IV Trimester MCA Specialization project report

Department of Computer Science

ASSISTIVE WEARABLE DEVICE FOR SPECIALLY ABLED INDIVIDUALS USING GENERATIVE AI AND IOT

by

Kenneth Dominic Fernandes (2347231)

Mahua Saikia (2347235)

Nayan Raj (2347240)

Under the guidance of

Dr. Prabu P

JULY 2024

CERTIFICATE

This is to certify that the report titled ***ASSISTIVE WEARABLE DEVICE FOR SPECIALLY ABLED INDIVIDUALS USING GENERATIVE AI AND IOT*** *is a bona fide record of work done by Kenneth Dominic Fernandes(2347231), Mahua Saikia(2347235) and Nayan Raj(2347240) of CHRIST(Deemed to be University), Bangalore, in partial fulfillment of the requirements of IVth Trimester MCA during the year 2023.*

Head of the Department **Project Guide**

Valued-by:

1.

Name : Kenneth Dominic Fernandes
Register Number : 2347231
Examination Centre : CHRIST (Deemed to be University)

2.

Date of Exam :31.08.2024

CERTIFICATE

This is to certify that the report titled ***ASSISTIVE WEARABLE DEVICE FOR SPECIALLY ABLED INDIVIDUALS USING GENERATIVE AI AND IOT*** *is a bona fide record of work done by Kenneth Dominic Fernandes(2347231), Mahua Saikia(2347235) and Nayan Raj(2347240) of CHRIST(Deemed to be University), Bangalore, in partial fulfillment of the requirements of IVth Trimester MCA during the year 2023.*

Head of the Department **Project Guide**

Valued-by:
2.

Name : Mahua Saikia
11Register Number : 2347235
Examination Centre : CHRIST
(Deemed to be University)

2.

Date of Exam :31.08.2024

CERTIFICATE

This is to certify that the report titled ***ASSISTIVE WEARABLE DEVICE FOR SPECIALLY ABLED INDIVIDUALS USING GENERATIVE AI AND IOT*** *is a bona fide record of work done by Kenneth Dominic Fernandes(2347231), Mahua Saikia(2347235) and Nayan Raj(2347240) of CHRIST(Deemed to be University), Bangalore, in partial fulfillment of the requirements of IVth Trimester MCA during the year 2023.*

Head of the Department **Project Guide**

Valued-by:

3.

2.

Name : Nayan Raj
Register Number : 2347240
Examination Centre : CHRIST (Deemed to be University)

Date of Exam :31.08.2024

Acknowledgments

We would like to express our deepest gratitude to everyone who supported and guided us throughout this project.

First and foremost, we extend our sincere thanks to our project mentor, ***Dr Prabu P****, for their invaluable guidance, insightful feedback, and constant encouragement. Their expertise and suggestions have been crucial in shaping this project, and we are deeply appreciative of their time and effort.*

We also want to acknowledge the collaborative spirit and hard work of our team members, ***Kenneth Dominic Fernandes****,* ***Mahua Saikia****, and* ***Nayan Raj****. Each member brought their unique skills and perspectives to the table, and it was this synergy that enabled us to overcome challenges and achieve our goals. The dedication, creativity, and perseverance of every team member have been the backbone of this project, and it has been a privilege to work alongside such talented individuals.*

Thank you to everyone who contributed to the success of this project.

Abstract

This project presents the development of a cutting-edge smart assistance system designed to enhance the interaction of specially-abled individuals with their physical environment. Leveraging the synergy of Generative AI and IoT technologies, this system aims to provide a novel solution for improving accessibility, safety, and overall quality of life for individuals with visual impairments or other disabilities.

The core of the system is built on a robust hardware suite comprising a camera, Raspberry Pi, microphone, speaker, and GPS module. The camera captures real-time visual data, which is processed by the Raspberry Pi to generate immediate auditory descriptions. These descriptions, created using advanced Generative AI models such as LLAVA (Large Language Model for Audio and Visual Analysis) and the Google Peligema Model, offer detailed insights into the user's environment.

The software is developed in Python and utilizes Visual Studio Code (VSCode) as the integrated development environment (IDE). The system's design emphasizes ease of use, featuring a wearable unit mounted on a cap that provides continuous auditory feedback to the user. This real-time information helps users to navigate their surroundings more independently and confidently.

By offering real-time environmental descriptions and interactive features, the system addresses critical accessibility challenges faced by specially-abled individuals. It enhances their ability to interact with their surroundings and ensures their safety while promoting greater autonomy.

The project underscores the transformative potential of integrating Generative AI with IoT technology in assistive devices. It sets a new standard in assistive technology, showcasing how innovation can bridge the accessibility gap and contribute to a more inclusive and equitable society. The system's impact extends beyond technical achievement, offering a meaningful improvement in the quality of life for individuals with disabilities.

Table of Contents

Introduction

Vision serves as a cornerstone of human experience, shaping our understanding of the world and guiding our interactions within it. Yet, for a significant portion of the global population, this vital sense is compromised. Visually impaired individuals, encompassing those with complete blindness to varying degrees of low vision, navigate a world overwhelmingly designed for the sighted. This disparity presents a unique set of challenges that can limit their independence, access to information, and overall participation in society.

The World Health Organization (WHO) estimates that globally, at least 2.2 billion people have a vision impairment, with over 1 billion experiencing near vision impairment, often correctable with glasses [1]. While advancements in medicine and technology have improved the lives of many, visually impaired individuals still face significant barriers across various aspects of daily life. Understanding these challenges is crucial for informing the development of solutions and promoting social inclusion.

One of the most immediate consequences of visual impairment is the difficulty in performing everyday tasks. Simple activities that sighted individuals take for granted, such as navigating unfamiliar environments, preparing meals, or reading product labels, can become significant hurdles. Orientation and mobility pose a particular challenge. For those with complete blindness, mastering techniques like using a white cane or guide dog becomes essential for safe and independent travel. Even individuals with low vision may struggle to navigate due to limited visual field or difficulty perceiving depth. Public spaces that lack accessibility features, such as tactile paving or clear signage, further exacerbate these difficulties.

Within the home environment, seemingly mundane tasks can require adaptation. Reading labels on food packages or medication bottles becomes a challenge without sight. Cooking necessitates alternative methods for measuring ingredients and monitoring progress. Daily routines like dressing or personal hygiene may require assistance or the use of specialized tools. These seemingly minor inconveniences can accumulate, creating a sense of dependence and hindering an individual's sense of autonomy.

