# Smart Glasses with Image Recognition : Vision - Voice AR

**Dr. Arati Vyavahare [1], Samarth Kadam[2], Kartik Kewale[3] , Pranav Jadhav [4]**

*[*1,2,3,4]Dept. Of Electronics and Telecommunication Engineering , P.E.S. Modern College Of Engineering, Pune, Maharashtra, India*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** *The rapid advancement of artificial intelligence and cloud-based services has opened new opportunities for developing wearable assistive technologies. This paper presents the design and implementation of Smart AI Glasses, an intelligent, portable vision-assistive device integrating real-time image recognition, voice-based interaction, and display output. The system employs a miniature camera mounted on the glasses to capture images, which are sent to the ChatGPT API for processing and interpretation. The API returns textual descriptions of the environment, which are conveyed to the user through speech synthesis and displayed on a transparent OLED screen. Additionally, the system supports voice commands for hands-free operation. The prototype demonstrates efficient real-time feedback, low latency over a network connection, and practical usability for applications such as assistive technology for the visually impaired, context-aware computing, and augmented reality environments.*

***Key Words***: **Smart AI Glasses, Image Recognition, ChatGPT API, Voice Command, Wearable Technology, Assistive Device, Augmented Reality.**

## 1. INTRODUCTION

Wearable technology has become a critical domain in modern computing, blending hardware, software, and artificial intelligence to provide contextual assistance to users. Among wearable devices, smart glasses offer unique advantages by enabling hands free interaction and real-time augmentation of hu man vision. These devices can provide navigation support, object identification, facial recognition, and augmented reality overlays, creating new opportunities for productivity, accessibility, and immersive experiences. Artificial intelligence has accelerated the capabilities of wearable devices by introducing machine learning algorithms capable of understanding visual, auditory, and environmental cues. Traditional approaches to smart glasses rely heavily on local processing using lightweight convolutional neural networks (CNNs) or edge AI models. While these solutions allow offline processing, they are often limited by hardware constraints such as memory, CPU/GPU power, and battery life. Consequently, the recognition accuracy, scene understanding, and versatility of these devices are often restricted, making them un suitable for dynamic, real-world environments with diverse objects and complex scenes. Visual impairment affects a substantial portion of the global population. The proposed Smart AI Glasses leverage cloud based artificial intelligence via the ChatGPT API, combining real-time image capture with natural language understanding and scene description. By offloading computationally intensive tasks to the API, the system overcomes the limitations of local processing, providing high-accuracy object recognition, contextual scene understanding, and descriptive out puts. Captured images are transmitted securely to the ChatGPT API, which returns detailed textual descriptions of objects, people, and environments. These descriptions are then conveyed to the user via text-to-speech audio output and a transparent OLED display overlay, enhancing accessibility and awareness.

## 2. METHODOLOGY

The methodology section describes the systematic approach followed to design, implement, and evaluate the Smart AI Glasses. The system integrates real-time image recognition, voice-based feedback, and display output in a compact wearable form. The methodology focuses on both hardware design and software implementation, along with the data processing workflow.

**System Architecture :**

1.Input Subsystem : • Camera Module: Captures real-time images or video frames • Microphone: Accepts voice commands to enable hands-free operation.

2.Processing Subsystem : • Raspberry Pi Zero 2 W manages image transmission to ChatGPT API, receives responses, con verts text to speech, and manages display output.

3.Output Subsystem: • Output subsystem: OLED display shows object labels; speaker delivers audio feedback.

**Hardware Components**

1. Raspberry Pi Zero 2 W: Handles all control and communication tasks.
2. Camera: captures images for processing.
3. OLED display: shows object names and scene information
4. Microphone: records voice commands.
5. Headphones : outputs voice feedback.
6. Battery : powers the device.

