Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

AI-Integrated Smart Glasses for Enhancing Reading and Guidance Independence for the Visually Impaired

Rajichellam J.[1], Baskaruni Dinesh Balaji Kumar[2], Bommannagari Usha Sai[3], Sangaraju Siva Karthik[4]

[1]Assistant Professor, [2-4]UG Student Department of Artificial Intelligence and Data Science, Nehru Institute of Engineering and Technology, Anna University, Coimbatore, India

Email: [1]nietjrajichellam@nehrucolleges.com

Abstract

Since there are blind and partially sighted people all over the world, they face significant challenges when it comes to reading. In the past, Braille was the only solution for the blind and visually impaired to read and learn. However, there is a lack of training in Braille and a shortage of Braille machines and facilities. To address these issues, a smart reader for the blind integrated with Raspberry Pi technology has been proposed. The system utilizes YOLO (You only look once) for object detection, OCR (Optical Character Reader) for text recognition, and Google Text-to-Speech to capture and read text aloud through a speaker or headphone. Smart glasses offer real-time assistance and enhance the daily lives of visually impaired individuals. Through a combination of sensors, cameras, and audio feedback systems, these smart glasses enhance the reading capability of the visual impaired. Moreover, advancements in artificial intelligence and Deep learning algorithms further enhance the capabilities of these devices, enabling more accurate personalized assistance.

Keywords: Raspberry Pi, OCR, YOLO, Google Text-to-Speech, Reading Assistance.

DOI: https://doi.org/10.36548/jtcsst.2024.3.002
Received: 02.06.2024, received in revised form: 22.06.2024, accepted: 03.07.2024, published: 18.07.2024

1. Introduction

More statistics can be found online and in books. Louis Braille, the creator of the popular Braille script, faced limitations due to the specialized and expensive presses required for printing in that unique format. Affordable printing presses could make Braille books more accessible. However, the availability of such books in libraries remains limited. Non-visually impaired individuals may not always find Braille books in regular print editions readily and consistently accessible. Therefore, this study introduces a cost-effective system for individuals with visual impairments to quickly read any printed e-book. The system is designed so that the user simply places the book down, and it handles all subsequent tasks: scanning the text, turning pages, and reading aloud. Additionally, it assists users in understanding unfamiliar words by providing dictionary definitions upon request, enhancing the reading experience. In addition, the user can pause the process by pressing a special button. Firstly, let's consider the fundamental challenge of accessing written content. Traditional printed materials pose significant obstacles for those who are blind, limiting their access to educational resources, literature, and everyday information. Screen readers, which convert text to speech, help mitigate these challenges. Moreover, the need for reliable guidance in navigating physical environments is essential for blind individuals to move safely and confidently. From GPS-enabled smartphone apps that provide step-by-step directions to indoor navigation systems utilizing Bluetooth beacons, these advancements are reshaping how blind individuals interact with their surroundings. In summary, enhancing reading and guidance independence for blind individuals is a multifaceted endeavor that offers the visually impaired the benefits of improved reading abilities and enhanced navigation guidance.

2. Literature Review

Queka et al. [1] explores the role of assistive technology in education, focusing on its impact and applications in facilitating learning for individuals with disabilities. Alkhalaf et al. [2] presents an OCR-based electronic documentation management system, indicating efforts to digitize and manage documents efficiently using optical character recognition technology. Abdallah et al [3] Introduces a braille system specifically designed for accessing religious texts like the Quran, emphasizing efforts to make religious materials accessible to visually impaired individuals through braille technology. Sabab et al [4] presents the Blind Reader, described as an intelligent assistant for individuals who are blind, indicating efforts to develop technology to aid in various tasks and activities. Ravi et al. [5] introduced a smart reader for the blind population, integrating Raspberry Pi technology. The system synchronizes with the page turning mechanism and includes a dictionary feature to provide an interactive session for the user. Asakawa et al. [6] presents auditory and tactile interfaces designed to convey visual effects to blind users, facilitating intuitive recognition of visual content encountered on the web. The Clique project by Parente et al. [7] aims to innovate by adapting graphical user interface (GUI) applications into audio-based formats that provide users with a conversational audio display. This approach focuses on task-oriented audio representations rather than relying on visual GUI elements. Velmurugan et al [8] describes the design and implementation of a system that converts text information present in images into speech and translates speech input from users into text. Dakopoulos et al. [9] presents a comparative survey of portable and wearable obstacle detection and avoidance systems, a subcategory of Electronic Travel Aids (ETAs). The survey aims to inform the research community and users about the capabilities and advancements in assistive technology for visually impaired individuals. Encalada et al. [10] addresses the development of English language reading skills and subskills among visually impaired individuals using the Braille system as a fundamental tool. To support the visually impaired in reading, navigation, and to enhance their confidence and independence [11, 12], the proposed method introduces the use of smart glasses.