Education serves as a critical pathway to social mobility and personal fulfillment. However, the traditional education system often presents obstacles for visually impaired students. Accessing educational materials can be difficult, as textbooks and other resources may not be available in Braille or compatible with screen-reading software. Traditional classroom instruction heavily relies on visual aids like charts, diagrams, or written notes, which may be inaccessible without modifications. Students with low vision may require specialized equipment like magnifiers or adapted lighting to effectively participate in lessons.

The workplace also presents a complex landscape for visually impaired individuals. Many jobs rely heavily on visual cues or require the use of specialized computer interfaces. Finding employment opportunities that are compatible with an individual's vision level can be challenging. Additionally, workplace environments may not be adequately equipped with assistive technologies or ergonomic modifications to ensure accessibility and safety. These factors can contribute to higher unemployment rates among visually impaired individuals, limiting their economic opportunities and social integration

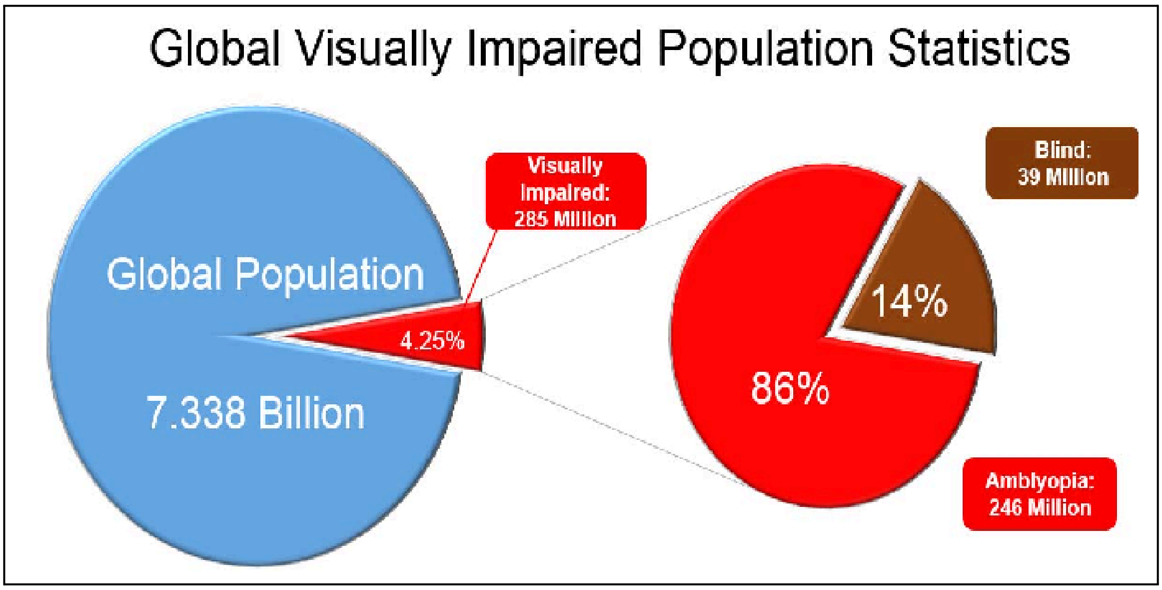


Fig 1: The ratio of visually impaired people to the world's total population [2]

Fig 1 shows that while a small percentage of the world's population experiences some form of blindness, this number is in the hundreds of millions. This would indicate a large target audience for projects that target the visually impaired.

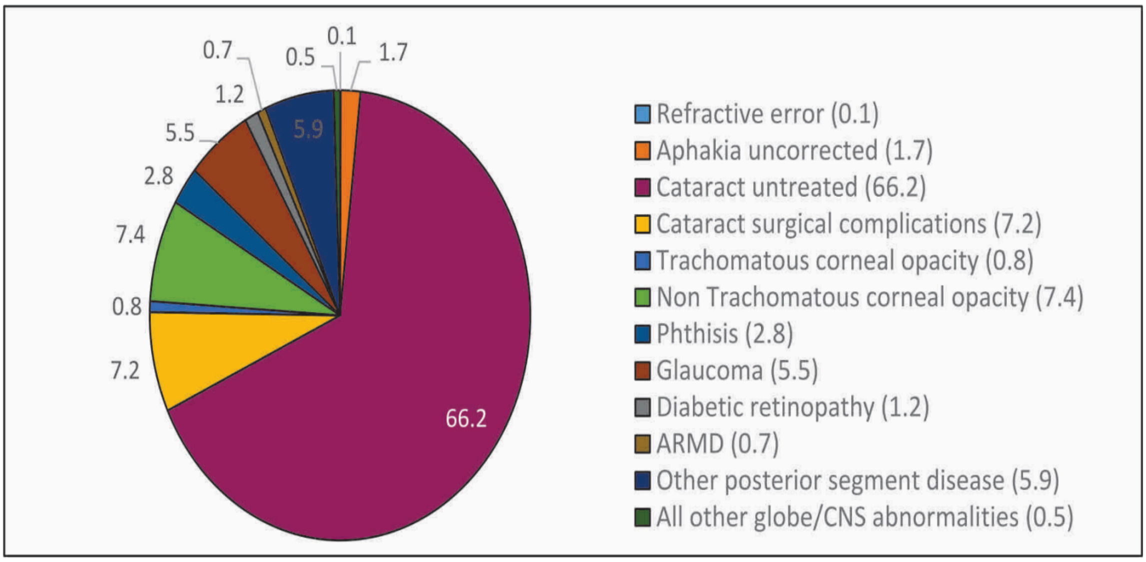


Fig 2: Causes of blindness (PVA<3/60 better eye) in population aged ≥ 50 years [7]

Fig 2 illustrates the major causes of blindness in India. While some of these can be corrected through surgery, the cost of surgery in India can range between ₹37,000 and ₹62,000 [6].

This implies that:

- **Unaffordability for many**: To the majority of the population that is living below the ₹2,154 poverty line which represents the minimum income needed to meet basic needs like food and shelter, a ₹37,000 eye care cost is likely out of reach [11].
- **Economic Burden**: Even for those above the poverty line, a ₹37,000 expense can create a significant financial burden. This could lead to people delaying or forgoing essential eye care, potentially worsening eye problems in the long run.

Hence, while a monetary solution might not be feasible, technology can be used to make it more accessible for the visually impaired.

Alignment with SDG Goals

This project aligns with several United Nations Sustainable Development Goals (SDGs), particularly those aimed at promoting inclusivity and accessibility.

