efficiency optimization to maximize performance while using the least amount of energy. These improvements aim to improve the system's dependability, efficiency, and accessibility for those who are blind or visually impaired .[19].

This device combines YOLO-v3, deep learning algorithms, and ultrasonic sensors to provide visually impaired people with intelligent navigation and effective multi-object recognition. With a high accuracy of 95.19%, it offers real-time obstacle recognition, guaranteeing dependable navigation support. Through personalized spoken instructions, users can receive real-time advice thanks to the system's scene segmentation and auditory feedback features. The gadget is also made to operate on inexpensive CPUs, which makes it a widely available and reasonably priced alternative.

Future enhancements will include text reading and facial recognition capabilities to allow users to read printed text in their surroundings and identify known faces. Enhancements in hardware minimization, algorithm optimization for faster processing, and user-friendly compact design are also planned to improve the system's efficacy and usability for a wider range of visually impaired users. [20].

2. AI-Based Indoor & Outdoor Navigation Systems

This system uses SSD MobileNet-V2 FPNlite for real-time object identification and audio support, enabling visually impaired individuals to move around indoor spaces on their own. Through the integration of an Android application, users can obtain voice-based guidance in real-time. This ensures high accuracy, rapid processing speed, and resource efficiency for seamless mobile deployment. By employing deep learning and AI-powered recognition, the technology enhances item detection capabilities, reducing dependency on external assistance and enhancing user autonomy.

Improved obstacle avoidance for safer travel, higher detection accuracy for more accurate item recognition, and expanding the system's functionality to support outdoor navigation will be the main goals of future research.

Real-time processing efficiency optimization and multi-sensor fusion techniques are two more improvements that increase the system's usability, dependability, and adaptability for people with visual impairments[7].

This method enhances Visual SLAM (VSLAM) and boosts localization accuracy in dynamic scenarios by integrating object detection, instance segmentation, and geometry-based filtering to remove dynamic feature points. The method combines semantic segmentation for static feature point extraction with a geometric verification module to ensure precise pose prediction and quick back-end optimization. Verified on the TUM and EUROC datasets, it demonstrated improved stability and reduced trajectory errors, outperforming existing state-of-the-art VSLAM systems in complex and dynamic conditions.

Future improvements will concentrate on improving real-time semantic mapping, enhancing deep learning-based detection models for more effective processing, and using multi-view techniques for more dynamic object filtering. Other improvements, like customizing the system for wearable navigation devices and enhancing its ability to handle shifting environmental conditions, guarantee greater dependability and real-time adaptation for visually impaired users[8].

This method provides precise navigation and obstacle recognition by combining visual SLAM with model-based localization using a monocular camera. Semantic segmentation is used to enhance real-time feedback, which helps visually impaired persons navigate unfamiliar circumstances. The system employs a novel global localization technique (VB-GPS) that blends visual SLAM with picture segmentation to improve scene comprehension and accurate placement. A localization accuracy of 0.27 meters was found during user trials, greatly increasing users' trust in obstacle recognition, navigation, and spatial awareness.

Future advancements will concentrate on improving performance in challenging scenarios, guaranteeing wearable device compatibility, and incorporating 5G/cloud-based processing for smooth real-time updates and increased accessibility. These developments will further improve navigation efficacy, security, and user autonomy .[14].

This method gives visually impaired people haptic and audio input by combining convolutional neural networks (CNNs) with ARKit to construct virtual routes for item detection and navigation. This system uses augmented reality (AR) technology to create established virtual trails, allowing for indoor and outdoor navigation without requiring environmental changes, unlike traditional systems that rely on actual markers like painted lines or QR codes. Real-time landmark, building, and monument identification is made possible by CNN-based object detection, which also improves environment recognition. Tests on heritage sites revealed high tracking accuracy and recognition rates of over 90%, proving the system's effectiveness in demanding real-world situations.

Future enhancements will include adding wearable technology for a better user experience, increasing real-world testing for validation in a range of circumstances, and deploying 5G/cloud-based connectivity for faster processing and remote updates. These improvements will ensure greater accessibility, usability, and independence for individuals who are blind or visually impaired. [15].