3. Proposed Methodology

Innovative technologies like Raspberry Pi, keyboards, cameras, and ultrasonic sensors are becoming pivotal in enhancing reading and guidance independence for blind individuals. The devices leverage cutting-edge algorithms such as YOLO (You Only Look Once) to identify objects and texts in real-time, providing instant feedback to users. Advanced IP scanners important for ensuring seamless network connectivity between devices. Additionally, software platforms like MobaXterm facilitate efficient remote access and control, further enhancing the usability of these tools. Moreover, integrating text-to-speech (TTS) capabilities through solutions like GTT (Google Text-to-Speech) enables the conversion of textual information into auditory cues, aiding in navigation and comprehension. Paired with intuitive interfaces developed using Pygame, these technologies offer a comprehensive solution to improve the autonomy and quality of life for visually impaired individuals, enabling greater independence in reading and navigation task. Figure 1 shows the block diagram of the proposed.

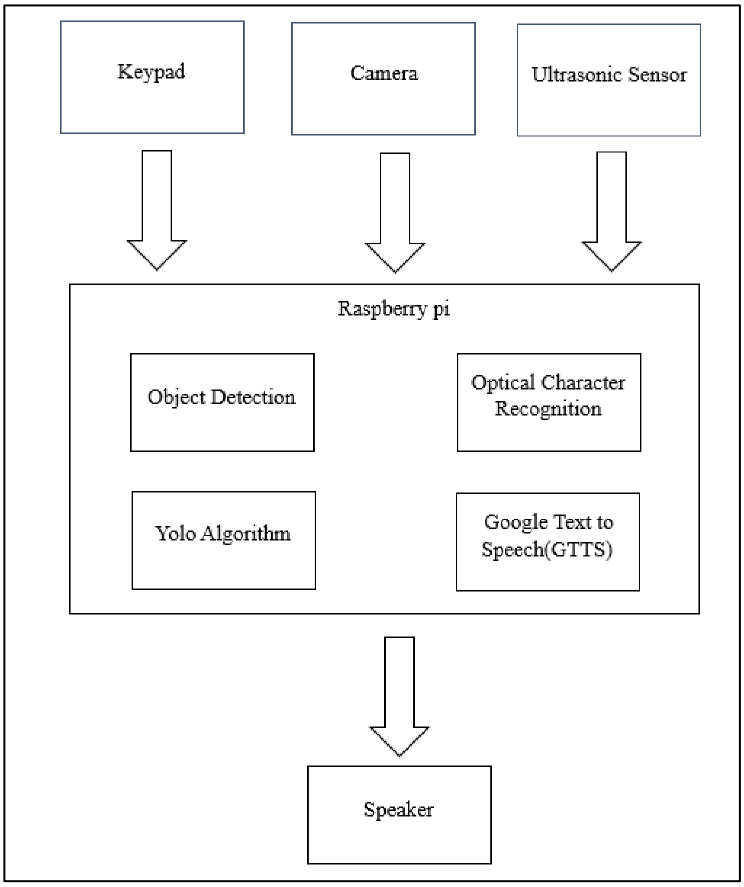


Figure 1. Block Diagram

The hardware and the software components used in the proposed smart glasses are depicted below in Table 1 and Table 2 respectively.