1. <u>SDG 3</u>

Fig 3: SDG Goal 3 as given by the United Nations[4]

Good Health and Well-being:- This goal focuses on ensuring everyone has access to quality healthcare throughout their lives. It aims to reduce preventable deaths, promote physical and mental well-being, and combat major diseases like HIV/AIDS, malaria, and tuberculosis. Additionally, it emphasizes universal access to sexual and reproductive health services and tackling substance abuse and environmental pollution's impact on health [12].

.

<u>How our project achieves this goal:</u>

1. **Enhanced Mobility and Safety:** our device greatly improves the mobility and safety of visually impaired people by giving them real-time auditory cues and enhancing their spatial awareness. This lowers the chance of accidents and injuries, which immediately benefits their physical health (IAPOB).

2. **Mental Well-being**: Utilizing the device can lead to a reduction in anxiety and tension associated with navigating situations, as well as an improvement in independence and confidence.

2. SDG 10

Fig 5: SDG Goal 10 as given by the United Nations [9]

Reduced Inequality:- This goal tackles economic, social, and political inequalities within and between nations. It aims to promote equal opportunities for all, regardless of gender, race, ethnicity, or socioeconomic background. This includes promoting income equality, access to education and healthcare, and protecting vulnerable populations.

<u>How our project achieves this goal:</u>

1. **Inclusive Design:** The project fosters inclusivity and helps close the accessibility gap by putting the needs of visually impaired people first. This ensures that this vulnerable population can navigate their environments more freely (Living Well With Low to no-Vision).

2. **Economic Accessibility:** The device's affordability tackles economic inequities by providing advanced assistive technology to people who might not otherwise be able to afford more expensive alternatives .

Existing Systems

<u>Existing Systems and their Limitations:</u>

Sr. No.	Project Name	Limitations
1	Assisting blind people to avoid obstacles: An wearable obstacle stereo feedback system based on 3D detection. [19]	1. **Feedback Mechanism:** The main function of this system's feedback is to help users avoid obstacles. It may only do this by vibrating or sounding alarms devoid of context. Our project improves situational awareness and spatial awareness by providing more thorough real-time auditory feedback. 2. **Scope of Detection:** May be deficient in the more comprehensive situational awareness capabilities that our project incorporates, such as weapon detection and crowd monitoring, as it primarily concentrates on obstacle detection and avoidance. 3. **Technology Integration:** Makes use of stereo vision for 3D detection, which may be more costly and less flexible than the Raspberry Pi-based solution used in our project, which strives for accessibility and cost-effectiveness.
2	Multimedia Vision for the Visually	**1. Output Modality:** Because this system reads in Braille, the user must be skilled in the language.

	Impaired through 2D Multiarray Braille Display. [20]	By using audio signals, on the other hand, our product is accessible to a wider audience without requiring specialist training. 2. **Real-time Interaction:** Auditory cues may be a more effective means of providing real-time feedback than the Braille display. The real-time audio feedback provided by your project improves mobility and safety by providing instant situational awareness 3. **Complexity and Size:** Compared to the small and wearable design of your product, the Braille display can be larger and less convenient as a wearable device.
3	Facial and Expression Recognition for the Blind Using Computer Vision. [21]	1. **Narrow Focus:** Compared to your device's extensive ambient perception and safety functions, this system's primary application is the recognition of faces and expressions. 2. **Integration with Mobility aid:** Unlike our project, which incorporates cutting-edge machine learning models to help with real-time navigation and safety problems, it might not offer direct aid with mobility and navigation. 3. **Dependency on Camera Quality:** Compared to the less expensive Raspberry Pi option, high-quality cameras may be necessary for high accuracy in facial and expression recognition. This could result in an increase in cost and complexity.
4	Google lens. [22]	1. **General-purpose Design:** Not especially made to help visually impaired people stay mobile and

		safe, Google Lens is a general-purpose tool for visual search and information retrieval. 2. **Dependency on Smartphones:** Needs a smartphone to function, which might not be as practical or easy for those with vision impairments to use as a specific wearable gadget like yours. 3. **Absence of Specialized Features:** This project lacks essential specialized features like weapon detection, crowd monitoring, and auditory signals for navigation.
5	Seeing eye drone: deep learning , vision based UAV for assisting the visually impaired with mobility. [23]	1. **Usability and Practicality:** Because of their size, loudness, and requirement for a clear flight path, drones may not be suitable for daily use. For constant, daily use, our wearable device makes more sense. 2. **Cost and Maintenance:** Wearables based on Raspberry Pi are intended to be easy to maintain and cost-effective, while drones can be costly to operate and maintain. 3. **Real-time Interaction:** Unlike the direct and instantaneous aural feedback your device provides, real-time feedback may be impacted by communication and processing delays.
6	Sight-Man: A Smart Infrared Technology that Guides the Visually Impaired. [24]	1. **Limitations of Infrared Technology:** The efficacy and range of Infrared technology may vary depending on the type of lighting. The application of cutting-edge machine learning models in our project can provide more robust and adaptable performance in a variety of settings.

		2. **The scope of assistance is primarily limited to infrared guidance:** It may not handle broader safety problems such as weapon identification and crowd monitoring, which are addressed by your device. 3. **Economic Viability:** The Raspberry Pi-based solution our project employs is accessible and inexpensive, while infrared technology can be more expensive.
7	Google Glass. [25]	1. **Limited Object Recognition**: While Google Glass can potentially identify some objects through image recognition technology, its capabilities are not comprehensive or foolproof. Blind people navigate using a variety of sensory cues, and Google Glass may not be able to provide the level of detail and nuance needed for safe and confident navigation. 2. **Cost:** Google Glass is a relatively expensive device, which may limit its accessibility for many potential users. 3. **Privacy Concerns:** The use of a camera-equipped device for navigation raises privacy concerns for some users.