3. Smart Cane and Wearable Devices with Sensors

This device employs a smart stick loaded with OpenCV and Python to assist those who are vision impaired. It uses the LBPH and Haar Cascade algorithms to identify faces and objects in real time. It efficiently recognizes known faces, traffic signals, and environmental impediments to guarantee safe navigation and gives feedback via vibration or sound. Combining computer vision with AI-driven recognition to improve situational awareness allows users to interact with their environment with more confidence. Additionally, the device has colour differentiation, which aids people with vision impairments in differentiating between surfaces like sidewalks and highways.

Adding Braille keyboards for improved accessibility, identifying cognitive processes for adaptive learning, and growing the training dataset to increase identification accuracy in various contexts will be the main goals of future developments. In order to provide visually impaired people with a more intelligent and user-friendly navigation experience, other improvements include real-time GPS tracking for location assistance and cloud-based AI processing for increased detection efficiency. [3].

Through the use of a smart cane and a walker equipped with many sensors and an SSD-RNN model that allows for real-time obstacle recognition, this technology assists people with visual impairments in navigating both indoor and outdoor environments. The Smart-Fold Cane has ultrasonic sensors, a camera module, and a water sensor, while the Smart-Alert Walker is a lightweight shoe with emergency obstacle detection capabilities. The system uses single-shot detection (SSD) for item categorization and an RNN-based frame generation model to achieve 95% accuracy indoors and 87% accuracy outdoors. Additionally, it provides low-latency feedback and high recognition rates, which significantly boost user confidence and mobility in a variety of contexts.

Future goals include conducting extensive user testing, increasing adaptability in a variety of environmental conditions, and fortifying the system's robustness through complex hardware integration. Other improvements include adding AI-based scene interpretation, enhancing obstacle identification in dynamic circumstances, and optimizing the device's energy economy to guarantee a more reliable, accessible, and intelligent navigation system for visually impaired users. [16].

This device uses a navigation aid that is carried in a backpack and features 3D stereo cameras to map environments and design routes for those with vision impairments. Navigation directions are transmitted via vibrotactile and bone-conduction acoustic signals, and the ambient data is processed in real-time by an embedded processing unit. The device significantly improved mobility and independence with a 90% signal identification accuracy and an average reaction time of 2.7 seconds. However, while moving rapidly, vibrotactile feedback performed less well, suggesting that further development is required.

Actuator location will be modified for better vibration feedback, extensive real-world testing will be conducted in a range of settings, and the system's ability to adapt to various walking speeds and movement patterns will be enhanced. Additional advancements aim to integrate AI-based adaptive feedback mechanisms and boost energy economy to give visually impaired users a more reliable and effective navigation system. [4].

Using multimodal data integrated with an attention-based deep learning model, this system combines a webcam and a haptic glove to classify items. It helps visually impaired people perceive their environment by integrating tactile and visual sensory inputs to provide precise object recognition and real-time feedback. Three attention mechanisms—temporal, channel-wise, and spatial attention—are incorporated into the system to improve the model's capacity to extract important information from multimodal inputs. With a classification accuracy of 99.75%, it demonstrates remarkable resilience in practical situations. Users can also learn about objects through touch and auditory descriptions, making it an immersive instructional tool.

Future developments will concentrate on enhancing accessibility through the system's user interface, including wearable technologies for real-time use, and optimizing attention mechanisms for increased recognition efficiency. In order to ensure wider adoption and practical application for visually impaired users, plans also call for improving tactile feedback for a more engaging experience, deploying cloud-based updates for adaptive learning, and growing datasets to increase identification accuracy [6].

4. Mobile-Based Navigation & AI-Assisted Applications

This system helps visually challenged individuals navigate their surroundings securely by processing camera images using an Android-based platform for voice-guided navigation and real-time obstacle identification. In order to identify barriers, routes, and pertinent landmarks while guaranteeing effective resource management for mobile deployment, it makes use of computer vision and deep learning algorithms. The system showed excellent accuracy and user-friendliness when tested using ISO usability and reliability standards, indicating that it is a promising assistive solution.

