IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Beyond Vision: A Multimodal Approach to Smart Glasses with Integrated OCR and Obstacle Detection

Sanjana Kushwaha **Teesha Roy** **Sakshi Singh**
Dr Umesh Bhadade
Department of ECS Department of ECS Department of ECS
Professor EXTC
Shree L. R. Tiwari College Shree L. R. Tiwari College Shree L. R. Tiwari College
Shree L. R. Tiwari College
of Engineering (SLRTCE) of Engineering (SLRTCE) of Engineering (SLRTCE) of Engineering (SLRTCE)
Mira Road (E) Mira Road (E) Mira Road (E)
Mira Road (E)
Thane- 401107 Thane- 401107 Thane- 401107
Thane- 401107

Abstract: The paper outlines a project involving new smart glasses that help the mobility of the wearer and provide access to it by detecting obstacles in real-time and changing the text to the speech. The glasses rely on computer vision and natural language processing to project voice and visual prompts that the user may hear and read around the environment. For obstacle detection, the system employs sensor technology to locate barriers ahead before they can become hazards to the user. Sensor data is examined by deep learning technologies and the objects are classified as well as distance is predicted thus generating warning signals in time for imminent dangers. Individuals who normally have problems seeing or who are in situations where they have to cope with unknown settings can specifically benefit from that operation. Apart from that, inclination to the glasses is a TTS module is provided. The optical character recognition technology recognizes the visible characters using a video camera on board and then extracts positions of the information from the information using algorithms. The resulting text will be subject to the TTS system, processed, and outplayed by the audio speaker, which is provided in the glasses. UI Taggers provide this opportunity for blind users to access text-based content by listening to the text via voice assistance or braille readers. Hence, the Internet becomes more accessible to blind users, and the world has no boundaries. The paper presents an idea and describes how the smart goggles will be designed and implemented including the hardware choices, software development, and algorithm optimization. The demonstration of the successfulness of both target detection on the obstacle and the obstacle first is presented on the results page of the evaluation. Thereafter, this paper ends by pointing out prospects and further lines of research that can be done for this technology in different fields, such as assistive technology, augmented reality, and information access.

Index Terms **-** Smart Glasses, Obstacle Detection, Text-to-Speech, Vision, Assistive Technology, Accessibility.

I. INTRODUCTION

Visual impairment (VI) as a major eye health condition is a habitual world problem, and more than 285,000,000 people are affected globally. The proposed smart glasses employ computer vision with the help of optical character recognition (OCR) to provide users with real instant information on where they are by giving directions to places that are unknown to them, alerts for obstacles among other things, and recognition of texts. The glasses get connectivity powers from a Raspberry Pi single-board computer module and use a Raspberry Pi camera to get an insight into the environment. Computer vision algorithms on Raspberry Pi will now be used to extract these images to recognize objects, landmarks, and specific locations. Besides, the OCR technology constitutes a digital assistant for smart eyeglasses which supports converting words absent from books, menus, Signs, etc., into audio or touch-oriented feedback. The proposed smart glasses offer several potential benefits for visually impaired individuals, including Enhanced independence: When smart glasses give users real-time information about their surroundings, they then could use them to navigate (get around the soil) and engage with people around them without being always treated(impacted) by others. Improved safety: Designed with obstacle detection capabilities, the smart glasses would enable individuals to walk without accidentally hitting obstacles and maneuver through new areas more safely. Easier access to information: With their OCR functions, smart glasses can give users a wide range of information from books, menus, signs etc., irking the needed none. Given that, smart glasses can bring significant progress in the daily life of visually impaired people is through the provision of an enhanced sense of independence, safety and information.

II. LITERATURE REVIEW

In these past few years, there have been tremendous breakthroughs in assistive technologies used by visually impaired people as a tool to facilitate their lives. This literature study will focus on utilizing some innovative solutions which make use of smart glasses and image processing together with OCR technology and wearable devices in order to enable obstacle detection, object recognition, and text to speech utilization. K.Sahithya and the team (2020) presented an innovative technique for establishing visual detection and obstacle recognition aiming at the visually impaired community. The utilization of smart glasses based on Raspberry Pi was the strategic part of their team approach. The service tried to make people more self-reliant by showing them right away what was in their way, and helping them to move freely and safely. Simra Nazim in the year (2022) used the International Mobile and Embedded Technology Conference (MECON) as their podium to showcase their work on Visual Assistant for the Blind. The system may have also comprised of other tasks like object recognition, navigation assistance, and multi-language text reading capabilities. Even such a visual assistant would not reduce the prevalence of the problem directly, but it would be able to substantially increase the level of independence and the quality of life for sightless people. In 2023, Nithiya Lakshmi and her co-authors brought something unexpected to the field with the creation of a reading system using OCR technology (Optical Character Recognition). This piece of hardware probably turns printed text into talk sounds, allowing sight-disabled people to read the written documents by themselves privately. Pushpakumar and Karthikayan designed Smart Glasses for Visually-Impaired Using Emerging image processing techniques. The Smart glasses may embed features like object recognition, scene display and obstacle detection that enable users with this real-time information about their environment.