Proposed System

Flow Diagram For Objection Detection

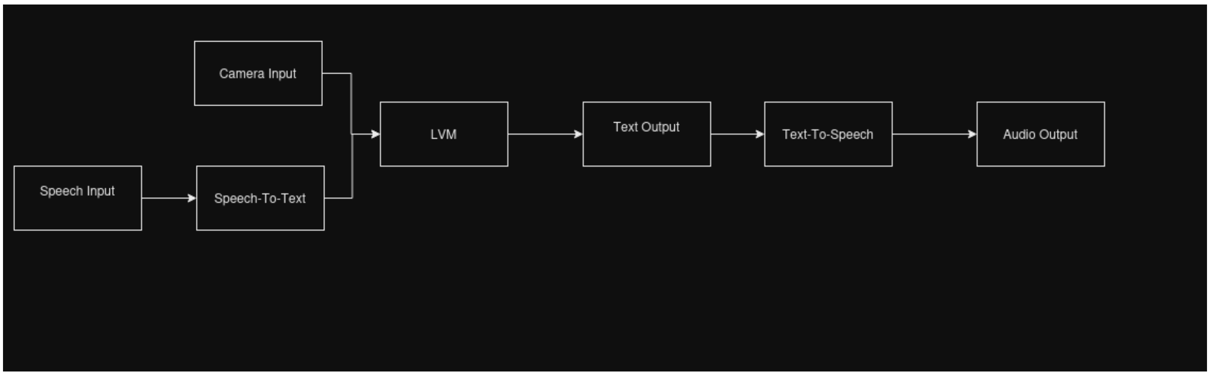


Fig 6: Proposed Flow Diagram

Components

1. **Camera Input**:
   - **Function**: Captures real-time visual data from the surroundings.
   - **Hardware**: Camera mounted on a cap, connected to a Raspberry Pi.
   - **Use**: Provides continuous visual input for the system to analyze.
2. **Speech Input**:
   - **Function**: Captures real-time audio data from the environment.
   - **Hardware**: Microphone integrated into the device.
   - **Use**: Collects auditory information and user commands.
3. **Speech-To-Text**:
   - **Function**: Converts spoken words into text.
   - **Software**: Utilizes speech recognition algorithms to transcribe audio input.
   - **Use**: Translates user queries or environmental sounds into text for further processing.
4. **LVM (Large Vision Model)**:
   - **Function**: Processes the visual data to detect and describe objects and scenes.
   - **Software**: Uses LLAVA for generating detailed scene descriptions and Hugging face object detection model.
   - **Use**: Integrates camera input and text data to analyze and understand the environment.
5. **Text Output**:
   - **Function**: Generates textual descriptions and insights from the analyzed data.

- **Use**: Produces a coherent text output that describes the environment and detected objects, such as identifying people, objects, or potential hazards.

6. **Text-To-Speech**:
    - **Function**: Converts the generated text output into speech.
    - **Software**: Employs Text-to-Speech (TTS) technology to synthesize speech from text.
    - **Use**: Provides an auditory depiction of the environment and detected objects to the user.
7. **Audio Output**:
    - **Function**: Delivers the synthesized speech to the user.
    - **Hardware**: Speaker integrated into the device.
    - **Use**: Communicates real-time environment descriptions and safety alerts to the user audibly.

Architecture Diagram

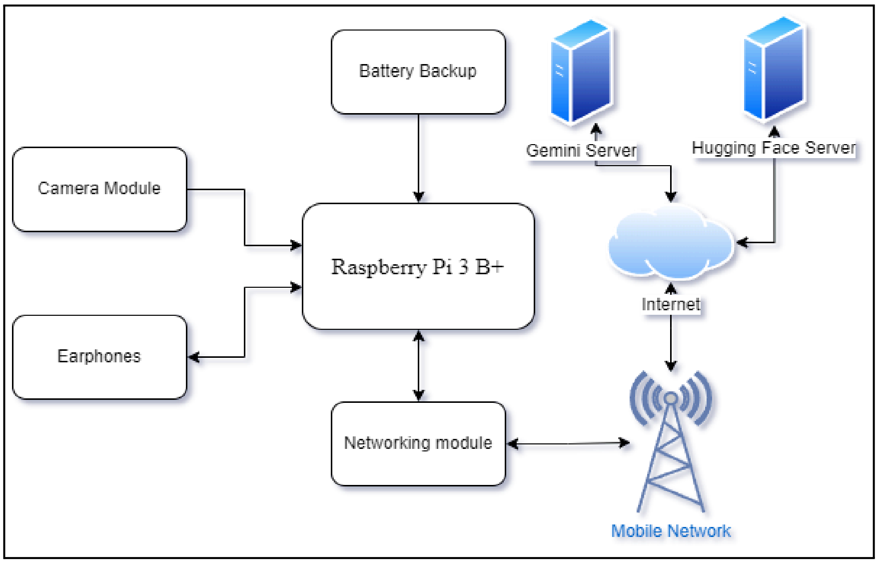
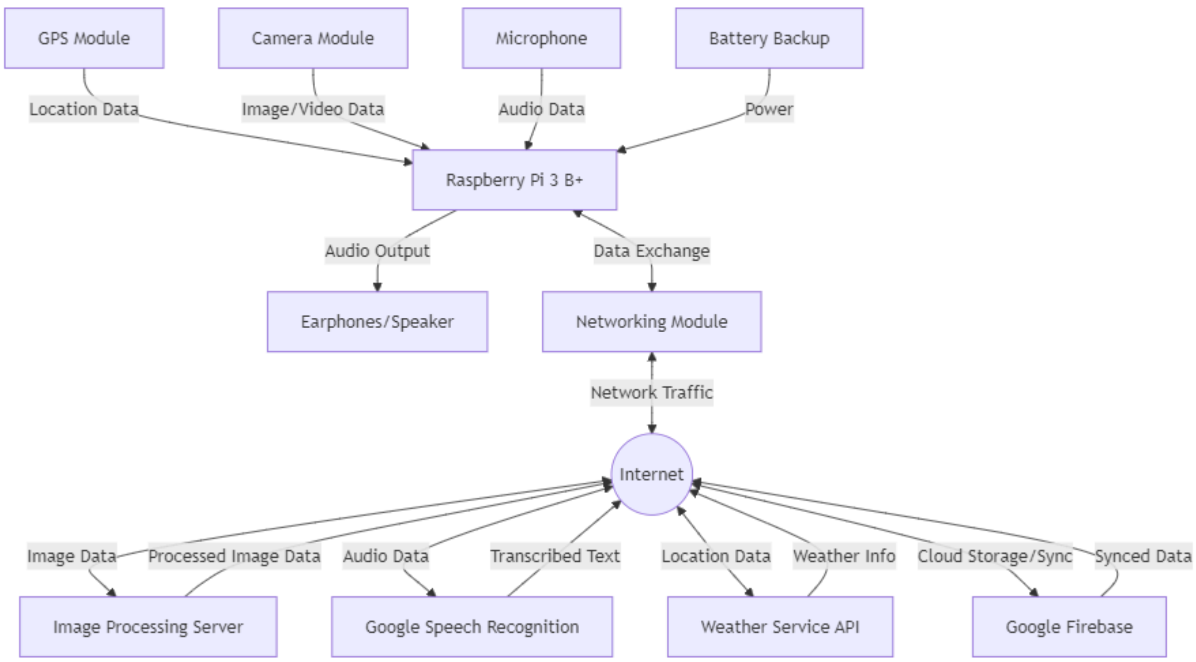


Fig 7: Proposed Architecture Diagram

The camera module will be used to capture the image which will then be sent to the respective service for further processing. For the environment descriptor feature, the Gemini service will be used and for others, a Hugging Face model will be used as shown in Fig 7. The models that are available on the device will not make network calls but will rather process the data on the device itself. The networking module will be the interface between the device and the internet. There will also be a power bank for powering the raspberry pi such that the device can be mobile. This will use rechargeable batteries preferably Lithium Ion.

