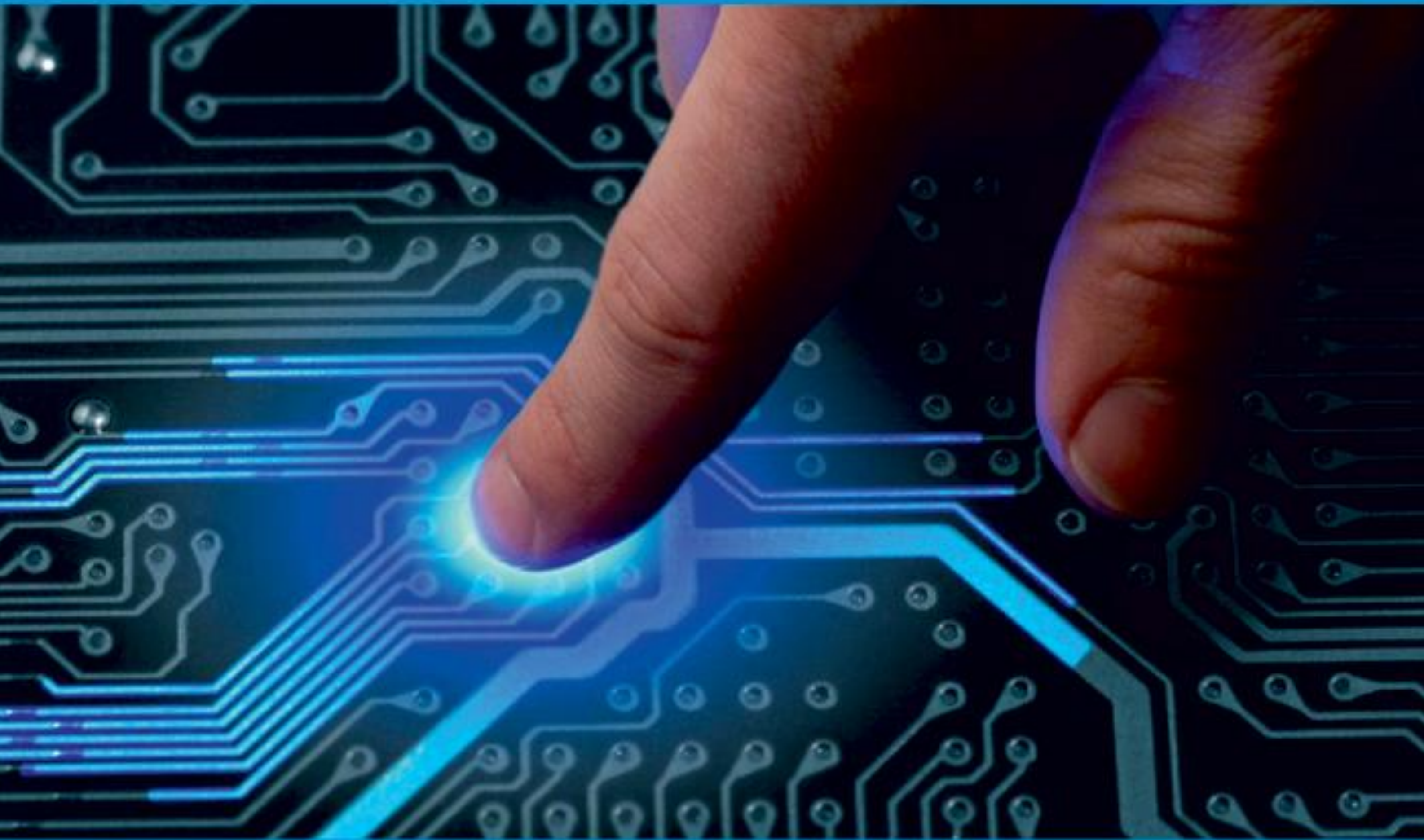
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

Implementation of the Raspberry-Pi Powered Optophone for the Vision-Impaired.

