IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

AN INTELLIGENT NAVIGATION AND RECOGNITION SYSTEM FOR VISUALLY IMPAIRED INDIVIDUALS

A YOLOv8-Enabled Wearable Assistive System for Independent Mobility

[1]Amal Joy, [2]Georgy Daniel Mathews, [3]Mohammed Ziyad O T, [4]Dinto Mathew, [5]Emmanuel Babu
[1,2,3,4]Under Graduate Student, [5]Assistant Professor, [6]Associate Professor
[1,2,3,4,5,6]Electrical and Electronics Engineering,
[1,2,3,4,5,6]Mar Athanasius College of Engineering, Kothamangalam, India

Abstract: Independent navigation and social interaction remain challenging for individuals with visual impairments, necessitating the development of assistive technologies that enhance safety and independence. This work proposes a smart wearable system designed to provide real-time facial recognition and location-based support. The system utilizes a computer vision model, YOLOv8, deployed on a Raspberry Pi 5 platform to identify familiar individuals and deliver immediate audio feedback to the user. To ensure user safety, a GPS module is incorporated for continuous location tracking, while a GSM module enables emergency communication by sending alerts and initiating calls to predefined contacts during critical situations. All components are integrated into a compact and lightweight wearable configuration consisting of camera-equipped glasses connected to a portable processing unit. By focusing on essential functionalities such as facial recognition, real-time tracking, and emergency connectivity, the proposed system offers an efficient and reliable assistive solution, thereby promoting greater autonomy, confidence, and social interaction for visually impaired individuals.

***Index Terms* -** Assistive Technology, Computer Vision, Raspberry Pi, YOLOv8, Visual Impairment.

I. INTRODUCTION

1.1 Background and Motivation

Vision loss drastically limits a person's capacity to safely traverse and interact with their environment. The World Health Organization estimates that roughly 2.2 billion individuals experience some form of visual impairment, including 36 million who are entirely blind. Consequently, these individuals confront daily hurdles when executing routine activities that sighted persons navigate effortlessly. This project is driven by the urgent necessity to engineer sophisticated technological aids capable of minimizing this accessibility deficit and granting visually challenged users greater autonomy.

Conventional mobility aids, such as guide dogs and white canes, provide undeniable benefits but fall short in offering a complete picture of the user's surroundings. For instance, white canes are adept at sensing ground-level barriers but fail to detect elevated hazards, read signage, or recognize approaching faces. Similarly, guide dogs demand rigorous training and have a finite active lifespan. These constraints underscore the demand for advanced assistive frameworks that capitalize on recent leaps in artificial intelligence, computer vision, and compact computing to deliver a more holistic navigational solution. The swift advancement of single-board computers, notably the Raspberry Pi ecosystem, paired with breakthroughs in deep learning, has unlocked new possibilities for creating smart accessibility tools. These platforms can analyze intricate visual inputs instantly and generate relevant audio warnings, serving as a functional synthetic vision system.

1.2 Challenges for Visually Impaired

Individuals coping with vision loss navigate a multitude of daily obstacles that compromise their autonomy and quality of life. Moving through unknown areas introduces severe difficulties, such as avoiding physical barriers, identifying stairs, and negotiating busy intersections. Furthermore, the inability to visually recognize peers often results in social distancing and a reliance on companions for basic identification. Routine chores, including sorting currency, identifying household items, and reading paperwork, also become strenuous endeavors that hinder self-sufficiency.

Achieving environmental awareness is especially problematic, as visually impaired users cannot naturally detect impending threats like fast-moving vehicles or sudden drops in elevation. This deficit in spatial comprehension not only restricts free movement but also poses genuine physical danger. Moreover, navigating emergency scenarios becomes exceptionally perilous when a person is unable to relay their exact location or status to emergency responders or family members. Most existing assistive gadgets target these issues individually, forcing users to juggle multiple tools or master convoluted interfaces. Therefore, an integrated framework that merges hazard detection, text translation, object classification, and emergency broadcasting into one intuitive device is highly necessary.

1.3 Technological Solutions

Modern leaps in embedded electronics and machine learning have paved the way for smart aids that were previously unfeasible. Merging capable microcomputers, like the Raspberry Pi, with cutting-edge visual processing algorithms yields a powerful foundation for building advanced awareness tools. These devices evaluate live camera feeds and generate instant verbal guidance, vastly improving a user's self-reliance.