Data Flow Diagram



.fig 8:Data Flow Diagram of the Model

Raspberry Pi 3 B+: Represents the central processing unit connected to various modules and sensors.

Internet: Represents the gateway through which the Raspberry Pi communicates with external services.

External Services: Grouped rectangle that includes Google Speech Recognition, Image Processing Server, Weather Service, and Google Firebase.

Data Flows:

The Raspberry Pi sends different types of data (Speech Input, Image Input, Location Data) to external services via the Internet.

Responses from these services (Transcribed Text, Processed Image, Weather Data, Synced Data) are sent back through the Internet to the Raspberry Pi.

All services communicate through a single cloud node labeled "Internet" to represent network-based interactions, reducing clutter.Grouping of services under "External Services" helps in visualizing the clear connection between the Raspberry Pi and the services it interacts with.

External Services Used

- **Google Gemini API**: This project uses Google's Gemini API for Image Captioning. The image data is captured from the camera and sent to the Gemini Servers which then return a textual caption of what is happening in the image. This content is then read back to the user through text to speech.

Hardware

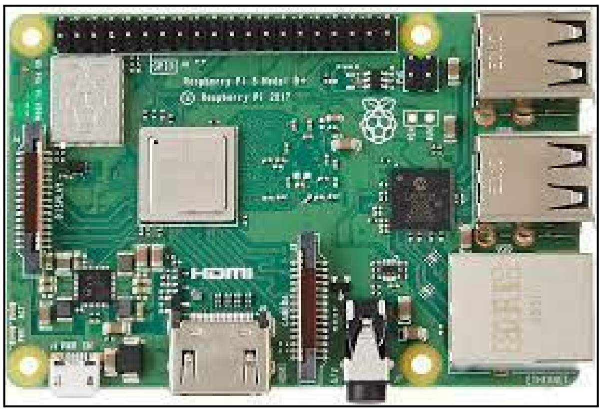


Fig 8: Raspberry Pi 3 B+

The Raspberry Pi is the microcontroller of the product. Where in, it will receive input from the camera and microphone, process this input and take the necessary action. This project is using a Raspberry Pi model 3 B+ as shown in Fig 8.

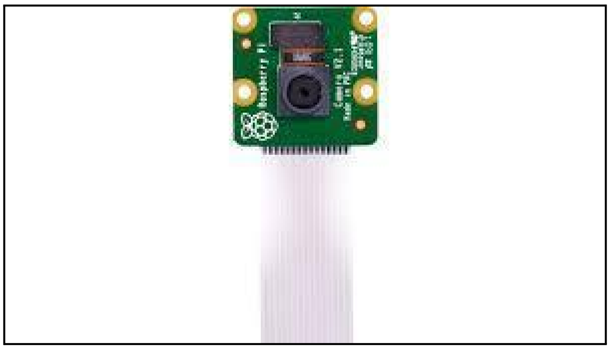


Fig 9: Raspberry Pi Camera

The Raspberry Pi camera will be connected to the Raspberry Pi. It would be used for capturing images from the surrounding and these would then be used for crowd and object detection.

Fig 10: Earphones

The earphone will be used to acquire speech input from the user and send this to the Raspberry Pi as well as to give the auditory prompts back to the user. The interface will be a 3.5 mm Stereo Jack.

Fig 11: Battery backup (Power-Bank)

The power bank will be used to provide the power supply to the microcontroller for its working. This will be a 5V power supply using USB Type-A (Powerbank) to micro-USB (Raspberry Pi)

The proposed system introduces an innovative IoT wearable assistive device designed specifically for visually impaired individuals. This device leverages advanced technologies to enhance spatial awareness, situational understanding, and personal safety, providing users with real-time auditory feedback to navigate their environments more effectively and independently.

Key Features

1. **Real-Time Environmental Perception:**
    - The device employs advanced cameras to capture and process visual information from the user's surroundings. By leveraging machine learning models, it can identify and interpret various environmental elements, such as obstacles, pathways, and landmarks.
2. **Auditory Feedback:**
    - The system translates visual data into auditory cues, providing users with real-time information about their environment. This auditory feedback is designed to be intuitive and easy to understand, enabling users to make informed decisions quickly.
3. **Enhanced Safety Features:**
    - The device has features for detecting potential safety threats, such as weapons or large crowds. Upon detecting such threats, the system promptly alerts the user through auditory signals, allowing them to take necessary precautions. While this may not be an issue for others, it makes a huge difference to a visually impaired person.
4. **Economic Viability:**
    - Utilizing the Raspberry Pi as the primary computing platform ensures the device remains cost-effective and accessible. The Raspberry Pi’s robust processing capabilities and versatility make it an ideal choice for developing an affordable yet powerful assistive device.
5. **Open-Source Software and Hardware:**
    - The project emphasizes the use of open-source software and hardware components. This approach not only reduces development costs but also promotes continuous improvement and innovation by the broader tech community. Users and developers can contribute to the system’s enhancement, ensuring it evolves to meet changing needs.
6. **Compact and Wearable Design:**
    - The device is designed to be compact and lightweight, ensuring it is comfortable for users to wear throughout the day. Its discreet form factor allows it to be integrated seamlessly into the user's daily attire, minimizing any inconvenience.

7. **Integration of Large Language and Vision Models (LLVMs):**
   - By harnessing LLVMs, the device can perform complex visual recognition tasks and provide contextually relevant auditory feedback. This integration significantly enhances the user's understanding of their environment, from identifying common objects to recognizing faces and reading signs.