Future advancements will include adding ultrasonic sensors to increase the reliability of obstacle detection, improving object detection capabilities for more accurate navigation, and bolstering text recognition for easier reading of signs and labels. Additional enhancements that focus on enhancing real-time processing for faster feedback, enhancing the user interface for accessibility, and expanding dataset training to accommodate a range of environmental conditions guarantee a stable and intuitive navigation experience for visually impaired individuals. [13].

For people who are visually impaired, this system enhances navigation with ICT-based teleguidance technology. It makes remote assistance possible by using a live video broadcast from the user's smartphone camera. With the aid of a remote caregiver who provides real-time navigation support via a smart cane that uses haptic feedback and voice instruction, visually impaired people can safely navigate both interior and outdoor environments. In user studies, users showed high levels of satisfaction and gave the system an overall dependability score of 4.4/5, demonstrating its value in improving mobility and independence.

Future enhancements will concentrate on strengthening system reliability for smooth operation, boosting safety measures to manage a wider range of navigation scenarios, and broadening cultural applicability to accommodate a variety of user needs. Additional developments include bettering real-time video processing, testing in a range of real-world situations, and positioning haptic feedback to guarantee greater accessibility, adaptability, and user confidence in the system[17].

5. Reviews & Comparative Studies on Assistive Technologies

To better understand how AI, deep learning, and the Internet of Things might improve obstacle detection, navigation, and everyday activities for people with visual impairments, this study examines 101 papers on intelligent settings and assistive technologies. The study highlights the growing use of deep learning models for object detection, path recognition, and navigation systems, classifying assistive technologies according to their usefulness and guiding principles. The study also examines the relationship between Ambient Assisted Living (AAL) and Intelligent Environments (IE), demonstrating how both technologies contribute to the creation of more independent and hospitable living environments for people who are blind or visually impaired.

Future recommendations should focus on developing scalable and affordable solutions, particularly for low-income areas, to make assistive technologies accessible to a wider audience. The report also highlights the ethical concerns related to AI, including data protection, bias reduction, and user-centric design. Additionally, cloud-based AI solutions, multimodal sensory feedback, and AR/VR are new technologies that have been identified as potential advancements to further improve accessibility and independence for individuals with visual impairments. [9].

This evaluation examines the challenges faced by visually impaired individuals when navigating urban areas, highlighting significant deficiencies in improved pedestrian maps, traffic light recognition, and crosswalk alignment. It analyzes current smart navigation systems and highlights their drawbacks with regard to real-time hazard detection, accessibility in public areas, and communication with public transportation systems. The study emphasizes the need for AI-driven navigation systems, with an emphasis on deep learning-based item detection, scene segmentation, and dynamic route planning, to enhance mobility, independence, and safety in urban environments.

Future directions include developing more adaptable AI models to handle diverse and dynamic urban environments, improving real-time environmental awareness, and integrating navigation support with smart city infrastructures. Recommendations also focus on expanding multimodal feedback systems, enhancing wearable compatibility, and leveraging crowdsourced data to create incredibly detailed and accessible urban navigation maps in order to give visually impaired people a more inclusive and efficient mobility experience. [12].

This study explores indoor navigation possibilities for those with visual impairments using electronic travel aids (ETA) that do not use GPS. It analyzes the accuracy, affordability, and practicality of several ETA technologies, such as Bluetooth Low Energy (BLE), Ultra-Wideband (UWB), and Wi-Fi-based positioning systems. The study highlights the need for scalable and affordable solutions that address issues like environmental adaptation, infrastructure requirements, and localization accuracy in order to assist visually impaired people in navigating public institutions, supermarkets, office buildings, and homes.

Future recommendations focus on developing environment-neutral solutions that don't require major interior space modifications, developing advanced hardware to improve localization accuracy, and conducting extensive field testing for real-world validation. More developments include enhancing AI-driven indoor positioning algorithms, adding multimodal feedback mechanisms, and maximizing system reliability for practical deployment in order to give visually impaired persons more mobility and accessibility in a range of indoor contexts. [10].