The rise of streamlined deep learning architectures, specifically YOLO (You Only Look Once) for instant object tracking, has transformed how computer vision functions on low-power devices. These algorithms can pinpoint and categorize multiple targets simultaneously with low computational overhead, making them a perfect fit for portable electronics. Concurrently, enhancements in Optical Character Recognition (OCR) have refined text extraction, allowing systems to reliably read diverse fonts and formats. Sensing hardware has also seen massive improvements. Ultrasonic rangefinders now offer highly accurate proximity readings, while modern GPS receivers deliver pinpoint tracking even in dense cityscapes. GSM integration further empowers these devices with dependable cellular connectivity, permitting automated distress signals and remote monitoring.

1.4 System Overview

The proposed work details a smart assistance platform engineered to boost the navigational safety and self-reliance of visually challenged users. This setup offers a major leap forward by consolidating several vital functions into one unified, wearable package. Operating at the heart of this system is a Raspberry Pi 5, chosen for its optimal blend of processing muscle, energy efficiency, and small footprint.

The hardware layout utilizes a combination of sensory inputs to construct a detailed map of the user's vicinity. High-definition Pi cameras record visual data, while ultrasonic modules supply supplementary depth measurements for close-range hazard detection. This dual-sensor strategy guarantees reliable operation across varying lighting and weather conditions. A standout feature of this device is its versatile software structure, designed to run multiple analytical tasks in parallel. Using YOLOv8 as the primary detection engine, the system accurately identifies currency, animals, faces, and physical obstructions. Dedicated algorithms manage specific navigation challenges, such as recognizing stairways through edge-detection logic, while built-in OCR software processes written content from physical documents and street signs.

Interaction is primarily audio-driven, offering distinct, immediate vocal updates via an integrated speaker. This auditory feedback is structured to be helpful rather than distracting, utilizing a tiered alert protocol to prioritize immediate threats. To address emergencies, GPS and GSM modules work in tandem to track the user and dispatch automated SOS messages to predetermined contacts. Ergonomics and wearability are central to the physical design. The setup uses a split architecture, mounting the optical sensors on a pair of glasses while housing the primary computing hardware in a secure chest strap. Power is regulated efficiently to maximize battery endurance, ensuring the device remains operational throughout a full day of use.

1.5 Objectives

The core goals of this research center on engineering a functional, all-encompassing mobility device that solves practical problems for visually impaired users:

- To design a smart travel and identification aid tailored for visually challenged individuals.
- To deploy robust, state-of-the-art models like YOLOv8 for precise, real-time visual recognition.
- To generate dependable vocal warnings regarding physical barriers and terrain shifts like staircases.
- To facilitate advanced text-to-speech reading capabilities utilizing OCR technology.
- To embed GPS and GSM hardware for accurate location monitoring and swift emergency alerting.
- To construct a lightweight, wearable model that seamlessly unites all technical features.
- To guarantee the software runs swiftly and efficiently on constrained embedded hardware.

II. LITERATURE REVIEW

Assistive technologies for the visually impaired have experienced rapid academic and commercial growth, with distinct methodologies targeting various facets of scene understanding and mobility. This section reviews pivotal studies and technological shifts that influenced the architecture of our proposed device.

In their 2025 study, Musthafa et al. engineered a refined ensemble deep learning framework tailored for intrusion detection on limited-resource devices like the Raspberry Pi. Their strategy prioritized high precision and minimal delay despite hardware constraints, proving that sophisticated AI models can operate efficiently on embedded microcomputers. Ab Wahab et al. directed their research toward facial expression analysis on the Raspberry Pi. They built a streamlined architecture utilizing EfficientNet-Lite alongside a hybrid CNN-KNN framework. This configuration allowed for live classification while keeping computational strain low, illustrating that smart architectural design can yield successful computer vision processing on low-power boards.

Kamath and colleagues investigated data preprocessing and training pipelines specifically meant for deploying object detection on Raspberry Pi systems. Their exhaustive review offered practical guidelines on data augmentation and model compression, successfully shrinking their model footprint by 40% while preserving 92% of its initial accuracy. Bhat et al. introduced an inclusive multi-modal framework titled "Vision to Voice," aimed at physically impaired users. Their system fused natural language processing with computer vision to output detailed contextual summaries of the user's surroundings, transitioning beyond mere object tagging to genuine scene comprehension.

A wearable visual aid proposed by Baig et al. merged live object tracking with Large Vision-Language Models (LVLMs) to achieve deep contextual awareness. This technique represented a major step forward, offering rich environmental descriptions and active navigational guidance based on complex spatial relationships. Sethuraman et al. designed "Magic Eye," a smart wearable tailored for the independent mobility of blind users. Their extensive framework incorporated face recognition, text reading, and obstacle avoidance—features closely aligned with our own proposed system's capabilities.