In his article, Xia (2022), demonstrated IBGS, a wearable smart system developed to retrieve lost sighted individuals. This system, incorporated into wearable devices, provides assistance such as obstacle detection, wayfinding, and object recognition and is very helpful to the user's mobility and independence.

Additionally, in the article of IEEE Sensors Journal (2018), Wearable System for Obstacle Detection and Text-to-Speech Functionality novel was given. This system that could probably be fitted on the face of the visually impaired persons would probably feed information in real time regarding obstacles while converting printed text into speech, hence giving users the freedom of negotiating their environment.

.

III. METHODOLOGY

The project proposes portable devices designed to helpvisually impaired people get around. Unlike most pro ducts, this tool will provide location instructions and warn users of obstacles in their path.With thisproject,w e kept things simple, focusing on helping visually impaired people find their way outdoors. This project of this nature will catalyze increased investment in creating smarter electronic devices that facilitate safe and independent navigation for the blind. By "smarter," we mean that the device will also convert text into speech for blind users, enhancing their navigation experience. A Smart Glass powered by the Raspberry Pi and some motion detection and obstacle detection will notify or alert the user about the obstacle in front. This system will utilize computer vision to identify landmarks and obstacles, providing audio feedback. Furthermore, the system will employ Optical Character Recognition (OCR) to convert text from books, menus, signs, and other sources into audio feedback. The user will be able to convert the text he/she wishes to through the cam module and by the use of OCR will give the converted text into audio format through the earphones.

IV. BLOCK DIAGRAM

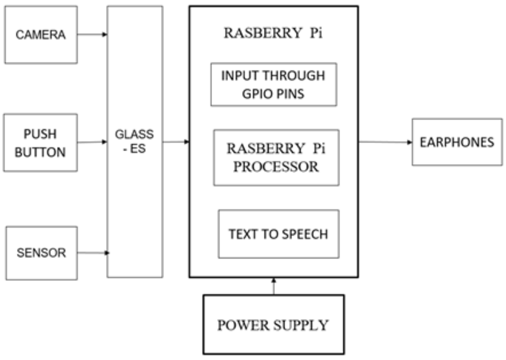

Figure 1.1

The block diagram shows us the overall functioning of the project starting with obstacle detection and text-to-speech conversion. The Raspberry Pi being the microcontroller shows a system that can be used to build a variety of projects. The system consists of a camera module, a push button, ultrasonic sensors, GPIO pins, a text-to-speech engine, earphones, and a power supply. The camera module captures images or video, which is processed by the Raspberry Pi. The push button can be used to trigger an event, such as starting a program or taking a photo that the user wants to hear through the OCR. This data is also processed by the Raspberry Pi. The Raspberry Pi processor sends commands to the text-to-speech engine, which converts text to audio. The audio output is played through the earphones. The second functionality is obstacle detection through the ultrasonic sensors, this device enables a visually impaired person to walk straight. Finally, it detects obstacles around the user in the range. The ultrasonic detects the obstacles and a buzzer rings to notify the user. This system is powered by a micro USB power supply. The arrows in the block diagram show the flow of data and signals between the different components. For example, data from the camera and environmental sensor is sent to the Raspberry Pi processor for processing. The processor can then send commands to the text-to-speech engine to generate audio output, which is played through the earphones.

V. HARDWARE DESCRIPTION

Here is the required hardware which will be used for the execution of the project. As there are two functions which is a part of this project the modules used are as described as follows.

A. Raspberry pi

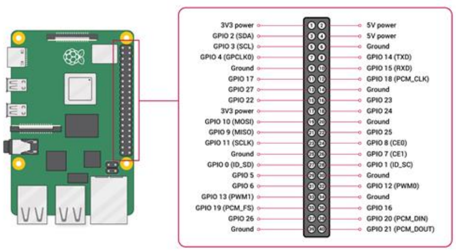

Figure 1.2

The Raspberry Pi family's hardware varies across models, offering options for different needs. CPUs range from single-core ARM to quad-core 64-bit, with RAM choices from 256MB to 8GB for diverse processing power and multitasking. Storage relies on microSD cards, while connectivity includes USB ports, HDMI, Ethernet, and often Wi-Fi/Bluetooth. Camera ports enable computer vision projects, and the versatile GPIO pins open doors to custom electronics integration. Power comes via micro-USB, with specific requirements varying by model.