With an emphasis on object recognition, navigation, and everyday life assistance, this assessment classifies assistive technologies for visually impaired people (VIPs) according to their functionality and performance. The study lists the benefits and drawbacks of a number of wearable, sensor-based, and AI-driven assistive technologies, showing that while high-performing tools like FingerReader are great in certain fields, no single device can completely satisfy VIPs' needs. A score-based analysis that ranks devices based on features including accuracy, real-time responsiveness, and indoor/outdoor adaption reveals significant gaps in device integration and user accessibility.

Future study should focus on developing unified and multipurpose assistive systems to promote seamless mobility, enhanced safety, and greater independence. Enhancing multimodal feedback systems to offer more dependable, adaptable, and user-friendly solutions for visually impaired people, including AI-based recognition models for improved environmental awareness, and making it more affordable for low-income areas are some of the recommendations [18].

IV. IMPLEMENTATION

A Raspberry Pi-based camera setup, a central processing unit, and a mobile application make up the modular architecture used to construct the BlindAid system. To provide visually challenged users with smooth object identification, navigation, and text recognition, each module communicates with the others via a common Wi-Fi network and cloud infrastructure.

A. Raspberry Pi Camera Stream Setup

To record live footage of the surroundings, a Raspberry Pi 4 with a regular camera module is utilized. The Pi is set up to use lightweight protocols to stream this camera feed over a public Wi-Fi network. It serves as a video source for the rest of the system rather than doing any processing locally.
The primary processing component, which runs on a desktop or laptop computer linked to the same network, uses this stream as its visual input.

B. Central Processing with Object Detection

A laptop or PC on the receiving end views the video feed and uses the YOLOv8n model to recognize objects in real time. Because of its optimal performance and capacity to process frames effectively on mid-range computers, YOLOv8n was selected. Every entering frame is examined to identify and label well-known items, like books, bottles, and bags.

After an object is identified, its position within the frame is used to determine its direction relative to the user. This spatial information is then translated into compass-based directions (e.g., "north-east" or "30° to the right"), based on the assumption that the mobile device is being held in a forward-facing orientation.

The identified object name and its corresponding direction are then transmitted to the user's mobile device via the Firebase Realtime Database.

C. Mobile Application and User Interaction

The mobile application acts as the primary interface for the end-user. It is designed with accessibility features to support voice-based interaction. The app continuously listens for voice commands. For example, when the user asks, "Where is my bottle?", the app uses the Gemini API to interpret the spoken query and extract the object of interest.

Using the object name, the app queries Firebase for the last known direction associated with it. It compares this with the current orientation of the mobile phone (which uses the built-in compass sensor) and computes the angle of rotation required to face the object. The app then guides the user with spoken directions such as "Turn left by 45 degrees" or "Turn right by 30 degrees." Once alignment is achieved, the user is notified with a confirmation like "You are now facing the bottle."

This interaction loop allows the user to independently locate and face objects without requiring external markers or Bluetooth beacons

D. OCR-Based Text Reading

In addition to navigation, BlindAid includes a feature to read textual content from the user's surroundings. When the user issues a voice command like "Read," the app captures a still frame from the camera feed and applies OCR (Optical Character Recognition) to extract the visible text. The recognized text is then converted into speech using the mobile device's text-to-speech engine, and read aloud to the user.
This allows the visually impaired to interpret printed signs, documents, and labels in real time, directly through their smartphone.

E. System Workflow Summary

The overall system functions as a coordinated pipeline involving hardware, cloud services, and mobile interfaces. The step-by-step data and control flow are outlined below:

1. **Environment Capture**

   - A Raspberry Pi 4, connected to a camera module, captures real-time video of the user's surroundings.
   - The video stream is transmitted over a common Wi-Fi network using a lightweight stream server.

2. **Object Detection and Direction Estimation**

   - A central processing unit (laptop/PC) receives the video stream and processes each frame using the YOLOv8n object detection model.
   - Detected objects are annotated and their relative positions within the frame are analyzed to estimate the direction with respect to the user.
   - The object name and its compass-based direction are uploaded to the Firebase Realtime Database.