Dr. Prabha R[1], Pranav B C[2], Rajmahmadpasha[3], Shreya K[4], Siddhartha Bhatta[5]

Associate Professor, Department of Computer Science and Engineering, Dr. Ambedkar Institute of technology, Bangalore, Karnataka, India[1]

Students, Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, India [2,3,4,5]

ABSTRACT: The visually impaired people face a numerous challenge in interacting and navigating in the surroundings. It is always a hustle for them. So, to overcome that we have come up with this Raspberry-Pi powered Optophone for the visually-impaired individuals. This paper gives an idea of how the implementing and design of the portable optophone system is equipped with real time object recognition capabilities and is tailored to help the visually impaired individuals in identifying objects and moving around in the environment surroundings safely. This proposed system utilizes the computer vision technology and the cutting-edge object recognition algorithm to effectively recognize and classify objects in the user's surrounding. A compact camera module is present which captures the images of the surrounding environment which are analyzed and processed by an object recognition model. The system then delivers audio feedback to the user, giving the information about the recognized object, their location and other additional contextual details. In addition to this object recognition, the optophone system has the face detection functionality to assist the blind people in identifying and interacting with the people present in the surrounding. The object recognition model is optimized for efficiency to provide a seamless performance. This proposed optophone has the potential to empower the blind people/individuals to navigate in their surroundings confidently and engage in various activities with an increased autonomy.

KEYWORDS: Optophone, visually impaired individuals, Raspberry Pi, Real-time object identification and detection, OCR technology, Face detection.

I. INTRODUCTION

With respect to the growing emphasis on knowledge acquisition within the diverse communities in regard to the social and cultural advancements, there arises a concern about the enhancement of cultural enrichment among the visually impaired. While there are institutions dedicated to model and train the visually impaired youth, they encounter obstacles in providing the adequate materials. Customized books tailored for them are usually scarce and costly. Also, the production processes for the audio materials and the text which are read by the blind called the braille texts are intricate and time-consuming thereby posing a challenge to be able to access the reading materials. Additionally, the old-fashioned and traditional braille books pose handling difficulties particularly for the children and young people. These limitations provide a substantial gap in the knowledge acquisition for the blind thereby creating an urgent need for an effective solution.

To handle these challenges, we propose the integration of face detection and the identification functionalities into the optophone which enhances the reading experience for the blind individual. This optophone translates the written text into audible sounds offering a novel approach to the traditional reading methods. It gives us a cost-effective, fast and user-friendly way of expanding the knowledge about the visually impaired. Unlike the previous iteration for the optophone which lacked widespread adoption, this research capitalizes on the versatility of the Raspberry-Pi module, a prevalent component of the domain of IoT landscape. While various optophone devices have been explored previously, they have gained significant attraction among the general population leaving the blind individuals with a restricted access to the reading materials.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 4, April 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1204318 |

Earlier the studies conducted in Bangladesh and Pernambuco encountered challenges such as limited applicability, complexity in the suboptimal efficiency, thereby hindering the widespread adoption. Therefore, the development of an affordable, portable and a user-friendly reading solution for visually impaired individuals to prevail over the reading material limitations is imperative.

This research involves an optophone design that leverages the Raspberry-Pi module, a cornerstone of IoT and integrates with smartphones via a barrier-free mode which is tailored for the visually-impaired. The optophone integrates the optical character recognition (OCR) and the neural network technologies to create a versatile reading device which can be readily embraced by the virtually impaired community. Diverging from the previous optophone iterations, this approach harnesses the robustness and the portability of the Raspberry-Pi significantly explaining the available knowledge sources for the vision impaired individuals.

The research paper offers a comprehensive and exhaustive review of the optophone, highlighting the key modules and presents a functional prototype that has been developed for validation.

II. RELATED WORK

In [1], the author gives us some insights into the blind journey experiencing the written through the void of the other. Similarly, the concept of "spectral disability" is explored. Here, the spectacles are viewed as either a prosthetic device or an assistive technology that postpones the specter of disability.

In [2], a braille terminal is constructed comprising of 10 braille components. An eBook is translated into braille and stored in an USB drive. When inserted into the braille terminal, the corresponding braille components elevate to display the text.

In [3], this smart device has a multimodal system that has the potential to convert any document to an accessible document format for the blind. The user can read the document by tapping on the words which is then presented audibly through the text-to-speech technology.

In [4], a software answer is developed to capture digital textual content and convert it into braille characters. The information converted is then transmitted to a hardware system which reproduces it into a tactile braille sign providing an intelligent interface.

In [5], a real-time execution of the optophone is conducted. The optophone translates images from a camera into audible sounds thereby offering a non-invasive method for visually impaired individuals to be aware of their environments.

