

AI Based Aid with Integrated Reading for Completely Blind

Dr.N.Rajalakshmi

Department of Biomedical Engineering,
Dhanalakshmi Srinivasan College of
Engineering, Coimbatore.
praniraji1@gmail.com

Ms. Swathy K

Department of Biomedical Engineering,
Dhanalakshmi Srinivasan College of
Engineering, Coimbatore.
swathyraji2226@gmail.com

Priyadharshini P

Department of Biomedical Engineering,
Dhanalakshmi Srinivasan College of
Engineering, Coimbatore.
priyadharshini0661@gmail.com

Rathipriya M

Department of Biomedical Engineering,
Dhanalakshmi Srinivasan College of
Engineering, Coimbatore.
rathipriya0108@gmail.com

Anbarasan

Department of Biomedical Engineering,
Dhanalakshmi Srinivasan College of
Engineering, Coimbatore.
aanbarasan008@gmail.com

Elangathir G

Department of Biomedical Engineering,
Dhanalakshmi Srinivasan College of
Engineering, Coimbatore.
gelangkathirelangkathir@gmail.com

Abstract—This document presents the development of an assistive technology tool designed to enhance environmental awareness and information accessibility for visually impaired individuals using AI-based features. The system integrates a hands-free camera with advanced deep learning algorithms to enable real-time object detection, scene understanding, face recognition, and text recognition within the user's immediate surroundings. It also incorporates GPS functionality for location tracking.

Identified objects and extracted text are conveyed through multimodal feedback, including synthesized speech and tactile cues, providing a more comprehensive understanding of the environment. Additionally, the system includes an Optical Character Recognition (OCR) module that allows users to seamlessly scan and read documents, labels, and signs. By minimizing environmental challenges, this assistive solution empowers visually impaired individuals with greater independence, moving beyond the limitations of existing assistive technologies. With its

adaptable and flexible framework, the system can effectively recognize common objects and extract readable text, offering strong potential for further development and real-world applications.

Keywords - AI Based visual Aid, blind people, Deep learning GPS, Object detection, face detection , Audio feedback

I. INTRODUCTION

Vision loss affects millions of people worldwide, significantly impacting daily life and limiting independence. Individuals who are completely blind face substantial challenges in navigating unfamiliar environments, recognizing objects, and accessing written information. These challenges range from safely crossing busy streets to independently reading food labels or using public transportation. Traditional assistive tools such as white canes and Braille offer only limited support in addressing these issues. While white canes help detect obstacles, they do not provide information about object identity or scene context. Similarly, Braille is not

universally available and does not support real-time environmental awareness.

However, major advancements in artificial intelligence (AI), particularly in computer vision and natural language processing, have created opportunities for innovative solutions. This work introduces an AI-powered visual assistance system aimed at enhancing environmental awareness and readability for individuals who are totally blind. By integrating a wearable camera with deep learning algorithms, the system delivers real-time object detection, scene interpretation, and text recognition, enabling a more comprehensive and natural understanding of surroundings. Advances in convolutional neural networks (CNNs) support accurate and rapid object identification, while deep learning-based optical character recognition (OCR) effectively extracts text from images. Such technologies hold significant potential to improve the quality of life for blind individuals by fostering greater independence and smoother social integration.

The objective of this paper is to present a novel visual aid system designed specifically for blind users. The key features of the proposed system are summarized as follows:

1. The device is designed for seamless integration into daily life, as it can be mounted on a pair of glasses. It functions both as a reading assistant and a navigation guide, making it suitable for indoor and outdoor use.
2. The system achieves high algorithmic performance while operating on minimal hardware resources.
3. Accurate real-time distance estimation is accomplished using camera-based measurements, enhancing precision while minimizing system cost through reduced sensor dependency.

Currently, the proposed device is capable of detecting and tracking both moving and stationary objects, providing real-time audio feedback to describe the environment to the user. In addition, it includes an automated reading assistant that reads printed documents aloud. The design, development, and performance evaluation of this visual aid system are discussed in the subsequent sections of this paper. Section II reviews the existing literature, Section III outlines the hardware implementation, Section IV explains the software components, and Section V describes the overall design of the proposed system.

