Mr. D. Nagaraju, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Real-Time Visual Assistance For The Visually Impaired Using Esp32-Cam

Mr. D. Nagaraju[1], K.Balaswamireddy[2], V. Jaya Vishnu[3], B.Balaji[4], U.Ganesh[5]
Assistant Professor[1],UG Students [2,3,4,5],Department of Electronic & communication Engineering
Sai Rajeswari Institute of Technology, Proddatur, Andhra Pradesh, 516306

Abstract- IoT based smart glasses system is designed for real-time environmental assistance and navigation for visually impaired people. The system is implemented using an esp32 microcontroller, integrating a camera module and a microphone for accurate scene capture and voice command processing. The camera continuously monitors the environment and converts visual signals into digital data from precise processing. Object, text and obstacles by analyzing the captured images through a cloud based AI platform. The esp32-cam uploads image and voice prompt data to the cloud platform through WiFi. Alerts recording detected objects, read text, and sense descriptions are sent to the user through an audio speaker using text-to-speech technology. The smart assistance system enhances user safety, improves navigation and provides independent living capabilities, making it cost effective and reliable solution for modern visually impaired users.

Keywoeds-ESP32 Microcontroller, Camera module, OpenAI vision, Text-to- Speech, Obstacle detection, Visual assistance

I. INTRODUCTION

With the rapid development of internet of things (IoT) and artificial intelligence technology, assistive systems have become more intelligent and efficient. Traditional walking canes provide approximate tactile feedback and lack environment awareness features. This system do not support remote AI processing, object detection or text reading. Lack of sense understanding leads to safety risks and inconvenience to visually impaired individuals. To overcome this limitations, this project proposes an Iot based smart glasses system that provides real-time scenes description and object detection. The proposed system uses esp32 microcontroller, camera module, microphone and cloud AI database to ensure continuous monitoring. By integrating speech recognition and text- to-speech platforms, users can receive instant voice alerts. This system improves independent navigation and environment awareness.

II. PROBLEM STATEMENT

Conventional assistive systems provide limited physical feedback and lack intelligent features. Visually impaired users must manually navigate environment cannot detect obstacles or read text. This system do not also support data processing and real-time descriptions. Navigation and object identification are major problem for blind individuals, leading to lack of independence. High-end AI smart glasses are expensive and not affordable for all users. Therefore, this is a need for low cost, reliable and smart assistive system. This Project aims to develop Iot based sense understanding and assistance system using esp32 and Open AI Chat GPT vision that offers accurate environmental descriptions, real time voice alerts, and interaction at an affordable cost.

III. METHODOLOGY

The proposed system is developed using an ESP32 microcontroller, Camera module, microphone and OpenAI ChatGPT vision. The camera module is mounted on the glasses to capture the scene

continuously when triggered. The esp32 reads the visual data and converts into digital image formats. Voice prompts are collected through the microphone and user commands. The processed image and voice data are transmitted to the ChatGPT vision using Wifi. The visual data is analyzed to detect objects, read text, and understand the environment. When the scene is analyzed, instant audio alerts and descriptions are sent to the user through a speaker using Text-to speech.

IV. BLOCK DIAGRAM

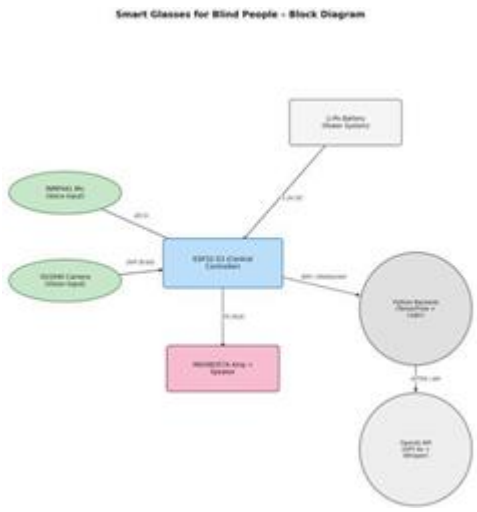

Figure: Block Diagram

V. COMPONENTSN USED

Esp32 Camera Module

The camera module captures the visual environment in front of the user as the scene changes, the image is converted into a digital signal. The controller process the data and transmits it to the ChatGPT vision. It helps in real- time monitoring and object detection.

Figure: ESP32 Camera Module

INMP441 Microphone

The microphone is used to capture the real-time voice commands of the user. It receives the audio signal from the user and converts them into text prompts. The voice data is sent ot the esp32 through a serial communication. This helps the user ask specific questions about their surroundings. It also supports interactive the scene understanding.

Figure: INMP441 Microphone

MAX98357 Amplifier Module

The MAX98357 is used to provide audible feedback to the user. It converts text descriptions generated by the google Text-to-speech into clear speech. This ensures the visually impaired user receives self and reliable real-time spoken descriptions of their environment without needing visual displays.

Figure: MAX98357 Amp Module

Openai Chatgpt Vision Ai

The ChatGPT vision AI is used to process image data, voice prompts, and scene and scene analysis. It receive highly accurate descriptions anytime through the internet. This helps in understanding the environment and improving user safety.

Figure: Chatgpt Vision

Power Supply

The system is powered using vehicle battery or rechargeable battery module. Voltage regulators are used to provide stable power to components.

Figure: Power Supply

VI. RESULT

VII. CONCLUSION

The developed system successfully captures the surrounding environment in real-time accurate audio descriptions. Objects and text are identified based on ChatGPT vision analysis. Voice interaction allows users to ask specific questions around their surroundings. Data is processed in cloud and delivered a speech feedback. The system operates reliably under normal daily conditions and provides timely alerts to users.

REFERENCES

1. A. Smith et AL., "Wearable Assistive Devices for the Visually Impaired Using IoT", IEEE Access, vol. 11, oo. 11234- 11245, 2023.
2. J.Doe and R. Roe, "Real-Time Object Detection and Scene Understanding for blind Navigation, " IEEE Transactions on Autonomous Science, vol. 18, no. 5, pp. 22107-2115, 2022.
3. S. Kumar et al., "Cloud-Based AI Integration in Wearable Systems for Visually Impaired People,"International Conference on smart systems, pp. 98-104, 2024.
4. M.patel and K.Desai, "Low Cost IoT Solutions for Real-Time Text to Speech Conversion,"

International Journal of IoT and Embedded Systems, vol. 14, no.2, pp. 78-89, 2023.
5. L.Zhang, "speech Interaction Interfaces in Wearable Devices,"IEEE Sensors Journal, vol. 23, no. 8, pp. 9245-9258, 2023
6. M. Patel and K. Desai, "Cloud-Based Analytics for Vehicle Fleet Management Systems," IEEE Transactions on Vehicular Technology, vol. 72, no. 4, pp. 4532–4545, 2023.