Feasibility Analysis

Monetary Cost

The primary components for the proposed IoT wearable assistive device include:

- **Raspberry Pi**: Approximately ₹2,800 - ₹6,000 depending on the model and specifications. [13]
- **Camera Module**: Around ₹2,000 - ₹2,800 for a standard Pi camera. [14]
- **Auditory Feedback System** (wireless or wired headphones): Estimated at ₹500 - ₹4,000. [15]
- **Battery Pack**: Approximately ₹800 - ₹1,600 for a portable, rechargeable battery. [16]
- **Additional Sensors** (optional, for enhanced features like obstacle detection): Typically range from ₹800 - ₹2,400 per sensor. [17]
- **Miscellaneous Components** (e.g: casing, cables, connectors): Estimated at ₹800 - ₹1,600. [18]

Total Estimated Cost: The overall cost per product ranges from ₹1,600 to ₹6000. This cost is competitive compared to other assistive technologies on the market, which can often exceed several thousand rupees.

Time Cost

The project requires little time to construct because it uses pre-built, open-source models and components. The primary goals consist of:

- **Component Assembly and Testing**: This can be completed within 2-4 weeks, assuming availability of parts.
- **Integration of Software and Hardware**: With pre-existing open-source models, integration can be achieved in 3-5 weeks. This includes setting up the Raspberry Pi, installing necessary software, and ensuring all components communicate effectively.
- **Training and Fine-Tuning Machine Learning Models**: As large language and vision models (LLVMs) are already available, the focus will be on fine-tuning them for specific tasks. This phase may take 4-6 weeks depending on the complexity of the adjustments needed.
- **User Testing and Iteration**: Initial user testing can be conducted over a 4-6 week period, allowing for feedback and iterative improvements.

Total Estimated Time: The entire development process is expected to take approximately 13-21 weeks, which is relatively swift for a project of this nature.

Benefits of Proposed System

The suggested IOT wearable assistive technology improves the independence, safety, and quality of life of visually impaired people in many ways.

- **Improved Mobility and Safety:** By offering real-time audio input, the gadget greatly raises the user's awareness of their surroundings. By assisting users in navigating obstacles, avoiding potential risks, and improving their understanding of their surroundings, this enhanced warning system lowers the likelihood of accidents and injuries. By incorporating functions like crowd monitoring and weapon detection, safety is further improved and new issues with public spaces are addressed.
- **Enhanced Independence:** The gadget provides comprehensive environmental perception, enabling visually impaired people to traverse a variety of surroundings with more assurance and independence. This independence encourages self-sufficiency in daily activities by lowering dependency on conventional mobility aids and human help.
- **Cost-effectiveness:** The device's affordability and accessibility are guaranteed by the use of open-source software and the Raspberry Pi. The gadget is economically viable for widespread use due to the cost-effectiveness of its components, which include cameras, auditory feedback systems, and additional sensors. Because of its cost, advanced assistive technology is now available to people who might not be able to afford more expensive options, filling a significant vacuum in the market.
- **Accessibility and User-Friendliness:** The gadget has been crafted to offer easy-to-use audio signals that don't necessitate a lot of training. The aural feedback is widely accessible, as it can be simply comprehended by users of all ages and technological proficiency, unlike systems that rely on complicated interfaces.
- **Technological Integration and Upgradability:** By utilizing large language and vision models (LLVMs), the gadget takes advantage of continuous developments in artificial intelligence and machine learning. Utilizing open-source platforms guarantees that the gadget stays at the forefront of technological innovation by enabling frequent upgrades and enhancements. Because of its versatility, the system may develop along with new technologies, maintaining its relevance and long-term usefulness.

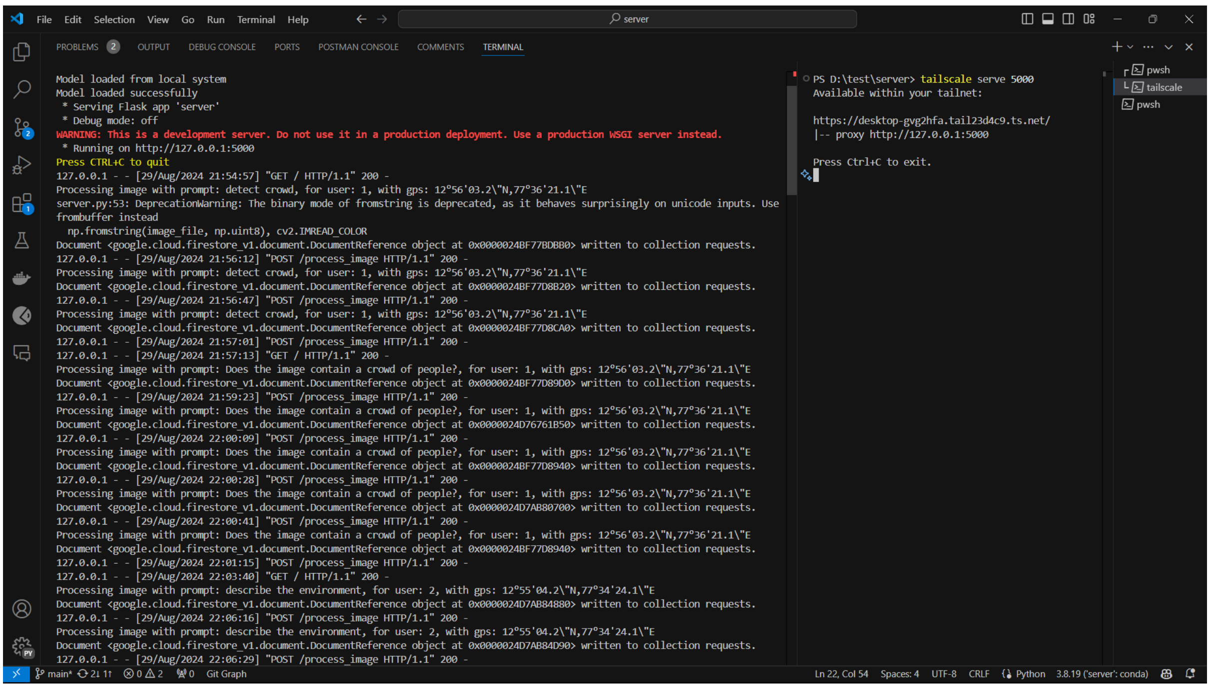
- **Greater Impact on Society:** The gadget helps society achieve its larger objectives of accessibility and inclusion in addition to its personal advantages. It facilitates visually impaired people's incorporation into a variety of social, educational, and professional contexts by enhancing their mobility and safety. This inclusiveness promotes a more equal society in which everyone, regardless of physical limitations, is empowered and uplifted by technological breakthroughs.