3. **Mobile Application Interaction**

   - The user interacts with the BlindAid mobile application using voice commands.
   - Upon querying for a specific object (e.g., "Where is my bottle?"), the app uses the Gemini API to extract the object keyword from natural language input.
   - The app fetches the corresponding object direction from Firebase and compares it with the phone's current orientation.
   - Real-time voice instructions (e.g., "Turn right by 30 degrees") are provided to help the user align with the object.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

4. **Text Recognition and Audio Feedback**

   - When prompted by commands like "Read," the system captures a frame and performs OCR to extract text content.
   - The extracted text is converted into speech and read aloud via the phone's audio output.

5. **End-to-End Experience**

   - All data-intensive and AI-based tasks (object detection, language parsing, OCR) are handled by the laptop/PC, keeping the mobile application lightweight.
   - Firebase ensures low-latency communication and synchronization between components, even in limited connectivity scenarios.

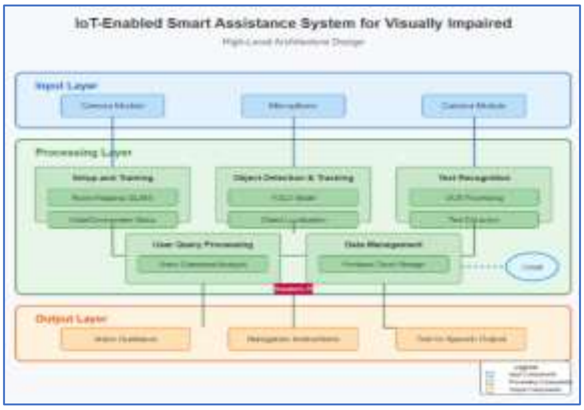

Fig 1. Blind Aid High level architecture design

Fig 2. Sequence diagram for Blind Aid

V. FUTURE WORK

While the current implementation of BlindAid offers reliable object recognition and voice-guided navigation within a localized environment, there are several avenues for future enhancement to improve system accuracy, scalability, and utility in diverse real-world scenarios.

A. Enhancing Directional Precision with Sensor Fusion

The current method of estimating object direction relies primarily on camera-relative positioning and mobile compass orientation. However, this estimation can be significantly improved through sensor fusion techniques, by integrating additional phone sensors such as the gyroscope, accelerometer, and magnetometer. By leveraging these internal sensors, the system can provide more precise angular calculations, thereby improving alignment accuracy when guiding users toward objects. This would be especially useful in environments with subtle object placements or tight spaces where fine orientation is critical.

B. Improving Object Localization Accuracy

At present, object localization is relative to the user's orientation during detection. This can be enhanced by refining the object recognition module using higher-resolution models or deploying depth estimation algorithms to get better spatial context. Incorporating stereo vision (with dual cameras) or AI-based depth estimation can help in calculating the object's distance, allowing the system to generate more accurate navigation cues such as, "Take three steps forward and turn 45° to the right."

C. Integration of GPS for Long-Range Object Tracking

One major limitation of the current BlindAid system is its operational range, which is confined to real-time proximity detection. To address this, we propose integrating the smartphone's GPS module to store the last known location of detected objects. For example, if a user leaves a personal item (e.g., charger or helmet) at the office, the system can record the GPS coordinates at the moment the object was last recognized and stored in the cloud.

Later, when the user is in a different location—such as home—and queries the system for that object, BlindAid can respond with meaningful feedback like:

"*Your charger was last seen at your office, 10 km away.*"

This feature would significantly reduce uncertainty and anxiety for visually impaired users by confirming the object's whereabouts even when outside their current environment. Additionally, the app can help users plan how to collect the thing by incorporating maps or route ideas.

D. System Scalability and Environment Independence

The BlindAid system can progress from an indoor assisting device to a city-scale object tracking assistant by incorporating GPS-enabled object memory and cloud synchronization. This makes it possible to expand its use in a variety of settings, including homes, offices, and public spaces, without requiring constant camera feeds or human updates. Users can manage and recover misplaced items no matter how far away they are by using the ability to recall and locate objects recorded in one area.