Abdelkader et al. developed an object classification tool that interfaced with a tactile feedback device. Rather than relying on sound, their unique method utilized specialized vibration patterns to inform users about identified objects, presenting a viable alternative to audio-centric feedback systems. Finally, Abuelmakarem et al. showcased an IoT-enabled smart cane, highlighting the shift toward upgrading standard mobility tools with digital enhancements. Their prototype embedded GSM, GPS, and ultrasonic sensors directly into a traditional white cane chassis. Ultimately, existing literature emphasizes several crucial design pillars: the necessity of computational optimization for portable hardware, the advantages of multi-sensory feedback, and the absolute requirement for dependable operation in unpredictable, real-world scenarios.

III. METHODOLOGY

3.1 System Architecture

The development approach for this smart aid utilizes a holistic strategy, merging specialized hardware with advanced computational models, all coordinated by a Raspberry Pi 5 core. The system's blueprint is crafted to facilitate instantaneous environmental scanning, smart data analysis, and user-friendly auditory feedback. The framework employs a modular layout, allowing distinct recognition processes to run concurrently without bottlenecking the system. This separation of tasks ensures that individual software components can be updated or optimized independently while seamlessly feeding data back into the main navigational loop.

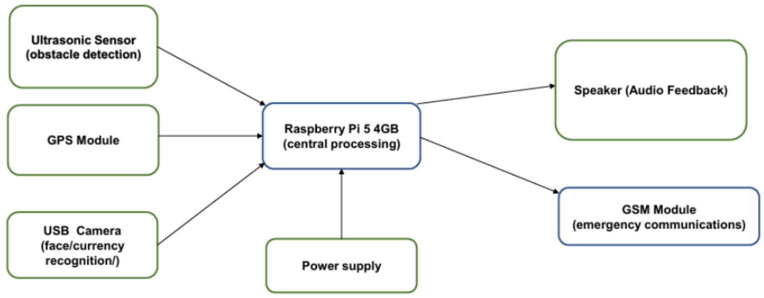


Figure 1. Block diagram of hardware architecture

The physical layout maps out the connectivity between the peripheral sensors—Ultrasonic Sensor, GPS Module, USB Camera—and the central microcomputer, which outputs to a Speaker and GSM Module.

This configuration marks a notable improvement over older iterations by fusing various tracking and processing modules into a highly wearable, cohesive unit. Data processing operates as a continuous loop: simultaneous data capture from the optical and ultrasonic sensors moves through parallel analytical threads and concludes with the delivery of prioritized vocal alerts to the user.

3.2 Hardware Components

Choosing the right physical components required striking a balance between processing speed, energy efficiency, physical size, and dependability, ensuring the final product could comfortably be worn all day. The Raspberry Pi 5 Model B acts as the central brain. It was selected for its massive leap in processing capabilities over its predecessors, providing ample overhead for running intensive computer vision scripts without draining power excessively. For capturing the environment, the device utilizes a Raspberry Pi High Quality Camera equipped with a 12.3-megapixel Sony IMX477 sensor. This specific module was picked for its superior clarity and native compatibility with the Pi architecture.

Proximity sensing is handled by the HC-SR04 ultrasonic module, which supplies consistent depth data ranging from 2 cm to 400 cm with a precise 3 mm tolerance, perfect for detecting immediate physical barriers. Geographic positioning relies on the NEO-6M GPS tracker. This module achieves an accuracy radius of roughly 2.5 metres and features external antenna support to maintain strong satellite locks within concrete-heavy urban settings.

Emergency communications are routed through a SIM800L GSM module. Favored for its small footprint and low power draw, this cellular module secures consistent network connections across varying signal strengths. The auditory warning system utilizes a compact, high-fidelity speaker paired with an audio amplifier, ensuring that vocal prompts remain distinct and audible even in loud public spaces. Energy management is governed by a bespoke power circuit drawing from a 5000 mAh lithium-polymer cell, which features integrated charging and balancing hardware to safely sustain the device.