II. RELEVANT WORKS

The evolution of assistive technologies for individuals with visual impairments has a long and progressive history, with artificial intelligence playing a transformative role in recent years. Early developments primarily focused on sensory substitution methods, such as tactile displays and auditory signals, to transmit basic environmental information. For example, the Tactile Vision Substitution System (TVSS) employed a camera to convert visual input into tactile stimuli delivered to the user's back. Although groundbreaking at the time, such systems were often bulky and challenging for users to master.

With advancements in technology, more recent research has shifted toward computer vision-based approaches for

object detection and scene interpretation. Numerous studies have explored wearable visual assistance devices that integrate cameras with processing units. Devices such as OrCam My Eye apply computer vision techniques for text recognition and object identification, delivering information through audio feedback. While this technology has demonstrated effectiveness in tasks like reading and facial recognition, it often falls short in providing comprehensive scene analysis and multimodal feedback beyond speech. Most existing systems focus on identifying individual objects rather than offering a complete contextual understanding of the surrounding environment.

Alternative approaches have concentrated on specific functionalities, such as navigation support through GPS and obstacle detection using image processing techniques. Systems like the NavBelt and various smart canes primarily assist with obstacle avoidance but do not provide detailed contextual information about the environment. For instance, the NavBelt employed ultrasonic sensors to identify obstacles and alert users through tactile feedback; however, it did not classify detected objects or describe the broader surroundings. Likewise, standalone OCR devices are available for reading printed text, but they typically operate independently from environmental awareness systems, requiring users to switch between different devices or operational modes.

The incorporation of deep learning has considerably enhanced the functionality of visual assistance systems. Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in object recognition, enabling real-time identification of diverse

objects. Detection frameworks such as YOLO and Faster R-CNN have played a pivotal role in achieving high detection accuracy and processing speed. Additionally, Recurrent Neural Networks (RNNs) and Transformer-based models are being investigated for advanced scene interpretation and the generation of more natural language descriptions. These architectures allow systems to analyze relationships between objects and deliver contextualized information—for example, describing “a person walking on a sidewalk” rather than separately identifying “person” and “sidewalk.”

The importance of multimodal feedback has also been widely acknowledged in enhancing user experience. Integrating auditory and tactile signals provides a more comprehensive and intuitive understanding of surroundings compared to relying on a single sensory channel. Ongoing research in haptic interfaces and spatial audio techniques aims to communicate spatial attributes and object characteristics more effectively. For example, tactile cues can indicate an object’s direction and distance, while spatial audio can convey its relative position within the environment.

Building upon these developments, the present work proposes an integrated system that unifies real-time object detection, detailed scene understanding, and embedded text recognition within a single platform. Moreover, it incorporates multimodal feedback mechanisms—such as synthesized speech and tactile signals—to deliver a more complete and accessible experience for individuals who are completely blind. This approach seeks to overcome the limitations of existing solutions by providing a comprehensive

system for both environmental awareness and information accessibility.

III.HARDWARE IMPLEMENTATION

A.Node Microcontroller(ESP8266)

An example Microcontroller is the ESP8266. It is a System-on-a-Chip (SoC) which integrates a 32 bit processor, RAM, WiFi and even a basic operating system, making it ideal for Internet of Things (IoT) applications. The Node MCU is an open-source development platform based on the ESP8266 SoC. It supports programming like Arduino and it allows interfacing with sensors, external devices through the GPIO pins.



Figure 1 Node microcontroller (ESP8266)

B. Ultra sonic sensor

Ultrasonic sensors are non-contact devices used to detect objects or measure their distance by transmitting high-frequency sound waves and analyzing the reflected echo. Their operating principle is similar to that of radar and sonar systems.