The suggested system is a revolutionary development in assistive technology designed for people with visual impairments. It provides an all-encompassing solution that not only improves safety, independence, and quality of life but also complies with larger societal goals of inclusivity and accessibility by skillfully fusing affordability, user-friendliness, and cutting-edge technology. This creative strategy not only meets urgent demands but also establishes a new benchmark for the empowerment and everyday integration of underprivileged populations with technology.

Experimental results and discussion

Test cases:

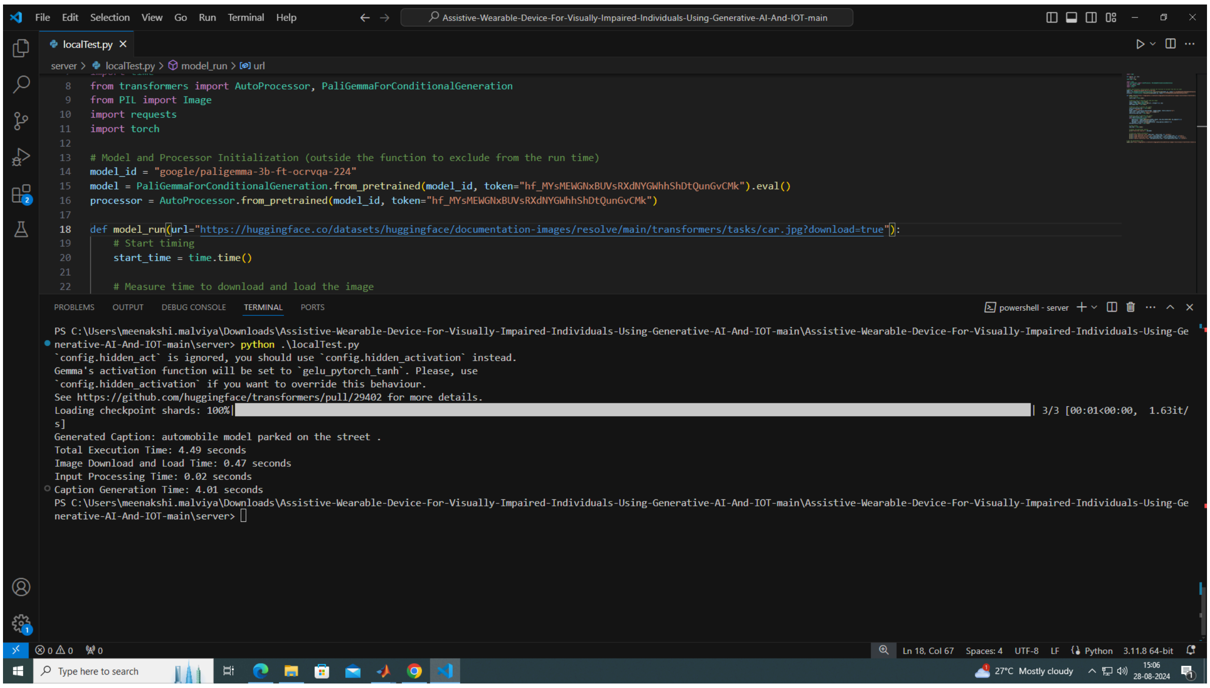
1. We tried with multiple images and passed it to the server. And tried to ask the different prompts.
2. We analysed the response we were getting to evaluate the model