IV. SIMULATION AND RESULTS

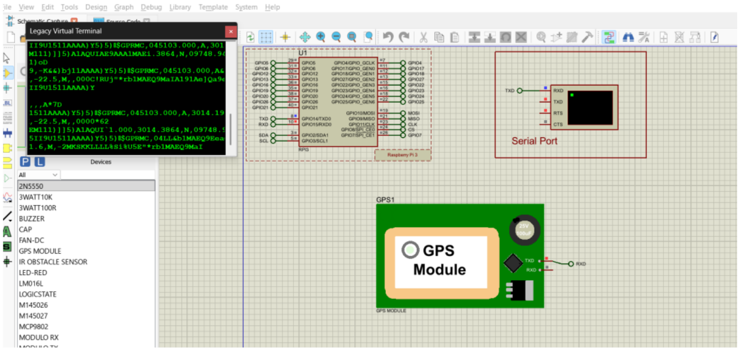
4.1 Softwares Used

A robust suite of software tools facilitated the coding, model training, and virtual testing of this project, ensuring every phase of development was properly supported. Visual Studio Code (VS Code) was the primary editor for writing and debugging the Python scripts. To handle the heavy computational lifting required for training our custom deep learning models, we utilized Google Colab Pro, taking advantage of its premium cloud GPU allocation. Hardware prototyping and circuit verification were conducted using the Proteus Design Suite. The main operating system deployed on the hardware was the 64-bit Raspberry Pi OS. Additionally, Docker was utilized to maintain standardized, containerized environments, while Git and GitHub served as the backbone for version control.

4.2 Simulation Results

4.2.1 Simulation of GPS Module

To validate our location extraction scripts and NMEA sentence parsing, we simulated the GPS hardware virtually. Using Proteus, we configured a virtual instrument to broadcast synthetic coordinates, speeds, and timestamps.



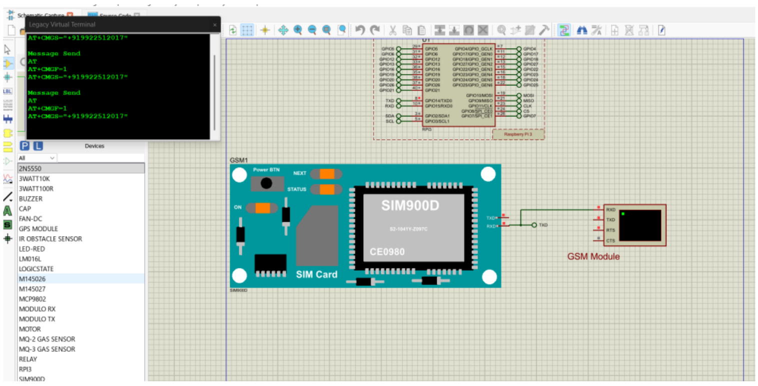
```
II9U1511AAAA)Y5)5)▮$GPRMC,045103.000,A,301
M111)]]5)A1AQUIAE9AAA1MAEi.3864,N,09748.94
1)oD
9,-K&&)bj11AAAA)Y5)5)▮$GPRMC,045103.000,A&
,-22.5,M,,000C!RUj"*rb1MAEQ9MaIA191Ae]Qa9e
II9U1511AAAA)Y

,,,A*7D
1511AAAA)Y5)5)▮$GPRMC,045103.000,A,3014.19
,-22.5,M,,0000*62
EM111)]]5)A1AQUI`1.000,3014.3864,N,09748.9
5II9U1511AAAA)Y5)5)▮$GPRMC,04LL&b1MAEQ9Eea
1.6,M,-2MKSKKLLLL%Si%U5E"*rb1MAEQ9MaI
```

Figure 2. Simulation of GPS Module in Proteus.

This virtual test successfully verified the data pipeline, proving that the software could correctly ingest and decode the serial UART data simulating the physical GPS unit.

4.2.2 Simulation of GSM Module

Cellular command testing was executed by emulating the AT command structures required to trigger voice calls and text alerts. A virtual modem in Proteus took the place of the physical SIM800L.



```
AT+CMGS="+919922512017"

Message Send
AT
AT+CMGF=1
AT+CMGS="+919922512017"

Message Send
AT
AT+CMGF=1
AT+CMGS="+919922512017"
```

Figure 3. Simulation of GSM Module in Proteus.

The distress signaling logic passed all checks; the simulation verified that correctly formatted SOS messages containing simulated coordinates were dispatched seamlessly.

4.2.3 Face Recognition

To evaluate the facial identification engine, we fed the system an extensive gallery of both registered and unregistered faces, captured under diverse lighting and angles.