E. Offline and Edge Computing Capability

By utilizing frameworks like TensorFlow Lite or Core ML to run lightweight object identification models directly on the mobile device, future iterations of the system might also enable offline processing. This would provide on-the-go help even in places with little connectivity by granting some independence from the central PC and internet.

BlindAid: An Intelligent Companion for the Visually Challenged offers a ground-breaking method of empowering people with visual impairments by addressing significant problems in text recognition, object tracking, and navigation. By integrating state-of-the-art technologies like IoT, SLAM, and OCR, the system offers real-time voice assistance with exceptional accuracy and reliability, enabling users to easily engage with their surroundings. Because of its offline capabilities and secure cloud synchronization, it offers flexibility even in places with inadequate access, so benefiting underserved areas. BlindAid emphasizes scalability, user-friendliness, and new features to bridge the gap between assistive technology and practical usage. This encourages self-reliance and significantly improves its users' quality of life. This study demonstrates how technology can lead to significant societal change and establishes a standard for assistive technology in the future.

REFERENCES

[1] Advanced Multimodal Navigation Perception System for the Visually Impaired Zhuyu Li, et al., 2024 7th International Conference on Electronics Technology (ICET), DOI: 10.1109/ICET61945.2024.10672952

[2] Zhuyu Li, Bing Xiong, Jialing Huang, Zhuojiang Xie, Tingli Guo, Hao Gu, "An Intelligent Navigation System for the Visually Impaired Based on Machine Vision," 2024 7th International Conference on Electronics Technology (ICET), Changsha University of Science and Technology, DOI: 10.1109/ICET61945.2024.10672952.T. K. De, P. Likhitha, J. Vamsi, et al., "Diabetes prediction using machine learning," International Journal of Advanced Research in Computer and Communication Engineering, vol. 13, no. 3, Mar. 2024.

[3] Muhammad Shuaib Qureshi, et al., 2023 IEEE Smart Cities Conference (ISC2), DOI: 10.1109/ISC257844.2023.10293380K. Ramudu, V. M. Mohan, D. Jyothirmai, D.V.S.S. Prasad, R. Agrawal, and S. Boopathi, "Machine learning and artificial intelligence in disease prediction: Applications, challenges, limitations, case studies, and future directions," Aug. 2023.

[4] Julian Sessner, et al., 2022 IEEE/SICE International Symposium on System Integration (SII), DOI: 10.1109/SII52469.2022.9708751

[5] S.R.D. Sudaraka Keshara, Jathusanan Ellankovan, W.R.J.M. Weragoda, Madushan W.A, and Sanjeevi Chandrasiri. Real-time Smart Navigation System for Visually Impaired People. 2022 4th International Conference on Advances in Computing (ICAC), IEEE. DOI: 10.1109/ICAC57685.2022.10025323K. Reshma, P. Niharika, J. Haneesha, K. Rajavardhan, and S. Swaroop, "ML applications for disease diagnosis," 2023.

[6] Xinrong Li, Meiyu Huang, Yao Xu, Yingze Cao, Yamei Lu, Pengfei Wang, and Xueshuang Xiang. AviPer: Assisting Visually Impaired People to Perceive the World with Visual- Tactile Multimodal Attention Network. CCF Transactions on Pervasive Computing and Interaction (2022), 4:219–239.A. Govindu and S. Palwe, "Early detection of Parkinson's disease using machine learning," Procedia Computer Science, 2023.

[7] A. Ajina, Showghi R B K, Lochan R, S. Harini, and Mohini Saha, "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, IEEE, DOI: 10.1109/ICETCS61022.2024.10543550.

[8] Shaohu Li, Jason Gu, and Ying Feng, "Visual SLAM with a Multi-Modal Semantic Framework for the Visually Impaired Navigation-Aided Device," 2023 8th International Conference on Advanced Robotics and Mechatronics (ICARM), IEEE, DOI: 10.1109/ICARM58088.2023.10218788 H. A. Al-Alshaikh, P. Prabu, R. C. Poonia, A. K. J. Saudagar, M. Yadav, H. S. AlSagri, and A. A. AlSanad, "Comprehensive evaluation and performance analysis of machine learning in heart disease prediction," Scientific Reports, Mar. 29, 2024.