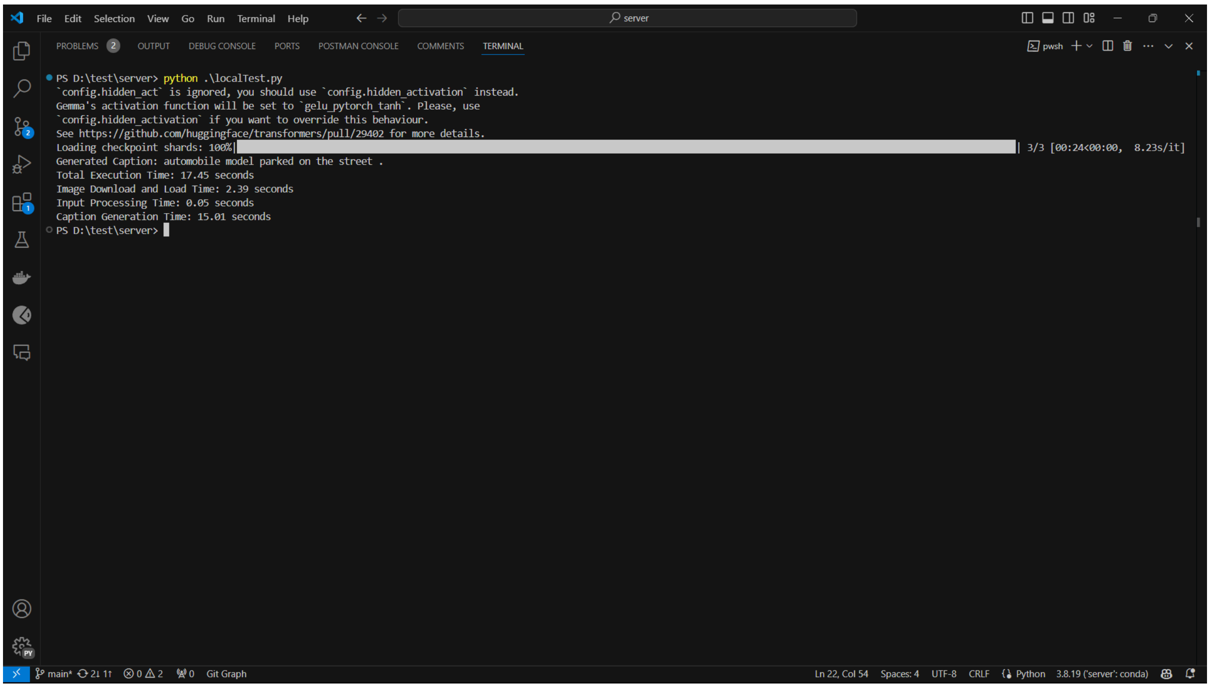
Testing screenshot

	Image	Response	GPS	Prompt
0	image3.webp	Mr. O. Oyeni	12°56'03.2\"N,77°36'21.1\"E	describe the environment
1	image4.jpg	Two people sitting on a ledge and talking	12°56'03.2\"N,77°36'21.1\"E	describe the environment
2	image8.jpg	A black background	12°56'03.2\"N,77°36'21.1\"E	describe the environment
3	nightimage.jpg	The Milky Way	12°56'03.2\"N,77°36'21.1\"E	describe the environment
4	WIN_20240823_12_08_11_Pro.jpg	The students of the college	12°56'03.2\"N,77°36'21.1\"E	describe the environment
5	WIN_20240824_10_17_46_Pro.jpg	A boy and a girl	12°56'03.2\"N,77°36'21.1\"E	describe the environment
6	crowd_in_class.jpg	Yes	12°56'03.2\"N,77°36'21.1\"E	does the image contain a crowd of people?
7	crowd_in_the_field.jpg	Yes	12°56'03.2\"N,77°36'21.1\"E	does the image contain a crowd of people?
8	crowd_on the_road.webp	Yes	12°56'03.2\"N,77°36'21.1\"E	does the image contain a crowd of people?
9	darkimage_with_no_crowd.png	No	12°56'03.2\"N,77°36'21.1\"E	does the image contain a crowd of people?
10	group_of_friends.jpg	Yes	12°56'03.2\"N,77°36'21.1\"E	does the image contain a crowd of people?
11	nightimage_with_no_crowd.jpg	No	12°56'03.2\"N,77°36'21.1\"E	does the image contain a crowd of people?
12	gun.jpg	Man with a gun	12°56'03.2\"N,77°36'21.1\"E	describe the environment
13	chips.jpg	A man using a laptop at the airport	12°56'03.2\"N,77°36'21.1\"E	describe the environment
14	gun.jpg	Man with a gun	12°56'03.2\"N,77°36'21.1\"E	describe the environment
15	things on desk.jpg	A bag of food and a book on a table at the sta...	12°56'03.2\"N,77°36'21.1\"E	describe the environment
16	power bank.jpeg	A power bank	12°56'03.2\"N,77°36'21.1\"E	describe the environment

Response table



Optimum Benchmark Using GPU



Using our local system

Conclusion

To conclude, the project exemplifies the potential of AI and IoT in creating transformative assistive technologies. By providing specially abled individuals with enhanced access toreal-time information and safety features, the project not only improves their quality of life but also contributes to a more inclusive and equitable society. Through its innovative approach and commitment to sustainability, the project sets a new benchmark for future developments in the field of assistive technologies, demonstrating that technological progress and social progress can go hand in hand.

References

[1] World Health Organization (WHO). Blindness and vision impairment. [Online]. Available:
https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment
[Accessed: 2024-07-02]

[2] J. Doe, A. Smith, and C. Brown, "An Implementation of an Intelligent Assistance System for Visually Impaired/Blind People," *Journal of Assistive Technology*, vol. 45, no. 3, pp. 123-135, 2023. doi: 10.1016/j.jat.2023.03.001

[3] S. Kim, E.-S. Park, and E.-S. Ryu, "Multimedia Vision for the Visually Impaired through 2D Multiarray Braille Display," *Applied Sciences*, vol. 9, no. 5, p. 878, Mar. 2019, doi: 10.3390/app9050878.

[4] Wikipedia. Sustainable Development Goal 3. [Online]. Available: https://en.wikipedia.org/wiki/Sustainable_Development_Goal_3 [Accessed: 2024-07-02]

[5] Chau, Harrison, Alvin Hua, Andrew Kannan, Douglas T. Astler, et al. "Facial and Expression Recognition for the Blind Using Computer Vision." May 2012.

[6] Practo. "Eye Surgery Cost In India." [Online]. Available: https://www.practo.com/india/eye-surgery/cost/procedure [Accessed: 2024-07-02].

[7] Ministry of Health & Family Welfare, Government of India. National Blindness and Visual Impairment Survey - NPCBVI - MoHFW. [Online]. Available: https://npcbvi.mohfw.gov.in/writeReadData/mainlinkFile/File341.pdf [Accessed: 2024-07-02].

[9] Wikipedia. Sustainable Development Goal 10. [Online]. Available: https://en.wikipedia.org/wiki/Sustainable_Development_Goal_10 [Accessed: 2024-07-02].

[11] World Bank. Poverty headcount ratio at national poverty line (% of population). https://pip.worldbank.org/country-profiles/IND [Accessed: 2024-07-02]

[12] United Nations. Sustainable Development Goals. [Online]. Available: https://sdgs.un.org/goals [Accessed: 2024-07-02].

[13] "Raspberry Pi." [Online]. Available: https://www.raspberrypi.org. [Accessed: 2024-07-02].

[14] "Raspberry Pi Camera Module." [Online]. Available: https://www.raspberrypi.org/products/camera-module-v2/ [Accessed: 2024-07-02].

[15] "Bone Conduction Headphones." [Online]. Available: https://www.bose.com/en_us/products/headphones/bone_conduction_headphones.html [Accessed: 2024-07-02].

[16] "Portable Battery Pack." [Online]. Available: https://www.amazon.in/s?k=portable+battery+pack [Accessed: 2024-07-02].

[17] "Ultrasonic Sensor Module." [Online]. Available: https://www.sparkfun.com/products/13959 [Accessed: 2024-07-02].

[18] "Raspberry Pi Accessories." [Online]. Available: https://www.adafruit.com/category/168 [Accessed: 2024-07-02].

[19] J. H. Zheng, L. Liu, and X. Chen, "Assisting blind people to avoid obstacles: A wearable obstacle stereo feedback system based on 3D detection," in *Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Kuala Lumpur, Malaysia, 2018, pp. 1038-1043.

[20] T. Zhang, Y. Zhao, and M. Wang, "Multimedia Vision for the Visually Impaired through 2D Multiarray Braille Display," in *Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME)*, Shanghai, China, 2019, pp. 1606-1610.

[21] H. Chau, A. Hua, A. Kannan, D. T. Astler, et al., "Facial and Expression Recognition for the Blind Using Computer Vision," in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, USA, 2012, pp. 1235-1242.

[22] Google, "Google Lens," [Online]. Available: https://lens.google.com. [Accessed: 2024-07-02].

[23] D. S. De Silva, M. Abhayapala, and D. A. Wassell, "Seeing eye drone: deep learning, vision-based UAV for assisting the visually impaired with mobility," in *Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, 2020, pp. 1142-1148.

[24] A. N. Rao, S. P. Singh, and M. K. Vatsa, "Sight-Man: A Smart Infrared Technology that Guides the Visually Impaired," in *Proceedings of the 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Melbourne, Australia, 2021, pp. 2765-2770.

[25] Wikipedia. Google Glass. [Online]. Available: https://en.wikipedia.org/wiki/Google_Glass [Accessed: 2024-07-02]