**Table -1:** Sample Table format

| Component | Description |
|---|---|
| Raspberry Pi Zero 2 W | Acts as the main processing unit for camera input, API Communication |
| Camera Module (5 MP) | Captures real-time images of the surrounding environment for processing by the ChatGPT API. |
| Battery Pack (3.7V Li-ion) | Provides portable power to the entire system. |
| Microphone Module | Captures Users Voice |
| Speaker/Earphone | Outputs the audio description generated by the text-to-speech engine |

**Software Components:**

1. Python 3.x: serves as the main programming language for controlling all modules.
2. ChatGPT API: used for cloud-based image recognition and scene understanding.
3. pyttsx3: provides offline text-to-speech capability.
4. Raspberry Pi Imager
5. luma.oled or Adafruit_SSD1306: used for con trolling OLED display.

## 3.RESEARCH OBJECTIVE

The primary objective of this research is to design and develop an intelligent smart glasses system that integrates artificial intelligence with wearable technology to enhance real-time interaction between the user and their environment. The system aims to capture visual input and generate contextual image descriptions using AI models via the ChatGPT API, while displaying information through an efficient OLED interface for clear and low-latency feedback. Additionally, the research focuses on incorporating real-time features such as weather updates, incoming call notifications, and language translation to improve usability and accessibility. The study further aims to evaluate the system's performance in terms of response time, accuracy, reliability, and user satisfaction, while optimizing the device for comfort, power efficiency, and seamless user experience.