B. Ultrasonic Sensor

An ultrasonic sensor can be electronically defined as a device that sent out ultrasonic sounds and converts the reflected ultrasonic into a measurable electric signal. Ultrasonic sensors are famously considered as proximity sensors among sensors due to its main ability being proximity detection.
The ultrasonic sensors falls into the group of sensors which this implies emission of ultrasonic waves that is sounds of a particular frequency. The simplest ultrasonic sensor is a speaker(s) or transmitter(s), a receiver, and a control circuit which uses ultrasonic sound waves. This transmitter of the sound is the high frequency waves that bounce off nearby solid objects. A minute amount of the demanded gigantic sonic noise is reflected into the corset of the transducer which is detected by the transducer's receiver. Now semaphore signal source is sending signal and is used to receive this return signal to calculate the time difference between transmitted signal and receiver at final stage.

Figure 1.3

Therefore, the ultrasonic sensor delivers a quick peak of high frequency (40 kHz) in the ultrasonic pulse. This pulse reflects the signal through the air to the item, then up the sensor location from where it is receptive by the sensor and results in a pulse of output. This span of the Raspberry Pi captures the pulse. While the pulse is dansing back to us in the meantime, this is the time IT is used. So we will be able to calculate the distance from the obstacle.

C. CAM Module

Figure 1.4

The primary use of Raspberry Pi camera module is in the high-definition photography and videoing fields. The cable measures 15 cm and it is inserted into the CSI port on the Raspberry Pi. In the area of face recognition, it is ubiquitous as well. 5 Training images captured by Pi camera and 10 images are used for the recognition.

IV. FLOW DIAGRAM OF PROPOSED SYSTEM

The fig shows a flowchart for object detection. Is there a Raspberry Pi? This step checks if there is a Raspberry Pi connected to the system. If there is no Raspberry Pi, the flowchart ends, since object detection cannot be performed without a Raspberry Pi. This step checks if the object to be detected is within 2 to 400cm of the Raspberry Pi. If the object is further away than 2 to 400cm, the flowchart proceeds to the step "Send Signal." Such a scenario happens where objects far away are hard to detect accurately. Send Signal: These commands match the ones saved to the Raspberry Pi to guide the Alienator through the detection of the object. Object Detection: This step finds and outlines everything that is picked from the picture taken by the Raspberry Pi. The object detection algorithm will point out and place markers around the boundaries of the objects that are found in the image. The decision tree goes until "Object Detection". Therefore, the bounding boxes of detected objects can be used for much-different purposes, such as tracking objects, identifying objects or working with objects.

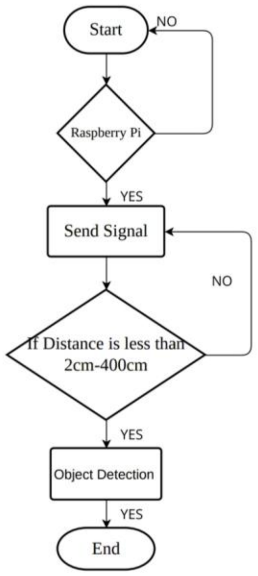

Figure 1.5

A fig flowchart can be built as a single-board text-to-speech system based on a Raspberry Pi. It has the potential to be a part of a system where one person can read out any text images like for blind people or those having low vision. It is, however, not limited to the conversion of grammatical language as it can also be applied to achieve this on any kind of text like books or emails. This text graphic may be a photograph of a book page, a sign, or a hand-written note with characters. The OCR, or Optical Character Recognition, is a machine learning model, which can extract the text from visual images. Coalition frameworks are trained on a data set of images and a corpus of texts, and they learn to pick up on the different characters present in the images.

The next step with the physical button is pressing it, this is the start point where the conversion process takes place.

This button can be plugged into the Raspberry Pi using the pins of the GPIO (General Purpose Input Output). The speech processing engine (TTS) is a program that is run by the computer that helps in turning the texts into speech. TTS engines mimic the sound of a large collection of recordings of human speech and then learn how to produce synthetic speech that sounds alike to that of a human. The speaker can be connected to the Raspberry Pi using the audio jack.