Figure 2 Ultra Sonic Sensor

The distance to an object is determined by measuring the time interval between the transmission of an ultrasonic signal and the reception of its echo. This pulse-echo principle enables the sensor to

emit and receive ultrasonic waves efficiently. Ultrasonic sensors operate using piezoelectric crystals, which convert electrical energy into ultrasonic sound waves and transform the returning echoes back into electrical signals.

These sensors are highly reliable under varying environmental conditions, as their performance is not significantly affected by factors such as weather, target color, transparency, or lighting. Owing to their accuracy and adaptability, ultrasonic sensors are widely used for detecting solid and liquid targets in applications such as industrial automation, robotics, and level measurement, often offering advantages over optical sensing technologies.

C.GPS (NEO-6 MODULE)

The Global Positioning System (GPS) is a satellite-based navigation system originally developed by the United States Department of Defense. It operates using a constellation of approximately 24 to 32 satellites positioned in medium Earth orbit, providing precise time and location information to users worldwide under all weather conditions. GPS determines a user's position by calculating the time taken for signals transmitted from multiple satellites to reach the receiver and applying triangulation techniques.

The system consists of three main segments: the space segment (satellites), the control segment (ground monitoring stations), and the user segment (receivers). Although initially intended for military applications, GPS is now widely utilized across civilian, commercial, and scientific fields.

This work also highlights the NEO-6 GPS module, known for its compact design and rapid signal acquisition. It is compatible with platforms such as Arduino and

Raspberry Pi, making it well-suited for portable and embedded system applications.



Figure 3 (GPS) Module

D. 4 Channel Iic I2C Logic Level Converter Bi-Directional 5V To 3.3V ET5244

The ET5244 is an ultra-portable bi-directional logic level converter that enables effortless communication between devices with contrasting voltages such as 5V and 3.3V, this makes it very useful in the world of microcontrollers. Its application is geared towards interfacing a microcontroller, for example an ESP8266 or arduino, with various sensors or displays that operate on a different logic level that requires a different level of logic. This module supports up to four independent channels, making it easy to use with many devices, modules or sensors at one time

E. PC or Laptop

The laptop or PC acts as the core processing unit of the system, responsible for executing and coordinating intelligent functions. It handles computationally intensive tasks such as image processing, object detection and recognition, facial

recognition, and optical character recognition (OCR). These processes require significantly greater processing power than microcontrollers like the ESP8266 can provide, making the laptop essential for real-time analysis and decision-making.

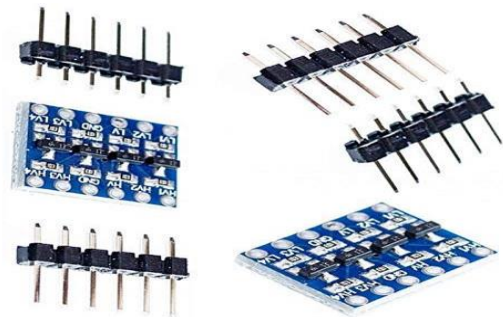


Figure 4 Channel Iic I2C Logic Level Converter Bi-Directional 5V To 3.3V ET5244

In addition, the laptop functions as a development and debugging platform. It facilitates code development, testing, and deployment to peripheral modules such as the ESP8266 using tools like Anaconda Navigator, Spyder IDE, and the Arduino IDE. Furthermore, it enables communication with sensors and microcontrollers through USB interfaces, utilizing technologies such as PySerial for transmitting commands and receiving data.

IV.SOFTWARE IMPLEMENTATION

A. Anaconda Navigator

The Anaconda distribution includes **Anaconda Navigator**, a graphical user interface (GUI) designed to simplify the management of Python environments and packages. It allows users to monitor and control applications, conda environments, and required packages without relying on command-line operations.

B. Spyder IDE

Spyder is a free, open-source, multi-language development environment designed primarily for scientific computing in Python. Projects such as a Blind Assistance System can significantly benefit from this IDE, as it supports widely used scientific libraries including NumPy, Pandas, Matplotlib, and Scikit-learn for data analysis and processing.