Table 1. Hardware Components

Components	Use	Specification
Raspberry Pi Board	Used for sensing, processing and connectivity.	Raspberry Pi 4 Model B
Keyboard	To get input commands or text (reading modes, object detection mode, adjusting settings etc)	4 Push Button Driver Module Switch
Camera	Capture text, objects.	OmniVision OV9282

Ultrasonic Sensor	Identify obstacles	HC-SR04 Ultrasonic Distance Sensor
Speaker	Delivers speech output	Samsung headphones.
Power Supply	Source of energy for the system.	Mean Well LRS-350-12 Power Supply

Table 2. Software Components

Components	Use
Python	used on the Raspberry Pi for writing code and integrating the components.
Pygame	Used for auditory feedback in the smart reader system
OpenCV	used for image and video processing tasks, including object detection with YOLO (You Only Look Once) for identifying objects in real-time captured by the OmniVision OV9282 camera.
Tesseract OCR	will recognize text from images captured by the OmniVision OV9282 camera.
Google Text-to-Speech (GTT)	convert the recognized text into spoken audio

3.1 Working of Object Detection

To start the Raspberry Pi, is connected to the power source through the cable. Once powered on, Advanced IP Scanner software is used to locate its IP address. The IP address is copied and pasted into MobaXterm's localhost for remote access. After connecting, user name and password are used to log in to MobaXterm. The all required packages and programs are pre-installed on the Raspberry Pi to run the code smoothly. The system uses a pre-trained program for real-time image detection, capable of identifying specific objects predefined in the program. To initiate Object Detection mode, press the first switch on the 4 push button driver module switch keyboard. The system captures an image using the camera and measures distance with an ultrasonic sensor. The captured image is processed by the YOLO model to identify pre-trained objects, displaying the object's name and measured distance through headphones to the user. To exit Object Detection mode, press the fourth switch on the keyboard module. Figure 2 shows how the YOLO model works.

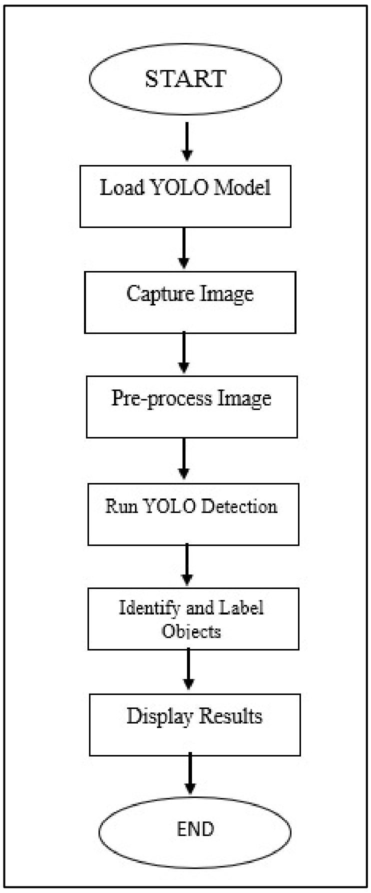


Figure 2. Flowchart of YOLO Model

3.2 Working of OCR

To activate the OCR mode, the second switch on the 4-push button driver module switch keyboard was pressed. Once activated, the system captured an image of English text using the camera. The captured image was processed through an Optical Character Recognition (OCR) model capable of identifying and extracting text. The recognized text was then translated into three different languages using Google Translate Text-to-Speech within the program. Users could select the desired language for audio output through earphones or speakers by pressing corresponding buttons on the keyboard module. This allowed them to hear the translated text in their preferred language. To exit OCR mode, the fourth switch on the 4-push button driver module switch keyboard was pressed.

This process provided a user-friendly interface for leveraging OCR capabilities and language translation using a Raspberry Pi, enhancing accessibility and functionality for users needing multilingual text recognition and audio output capabilities. The Figure 3 shows the flowchart of the OCR.