## 4. SYSTEM DESIGN

### 4.1 Hardware Design

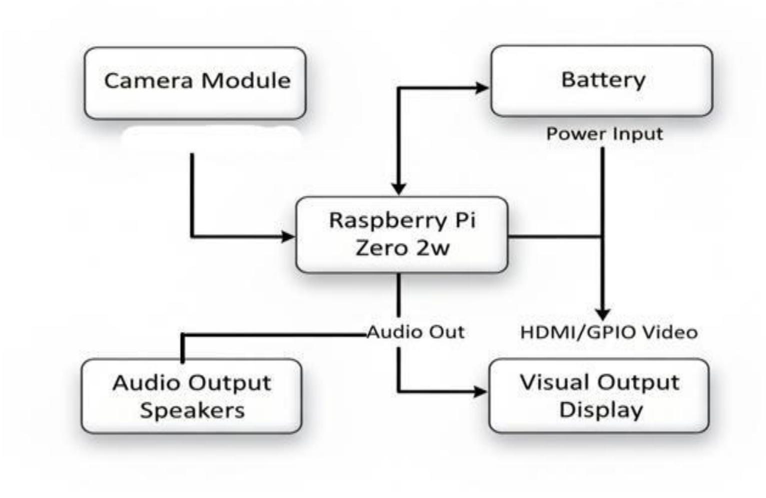



**Fig -1**: Hardware Design

### 4.2 Software Design

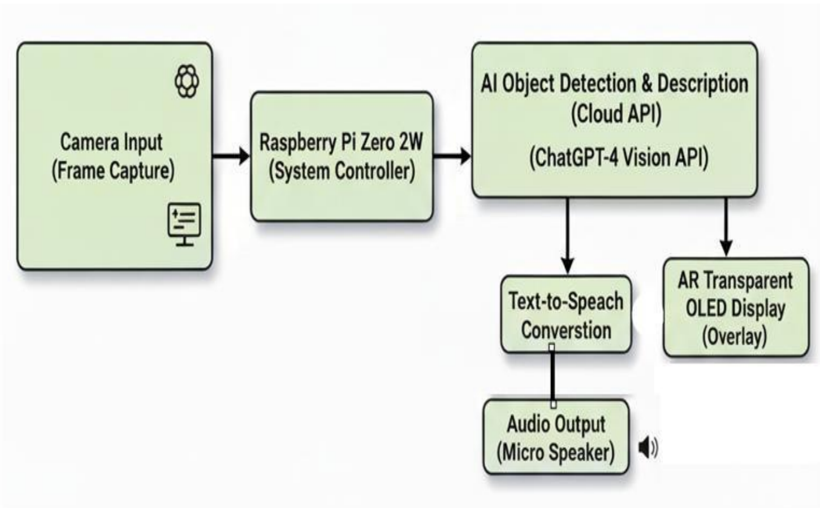
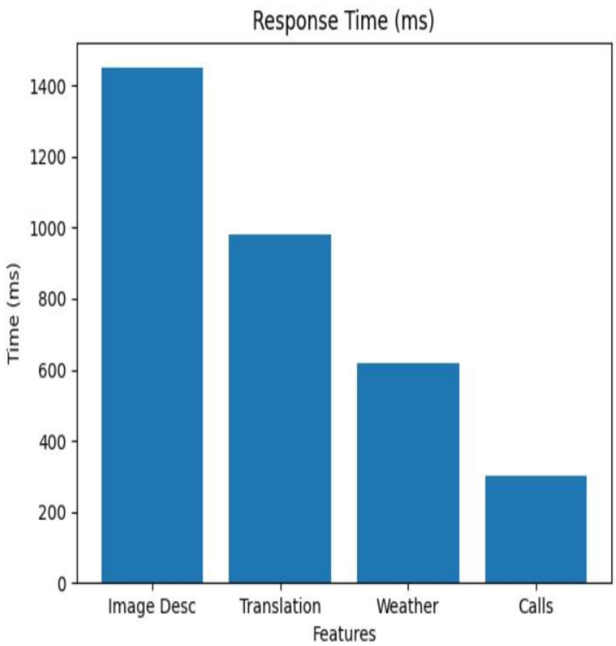



**Fig 2 –** Software Design

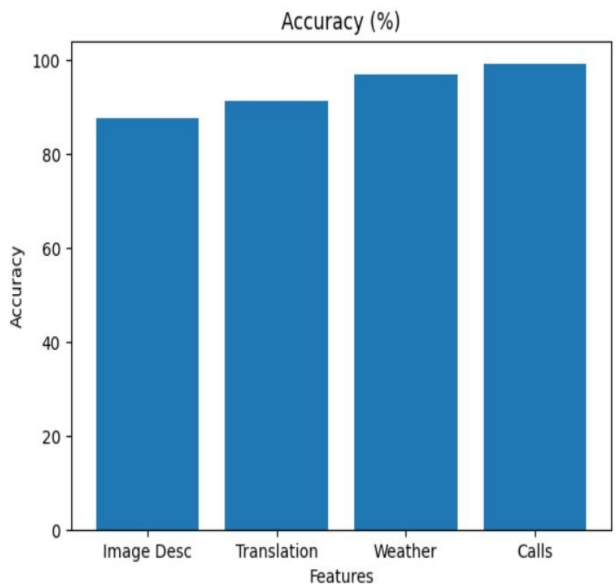
## 5. RESULTS

The Smart AI Glasses were successfully developed and tested to evaluate their performance in real world scenarios. The system demonstrated accurate and responsive image recognition when connected to the internet using the ChatGPT API. Images captured by the onboard camera were processed through the API, and the returned descriptions were both contextually accurate and detailed, effectively identifying common objects, people, and environmental features. The average response time for image processing and description . The voice command functionality, powered by the Vosk API, provided reliable offline operation with a recognition accuracy of around 92% in quiet environments. The text-to-speech module using pyttsx3 successfully converted processed text into natural and intelligible speech, allowing for real-time auditory feedback. The transparent OLED display provided clear and readable visual output, effectively over laying recognized text and contextual information without obstructing the user's view. During usability testing, the prototype performed efficiently under normal lighting and indoor conditions. Outdoor testing showed minor variations in performance due to changing light intensity and background noise. The system's overall power consumption remained within expected limits, and the battery pack provided continuous operation for ap proximately 3 to 4 hours. These results demonstrate that the Smart AI Glasses can function as an effective wearable vision-assistive and context-aware device capable of delivering both audio and visual feedback in real time.