Spyder offers an intuitive interface equipped with powerful, extensible tools such as the IPython console, variable explorer, real-time debugging, and profiling features, enabling efficient and interactive coding. It is commonly used alongside Anaconda and can interface with platforms like Arduino and Raspberry Pi, making it well-suited for developing smart sensor-based assistive technology systems.

C. Arduino IDE

Arduino IDE (Integrated Development Environment) is the software platform used to write, compile, and upload programs to Arduino-compatible boards. It features a user-friendly interface that allows individuals with different levels of experience to easily program and interact with hardware. The IDE supports coding in C/C++ through its customized dialect tailored for microcontroller applications.

The environment includes a built-in code editor, compiler, serial monitor for USB communication, and tools for uploading programs to the board. In addition to standard Arduino boards, it also supports other microcontrollers such as the ESP8266 and ESP32 through additional board manager packages. With its extensive library support, the Arduino

IDE enables seamless interfacing of microcontrollers with various electronic components, including sensors, actuators, motors, and displays. Tasks such as reading sensor data, controlling motors, and implementing wireless communication can be performed efficiently. Overall, the Arduino IDE is a valuable tool for embedded system development and interactive hardware applications, as it facilitates rapid prototyping and testing.

V. DESIGN AND OUTPUT OF THE PROPOSED SYSTEM

A. Object Detection Module

Real-time identification of objects in the user's immediate surroundings is achieved using deep learning models such as MobileNet, which is lightweight and optimized for edge-based applications. Live video frames captured by the camera are transmitted to the laptop, where they are processed to detect furniture, household items, obstacles, and signboards.

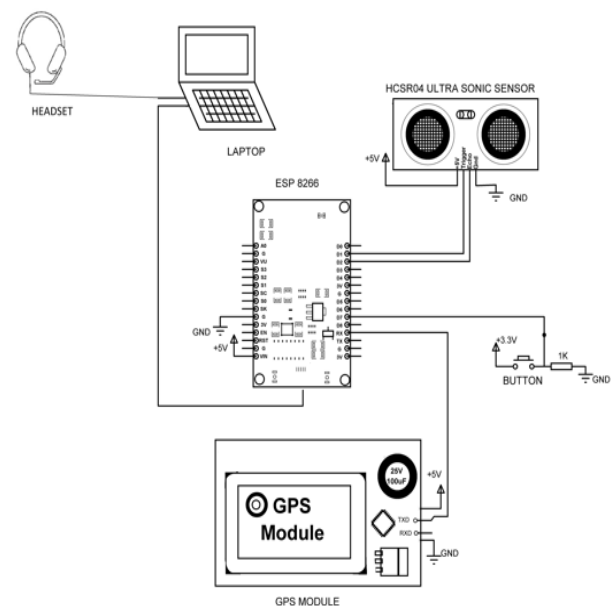


Figure 5 Proposed Setup

The system integrates OpenCV and TensorFlow for image pre-processing, object detection, and classification into predefined categories. Once an object is identified, it is highlighted and conveyed to the user through audio feedback. This module enhances the user's spatial awareness, enabling them to avoid obstacles, navigate safely, and gain a clearer understanding of their surrounding environment.

B. Face Recognition Module

Social awareness is enhanced through the automatic detection and voice announcement of familiar individuals present in the user's vicinity. This functionality utilizes deep learning algorithms to extract facial features and compare them with a pre-trained database of known faces. A webcam captures the images, which are then processed using face embedding techniques and classification models for accurate identification. Once a person is recognized, their name is announced to the user through headphones or speakers. This feature improves social interaction while also adding a layer of security by clearly distinguishing between known and unknown individuals.

C. Optical Character Recognition

The OCR module enables users to read text from signs, books, and other printed materials. It utilizes PyTesseract, a Python wrapper for the Tesseract OCR engine, to extract textual content from captured images. Before text extraction, the images undergo preprocessing steps such as grayscale conversion, thresholding, and noise reduction to enhance recognition accuracy.