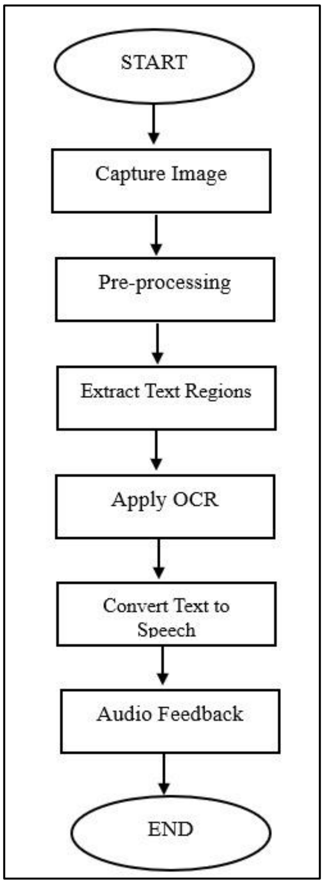
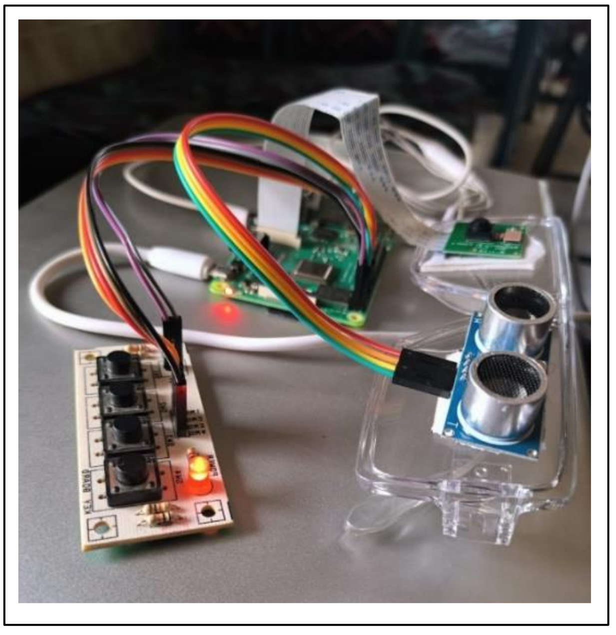


Figure 3. Flowchart of OCR Model

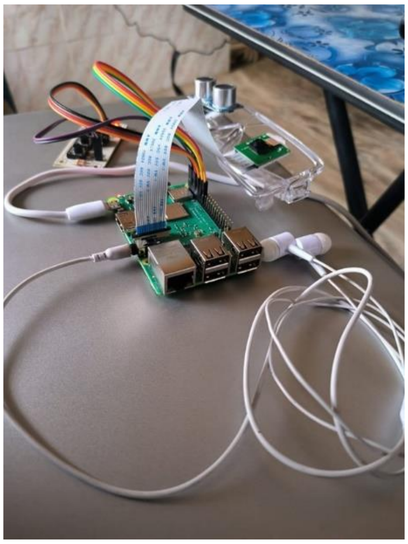
4. Results and Discussion

The proposed work utilized Python as the primary programming language for scripting and integrating essential software components in developing smart glasses. Pygame was employed to create an auditory feedback system. OpenCV handled image and video processing tasks, including object detection with YOLO using the OmniVision OV9282 camera. Tesseract OCR facilitated optical character recognition for text extraction from images, while Google Text-to-Speech (GTT) converted recognized text into spoken audio, delivered through connected headphones. Several key libraries and APIs such as NumPy for numerical operations, Pillow for image formats, PyAudio for audio input and output, SpeechRecognition for voice commands, Pygame Mixer for audio management, SQLite for local data storage, Socket programming for network communication, and GPIO Zero for Raspberry Pi GPIO control used enhanced the smart reader system, aiding visually impaired users in reading and navigation tasks. Figure 4 shows the hardware design of the product.