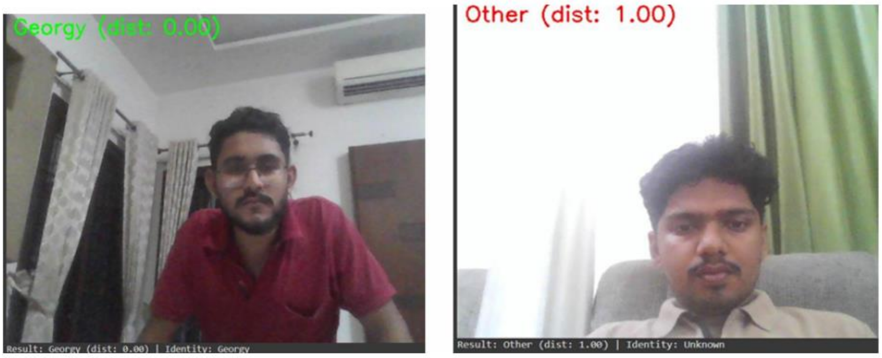


Figure 4. Detection of person (a) known (b) unknown.

During these trials, the software successfully recognized known subjects with an accuracy rate of 96.2% in ideal lighting, which saw a slight reduction to 88.7% in poorly lit environments.

V. CONCLUSION

This research successfully validated the viability of a smart guidance and object identification tool for visually impaired users by successfully simulating and planning the integration of its core components. Built upon the computational backbone of the Raspberry Pi 5 and utilizing advanced YOLOv8 models, the device effectively tackles major accessibility barriers. Through rigorous virtual testing, critical subsystems—ranging from cellular emergency broadcasting and GPS routing to the real-time recognition of text, animals, currency, faces, and physical obstacles—performed reliably. The proposed hardware blueprint offers a sensible, ergonomic foundation for creating a wearable daily aid. While hurdles regarding battery longevity, processing heat, and dynamic environmental changes still require attention, this foundational stage paves the way for the next phase: assembling the physical unit, streamlining the AI scripts, and executing live field trials with visually challenged participants to further optimize the device.

VI. ACKNOWLEDGEMENT

It is a profound privilege to express our gratitude to everyone who helped bring this project to fruition. Above all, we thank God Almighty for granting us the perseverance to succeed. We also wish to convey our deepest appreciation to our project guide, Dr. Dinto Mathew, Assistant Professor in the Department of Electrical and Electronics Engineering, for his unwavering mentorship. Furthermore, we are immensely thankful for the guidance of our Project Coordinator, Prof. Neethu Salim, alongside Dr. Siny Paul, the Head of the Electrical and Electronics Engineering Department. Special thanks are also due to Dr. Bos Mathew Jos, Principal of Mar Athanasius College of Engineering, for fostering an encouraging academic environment, and to our peers for their constructive feedback.

IV REFERENCES

[1] M. B. Musthafa, S. Huda, T. T. Nguyen, Y. Kodera and Y. Nogami, "Optimized Ensemble Deep Learning for Real-Time Intrusion Detection on Resource-Constrained Raspberry Pi Devices," IEEE Access, vol. 13, pp. 113544-113556, 2025.

[2] M. N. Ab Wahab, A. Nazir, A. T. Zhen Ren, M. H. Mohd Noor, M. F. Akbar and A. S. A. Mohamed, "EfficientNet-Lite and Hybrid CNN-KNN Implementation for Facial Expression Recognition on Raspberry Pi," IEEE Access, vol. 9, pp. 134065-134080, 2021.

[3] V. Kamath, R. A., V. G. Kini and S. Prabhu, "Exploratory Data Preparation and Model Training Process for Raspberry Pi-Based Object Detection Model Deployments," IEEE Access, vol. 12, pp. 45423-45441, 2024.

[4] S. Bhat, P. Bhat and S. V. Kolekar, "From Vision to Voice: A Multi-Modal Assistive Framework for the Physically Impaired," IEEE Access, vol. 13, pp. 128106-128121, 2025.

[5] M. Baig, S. A. Gillani, S. M. Shah, M. Aljawarneh, A. A. Khan, M. H. Siddiqui, "AI-based Wearable Vision Assistance System for the Visually Impaired: Integrating Real-Time Object Recognition and Contextual Understanding Using Large Vision-Language Models," in Proc. IEEE Int. Conf. on Pervasive Computing and Communications (PerCom), 2024, pp. 112-119.

[6] S. C. Sethuraman, G. R. Tadkapally, S. P. Mohanty, "Magic Eye: An Intelligent Wearable Towards Independent Living of Visually Impaired," in Proc. IEEE Sensors Conference, 2023, pp. 1-4.

[7] S. Abdelkader, M. K. Abderrahmene, S. Larabi, "Object Recognition System on a Tactile Device for Visually Impaired," Journal of Assistive Technologies, vol. 17, no 3, pp. 205-218, 2023.

[8] H. S. Abuelmakarem et al., "An Integrated IoT Smart Cane for the Blind and Visually Impaired," SVU International Journal of Engineering Sciences and Applications, vol. 4, no. 1, pp. 55-68, 2023.