[9] Leandro Rossetti de Souza, Rosemary Francisco, João Elison da Rosa Tavares, Jorge Luis Victória Barbosa.R. Bharti, S. Pande, A. Khamparia, P. Singh, M. Shabaz, and G. Dhiman, "Deep learning approaches in healthcare," Computational Intelligence and Neuroscience, Jul. 1, 2021.

[10] Bourne, R.R.A.; Flaxman, S.R.; Braithwaite, T.; Cicinelli, M.V.; Das, A.; Jonas, J.B.; Keeffe, J.; Kempen, J.H.; Leasher, J.; Limburg, H.; et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob. Health 2017, 5, e888–e897. Tapu, R.; Mocanu, B.; Zaharia, T. Wearable assistive devices for visually impaired: A state of the art survey. Pattern Recognit. Lett. 2018, in press. Vatansever, S.; Butun, I.

A broad overview of GPS fundamentals: Now and future. In Proceedings of the 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 9– 11 January 2017; pp. 1–6.C. A. Hassan, J. Iqbal, R. Irfan, S. Hussain, A. D. Algarni, S. S. H. Bukhari, N. Alturki, and S. S. Ullah, "Sensor- based diagnostics for diseases," Sensors, Sep. 23, 2022.

[11] Mohamed, I., Salah, M., & Farghal, A. (2022). Camera-Based Navigation System for Blind and Visually Impaired People. Sohag Engineering Journal, 3(1), 1-13. DOI: 10.21608/sej.2022.155927.1018.T. Miller and J. Grant, "AI models for disease management," Aug. 2023.

[12] ostales, J. A., Albino, M. G., & Vida, A. K. (2023). Mobile Based Navigation System for Visually Impaired Person Using Image Detection with Voice Translation. 2023 3rd International Conference on Information Communication and Software Engineering (ICICSE), DOI: 10.1109/ICICSE58435.2023.10211761.

[13] : El-taher, F.E.-z.; Taha, A.; Courtney, J.; Mckeever, S. A Systematic Review of Urban Navigation Systems for Visually Impaired People. Sensors 2021, 21, 3103

[14] Ping-Jung Duh, Yu-Cheng Sung, Liang-Yu Fan Chiang, Yung-Ju Chang, and Kuan-Wen Chen, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 23, 2021

[15] : Lo Valvo, A.; Croce, D.; Garlisi, D.; Giuliano, F.; Giarré, L.; Tinnirello, I. A Navigation and Augmented Reality System for Visually Impaired People. Sensors 2021, 21, 3061.

[16] Suman, S., Mishra, S., Sahoo, K. S., & Nayyar, A. (2022). Vision navigator: a smart and intelligent obstacle recognition model for visually impaired users. Mobile Information Systems, 2022(1), 9715891.

[17] Chaudary, B., Pohjolainen, S., Aziz, S., Arhippainen, L. and Pulli, P., 2023. Teleguidance-based remote navigation assistance for visually impaired and blind people—usability and user experience. Virtual Reality, 27(1), pp.141-158.

[18] Zafar, S., Asif, M., Ahmad, M.B., Ghazal, T.M., Faiz, T., Ahmad, M. and Khan, M.A., 2022. Assistive devices analysis for visually impaired persons: A review on taxonomy. IEEE Access, 10, pp.13354-13366.

[19] Ashiq, F., Asif, M., Ahmad, M.B., Zafar, S., Masood, K., Mahmood, T., Mahmood, M.T. and Lee, I.H., 2022. CNN-based object recognition and tracking system to assist visually impaired people. IEEE access, 10, pp.14819-14834.

[20] Joshi, R.C., Yadav, S., Dutta, M.K. and Travieso-Gonzalez, C.M., 2020. Efficient multi-object detection and smart navigation using artificial intelligence for visually impaired people. Entropy, 22(9), p.941