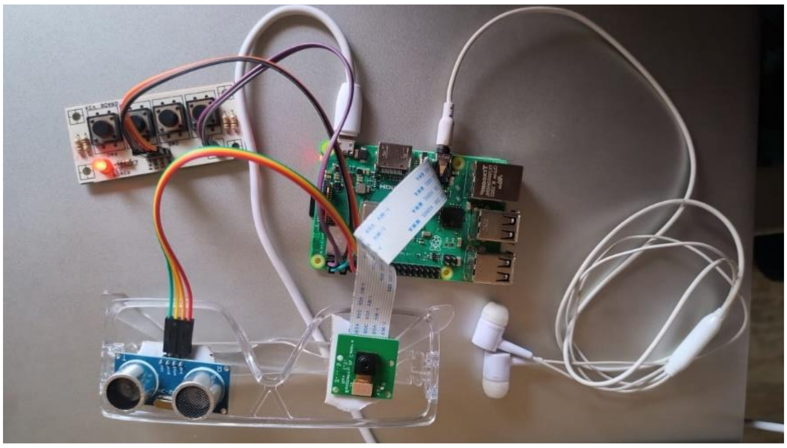
(a)



(b)



(c)



(d)

Figure 4. (a), (b), (c), and (d)Product Design

In the process of transforming images into digital data, object detection plays a crucial role. The Object detection algorithms identify and search for specific features in an image, first we have to identify bounding boxes and labels to each object detected. This process enables the extraction of relevant visual information from the image, such as identifying objects, or text. By utilizing object detection, images containing textual content, such as signs, documents, or labels, can be processed effectively, to extract the textual data. Second, optical character recognition (OCR) technology used to transform this textual data into digital data. OCR algorithms analyse the image regions containing text, recognizing individual characters and converting them into machine-readable text format. This enables the translation of printed or handwritten text. Finally, it gives result from earphones which is connected to Raspberry Pi. Figure. 5 displays the distance measured from the Ultrasonic sensor and the object from the camera.