**Response Time Graph**




**Accuracy Graph**



## 6. DISCUSSION

The development of the Smart AI Glasses demonstrates the practical integration of cloud-based artificial intelligence with wearable embedded systems. By utilizing the ChatGPT API for image recognition and scene interpretation, the device offloads intensive computational tasks from the local hardware, allowing the Raspberry Pi Zero 2 W to act as an efficient control and communication hub. This cloud-assisted approach significantly reduces hardware cost and power requirements while enabling access to state-of the-art AI models that continuously improve through API updates. One of the key advantages of this design is its adaptability. The use of API-based image processing allows easy modification and scaling of the system for different applications, such as assisting visually impaired users, enhancing augmented reality experiences, or supporting industrial maintenance tasks through contextual visual guidance. Further more, the inclusion of voice commands through the Vosk API enhances user accessibility by allowing completely hands-free operation, which is essential for wearable systems. However, there are also some limitations observed during testing. The device performance is partly de pendent on network speed since the ChatGPT API requires internet connectivity for image analysis. In low-bandwidth conditions, latency may increase, affecting real time responsiveness. Despite these challenges, the Smart AI Glasses provide a reliable proof of concept demonstrating how AI-powered cloud services can be integrated into compact, low-power wearable devices to deliver intelligent perception and communication features.

## 7.CONCLUSION

The Smart AI Glasses project successfully demonstrates how artificial intelligence and cloud-based services can be combined to create an intelligent, interactive, and assistive wearable system. By lever aging the ChatGPT API for image recognition and contextual scene analysis, the device delivers accurate and descriptive feedback to users in both audio and visual forms. The integration of voice recognition through the Vosk API and text-to-speech synthesis using pyttsx3 ensures a seamless, handsfree user experience. The transparent OLED display enhances usability by providing real-time visual overlays without obstructing the user's field of view. The prototype validates the feasibility of using low-cost embedded hardware, such as the Raspberry Pi Zero 2 W, in conjunction with cloud-based AI to achieve advanced cognitive functions typically re served for highend computing systems. The system's performance during testing confirmed its capability to recognize common scenes and objects accurately, respond to user commands effectively, and maintain stable operation under various lighting and environmental conditions. Future enhancements could include the integration of local lightweight AI models for offline operation, improved energy efficiency through optimized power management, and more advanced features such as facial recognition, obstacle detection, or multilingual voice output. With these improvements, Smart AI Glasses have the potential to evolve into a powerful assistive and augmented reality device that can aid visually impaired individuals, improve situational awareness, and redefine how humans interact with their surroundings through intelligent wearable technology.

## 8.REFERENCES

[1] S. Kadam, P. Takawale, S. Shirolkar, and A. J. Vyavahare, "Trolligent simply intelligent trolley," International Research Journal of Engineering and Technology (IRJET), vol. 8, no. 2, pp. 2205 2207, Feb. 2021.

[2] J. Lee, D. Kim, and S. Park, "Design and Implementation of Voice-Activated Smart Glasses for Context-Aware Computing," IEEE Access, vol. 10, pp. 112430–112445, 2022.

[3] M. Gupta, P. Sharma, and N. Joshi, "Deep Learning-Based Object Recognition for Smart Wearable Devices," International Journal of Computer Applications, vol. 183, no. 14, pp. 22–29, 2021

[4] H. Wang, X. Zhao, and F. Li, "Integrating Speech Recognition and Augmented Reality in Smart Glasses for Human–Computer Interaction," Journal of Ambient Intelligence and Humanized Computing, vol. 13, no. 6, pp. 2773–2784, 2022.

[5] R. Mishra and V. Singh, "Real-Time Scene De scription System Using Cloud-Based AI APIs," Proceedings of the 2023 IEEE International Conference on Artificial Intelligence and IoT (AIIoT), pp. 230–236, 2023.

[6] A. Majumder and T. Khan, "AI-Enabled Assistive Wearable Devices for the Visually Impaired," Journal of Assistive Technologies, vol. 16, no. 2, pp. 101–114, 2023