The extracted text is then converted into speech using text-to-speech libraries such as Google Text-to-Speech (gTTS) or pyttsx3. This functionality allows visually impaired users to independently access and understand written information in their surroundings.

D. Obstacle Detection Module

This module maintains up-to-the-second awareness of physical obstacles using an ultrasonic sensor (HC-SR04). This sensor is able to measure the distance between the user and nearby objects through the emission of ultrasonic pulses and measuring the time taken for the echo to return. A microcontroller (ESP8266) running the algorithms of this module processes this information and provides an audible warning when an obstacle within controllable detection range is sensed. This assists users in avoiding collisions with ultra-low lying and hard to detect objects which are usually missed by white canes.

E. GPS and Emergency Alert Module

In emergency situations, the user can activate an emergency button that triggers the NEO-6M GPS module. The module obtains the user's current location coordinates, which are then transmitted to a caregiver through a mobile application using the ESP8266 microcontroller and the Blynk IoT platform.

This feature plays a crucial role in ensuring user safety by enabling rapid assistance during critical situations. It also increases user confidence, especially when traveling in unfamiliar environments.

F. Audio Feedback Module

This module functions as the main output interface between the system and the user. All identified text, objects, faces,

and alerts are converted into audio using text-to-speech (TTS) engines such as gTTS or pyttsx3. The generated audio is delivered through headphones or a compact speaker integrated into the wearable device.

Real-time voice feedback is essential for assisting users in navigating their surroundings and communicating system information without the need for visual or tactile interaction.

VI. CONCLUSION

The proposed AI-Based Visual Aid with Integrated Reading Assistant presents an effective and practical solution to enhance the independence of completely blind individuals. By combining deep learning techniques with embedded hardware, the system enables real-time object detection, face recognition, and text-to-speech conversion, thereby improving environmental awareness and accessibility to information. The integration of ultrasonic sensors and GPS further strengthens the system by ensuring safe navigation and providing emergency support when required.

Unlike traditional assistive tools, the proposed system offers a unified platform that delivers contextual and meaningful feedback through audio, allowing users to better understand their surroundings. The implementation demonstrates reliable performance with efficient resource utilization, making it suitable for real-world applications.

In conclusion, this work highlights the significant potential of AI-driven assistive technologies in empowering visually impaired individuals. Future improvements such as device miniaturization, enhanced

accuracy, and multilingual support can further increase its usability and impact, contributing to a more inclusive society.

REFERENCES

- [1].N. A. NishaJith, J. Nivedha, S. S. Nair, and M. S. J., “Smart cap – Wearable visual guidance system for the blind,” 2018.
- [2].S. Patil, K. Narute, M. Jadhav, and A. Chougule, “Blind reader using Raspberry Pi,” 2019.
- [3].G. S. Kumar, N. Ramesh, and V. Teju, “Assistive technologies for blind people to describe indoor sensing by Raspberry Pi,” 2017.
- [4].M. T. Islam, M. Ahmad, and A. S. Bappy, “Real-time family member recognition using Raspberry Pi for visually impaired people,” 2020.
- [5].D. P. Khairnar, R. Karad, A. Kapse, and G. Kale, “PARTHA: A visually impaired assistance system,” 2020.
- [6].M. A. Khan, P. Paul, M. Rashid, M. Hossain, and M. A. R. A. Ahad, “An AI-based visual aid with integrated reading assistant for the completely blind,” 2020.
- [7].M. Poggi and S. Mattoccia, “A wearable mobility aid for the visually impaired based on embedded 3D vision and deep learning,” 2016.
- [8].R. Uma, P. S. Harish, N. G. Ragav, and A. Navin, “Third eye – A smart wearable glass with deep learning technology powered with artificial intelligence for visually impaired people,” 2023.
- [9].H. AlSaid, L. AlKhatib, A. AlOraidh, S. AlHaidar, and A. Bashar, “Deep learning assisted smart glasses as educational aid for visually challenged students,” 2016.