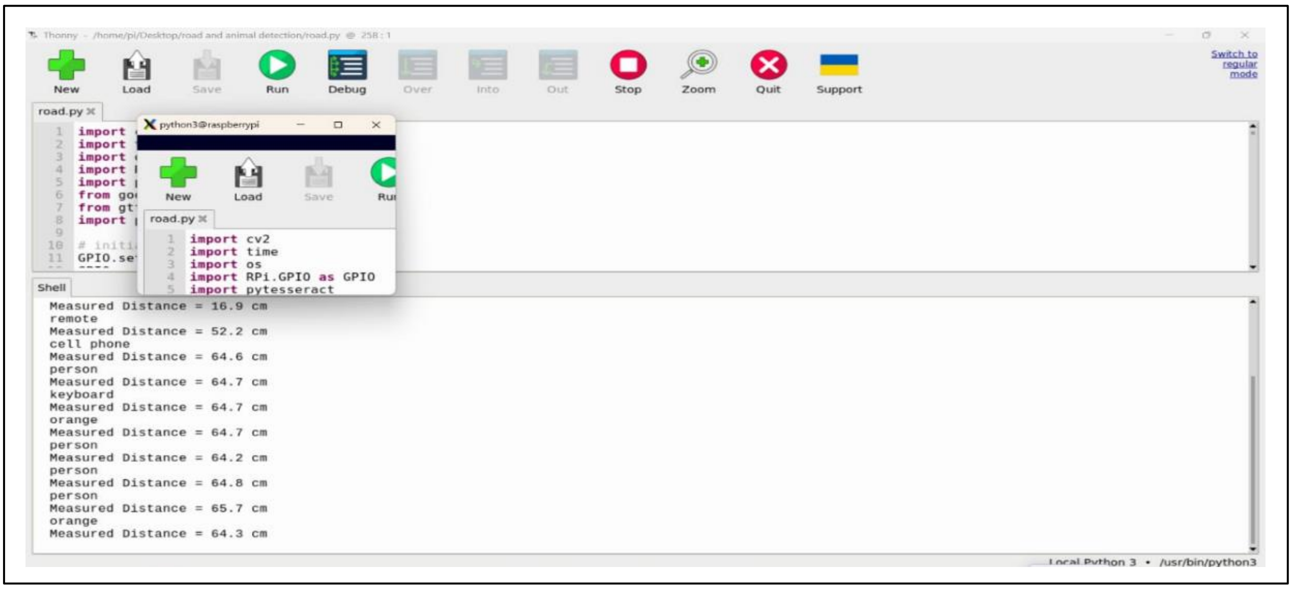


Figure 5. Displaying the Distance Measured from the Ultrasonic Sensor and the Object from the Camera

Google Text-to-Speech (gTTS) is a Python library and CLI tool designed for interacting with the Google Translate text-to-speech API. It simplifies converting text into speech with support for multiple languages and voices. This capability is valuable in applications requiring text-to-speech functionality, such as virtual assistants. Both OCR and gTTS empower applications to automatically extract and convert text from images or documents, aiding visually impaired individuals in accessing textual information and automating text-related task. Figure. 6 displays context of text which is done by using OCR and translating it through speaker.

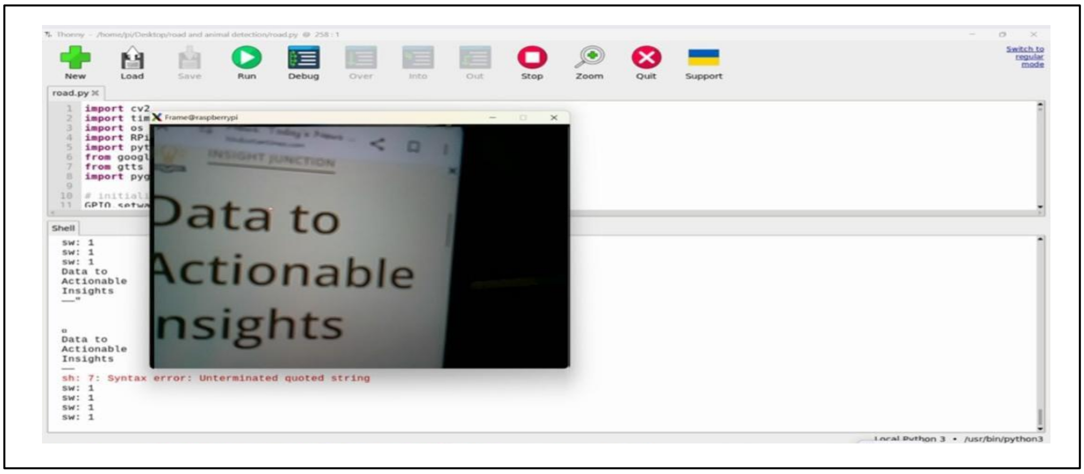


Figure 6. Displaying Context of Text which is Done by using OCR and Translating it Through Speaker

The integration of assistive technologies has significantly improved reading and independence for blind individuals. By combining hardware components and advanced software, researchers have made strides in addressing challenges faced by the visually impaired. Key tools like Braille displays, TTS systems, and OCR software effectively convert text, enabling independent access to printed materials and enhancing the reading capability.

Furthermore, in the realm of navigation and mobility, assistive technologies play a crucial role in enhancing the autonomy and safety of blind individuals. Ultrasonic sensors provide real-time feedback on surrounding obstacles, empowering users to navigate their environment with confidence. Similarly, computer vision algorithms, such as YOLO, offer robust capabilities in object detection and scene recognition, enabling blind individuals to interpret and interact with their surroundings more effectively.

The integration of hardware and software components further enriches the functionality and accessibility of assistive technologies. Raspberry Pi microcomputers serve as versatile platforms for deploying various applications, while cameras capture visual information essential for navigation and object recognition. Software platforms like MobaXterm facilitate seamless communication and remote access, enabling users to interact with assistive devices across different locations. Additionally, user-friendly interfaces developed using Pygame enhance the usability and acceptance of these technologies among blind users.