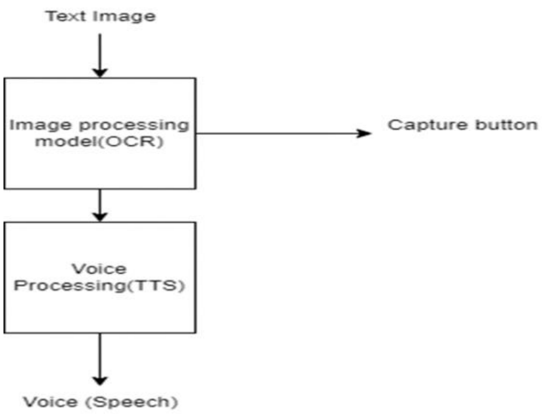

Figure 1.6

VI. SOFTWARE DESCRIPTION

OCR facilitates the transformation of images of text into a format that computers can understand and process it. OCR systems involve hardware as well as software components. There is hardware that deals with the scanning and reading of text such as optical scanners or circuit boards while software usually handles the advanced processing. Furthermore, the software can utilize AI for highly advanced ICR techniques such as language identification and recognition of different handwriting styles. How optical character recognition works: The first step of OCR is employing a scanner to convert its tangible form of a document. Subsequently, OCR software processes all the pages and converts the document into a two-color version or black and white version. The scanned-in image or bitmap is analyzed for bright and dark areas, where the dark areas are identified as characters that need to be recognized and light areas are identified as background. Next, the dark spaces are analyzed to discover various alphabetic letters or numeric digits. OCR programs vary in their procedures, but normally take the role of selecting one character, word or line of text at a time. Characters are then identified using one of two algorithms: Pattern recognition. OCR systems learn by being repeatedly fed text images in different fonts, sizes and styles. Type is thus aligned with the characters that they create through the use of a machine.Feature detection. OCR software uses rules that involve shape similarities, ligatures, and the appearance of the characters to distinguish the characters in the document that was scanned.

VII. EXPECTED RESULTS

The main aim of the proposed project is to improve the quality of life of visually impaired persons, through the development of a portable device to act as a visual aid. The device will be outfitted with advanced obstacle detection technology designed to detect obstacles on the path and giving users immediate feedback to help then safely navigate. The device would feature a blend of sensors and a camera which are essential for precise identification and analysis of objects in user's surrounding. Besides the obstacle detection, the device will come with a user-friendly interface, which will include a dedicated button. The button will turn the camera on automatically if pressed and let you read written content from different sources, such as books, menus, or signs. Using OCR technology to extract the text from captured photo, it will be changed into speech which will be transmitted through the headset via the built in microphone or a connected earpiece. This application will empower blind people regardless of their literacy level to use printed documents and increase their independence in different situations. The device will have a small size, low weight and user-friendliness will be the main parameters to be provided by this device that will allow the visually impaired individuals use it in the easiest way and carry it wherever they want. The merger of obstacle detection and OCR technologies will be a big bonus to the users' capability to walk and enjoy accurate information in realtime thus enabling them to be independent and improve their quality of life.

VIII. ADVANTAGES / DISADVANTAGES OF PROPOSED SYSTEM

Advantages	**Disadvantages**
Enhanced Mobility and Independence	Technical Limitations
Improved Safety	Accessibility and Cost
Greater Access to Information	Privacy and Security Concerns
Improved learning opportunities	Limited field of view
Increased social interaction	Distraction and Disengagement
Navigation Assistance	Health Concerns

Table no 1

REFERENCES

[1] K. Sahithya, G. Lakshmi Tejaswi, K. Hari Gopal, B. Pavan Karthik. “A new method for recognition and obstacle detection for visually challenged using smart glasses powered with raspberry pi” International Journal of Engineering Applied Sciences and Technology, 2020 Vol. 5, Issue 1, ISSN No. 2455-2143, Pages 408-412

[2] Simra Nazim, Swaroop Ramaswamy Pillai, Vinod Kumar Shukla, Saba Firdous
“A Visual Assistant for the Blind”. 2022 International Mobile And Embedded Technology Conference (MECON).

[3] A. Nithiya Lakshmi, Stephi Jacob, Pommayan Thileja, “A Peculiar Reading System for Blind People using OCR Technology” 2023 Second International Conference on Electronics and Renewable Systems (ICEARS).

[4] R. Pushpakumar, P.N . Karthikayan, “Smart Glasses for Visually Impaired Using Image Processing Techniques”, International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC).

[5] Kun Xia, Xueyong Li , Haiyang Liu, Mingli Zhou, And Kexin Zhu, “IBGS: A Wearable Smart System to Assist Visually Challenged” (IEEE Access, 2022).

[6] A Wearable System for Obstacle Detection and Text-to-Speech for Visually Impaired Individuals (IEEE Sensors Journal, 2018)