However, despite the significant progress made, several challenges and opportunities for improvement remain. Issues such as cost, usability, and interoperability continue to impact the widespread adoption of assistive technologies among blind individuals. Moreover, the need for ongoing research and development efforts to enhance the accuracy, reliability, and affordability of these tools is paramount.

5. Conclusion

In conclusion, enhancing reading and guidance independence for blind individuals is not just a matter of accessibility, but a fundamental step towards encouraging inclusivity and empowerment in society. Through the integration of innovative technologies like the Smart glasses integrated with sensors and advanced programming capabilities, we have the opportunity to break down barriers and create a world where blind individuals can navigate their surroundings with confidence, access information freely, and pursue their aspirations without limitations. By prioritizing inclusivity, collaboration, and ongoing innovation, we can continue to advance the cause of accessibility, ensuring that every individual, regardless of visual impairment, has the opportunity to participate fully in all aspects of life. Together, we can build a future where independence and equality are not just aspirations, but lived realities for all.

References

[1] Queka, Francis, Yasmine El-glalyb, and Francisco Oliveirad. "Assistive Technology in Education." (2015). Handbook of Science and Technology Convergence DOI 10.1007/978-3-319-04033-2_25-1# Springer International Publishing Switzerland 2015

[2] Alkhalaf, Khalaf S., Abdulelah I. Almishal, Anas O. Almahmoud, and Majed S. Alotaibi. "OCR-based electronic documentation management system." International Journal of Innovation, Management and Technology 5, no. 6 (2014): 465-469.

[3] Abdallah, J., M. Abualkishik, and Khairuddin Omar. "Quranic braille system." International Journal of Humanities and Social Sciences 3 (2009): 313-319.

[4] S. A. Sabab and M. H. Ashmafee, "Blind Reader: An intelligent assistant for blind," 2016 19th International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2016, pp. 229-234

[5] A. Ravi, S. Khasimbee, T. Asha, T. N. S. Joshna and P. G. Jyothirmai, "Raspberry pi based Smart Reader for Blind People," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 445-450.

[6] Asakawa, Chieko, Hironobu Takagi, Shuichi Ino, and Tohru Ifukube. "Auditory and tactile interfaces for representing the visual effects on the web." In Proceedings of the fifth international ACM conference on Assistive technologies, (Assets '02). Association for Computing Machinery, New York, NY, USA. 2002.pp. 65-72.

[7] Parente, Peter. "Clique: a conversant, task-based audio display for GUI applications." ACM SIGACCESS Accessibility and Computing 84 (2006): 34-37.

[8] Velmurugan, D., M. S. Sonam, S. Umamaheswari, S. Parthasarathy, and K. R. Arun. "A smart reader for visually impaired people using raspberry pi." International Journal of engineering science and computing 6, no. 3 (2016): 2997-3001.

[9] Dakopoulos, Dimitrios, and Nikolaos G. Bourbakis. "Wearable obstacle avoidance electronic travel aids for blind: a survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 40, no. 1 (2009): 25-35.

[10] Encalada, Edgar Guadia, Cristina del Rocío Jordán, Verónica Elizabeth Chicaiza, and Sarah Jacqueline Pazmiño. "Enhancing reading competence through the braille system for visually impaired people: a preliminary study." International Journal of Teaching and Learning 1, no. 1 (2021): 65-77.

[11] Kao, Matiekase A., and Patience J. Mzimela. "'They are visually impaired, not blind… teach them!': Grade R in-service teachers' knowledge of teaching pre-reading skills to visually impaired learners." South African Journal of Childhood Education 9, no. 1 (2019): 1-11.

[12] Gori, Monica, Giulia Cappagli, Alessia Tonelli, Gabriel Baud-Bovy, and Sara Finocchietti. "Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children." Neuroscience & Biobehavioral Reviews 69 (2016): 79-88.