In [6], The photo OCR, which stands for Optical Character Recognition has been explained. This optophone system extracts text from images, therefore traditional OCR system limitations such blur, low resolution, low contrast and high image noise. This paper integrates the modern data center-scale distributed language modelling to enhance OCR performance.

III. THE OPTOPHONE SYSTEM

The figure below shows us a representation of the component design of the optophone system. It showcases its division into two primary components: hardware and software. In the hardware segment, the Raspberry Pi capsule provides a central unit while the acquisition and the playback component interface with the Raspberry-Pi collectively forming a cohesion reading entity. This system facilitates tasks such as collecting and processing book information as playing the back text content.

On the contrary, the software component portrays a significant role in realizing the reader function. The network image recognition extracts the reading materials, while the communication and the processing module of the Raspberry Pi facilitates the conveying controller and facilities the processing and collection of the reading information. In addition, the controller module handles the conveyance and transmits the control information.

Further, the including of the facial recognition features adds an additional layer of functionality to the optophone. By executing the facial recognition capabilities, the system can identify and also provide necessary and appropriate

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

information required and hence providing a comfortable user experience and also making the usability of the device easier.

Finally, the mixture of the software and hardware components enables the optophone model to achieve its intended design objective efficiently.

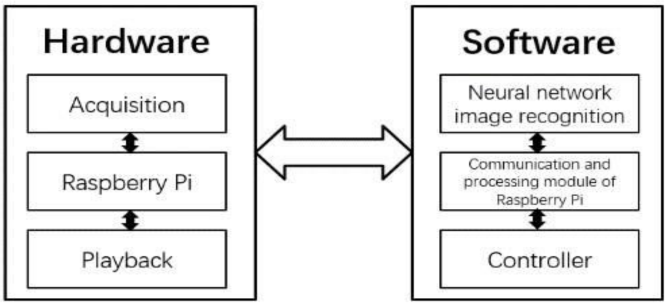


Figure 1: The Optophone system capsule design

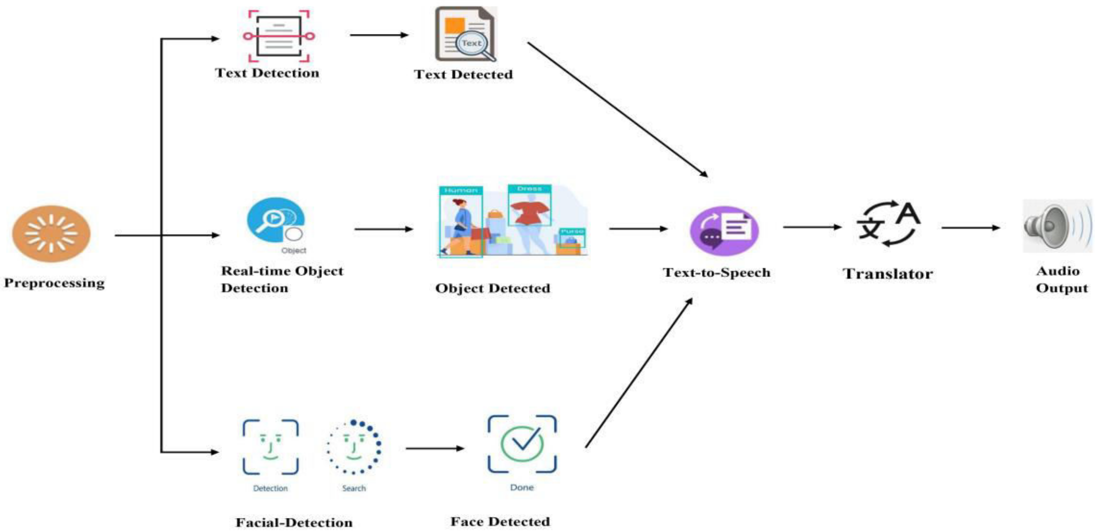


Figure 1(a): The system architecture of Optophone capsule.

IV. THE SYSTEM IMPLEMENTATION

A. *Hardware Requirements*

The fig below, fig 2. displays the hardware design, while fig 3. depicts the practical design of the optophone system. This section focuses the core function of the optophone which involves integrating the Raspberry Pi module, the acquisition module and playback module. However, both the acquisition and playback capsule are easily incorporated into the Raspberry-Pi module creating a unified unit within the hardware design.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

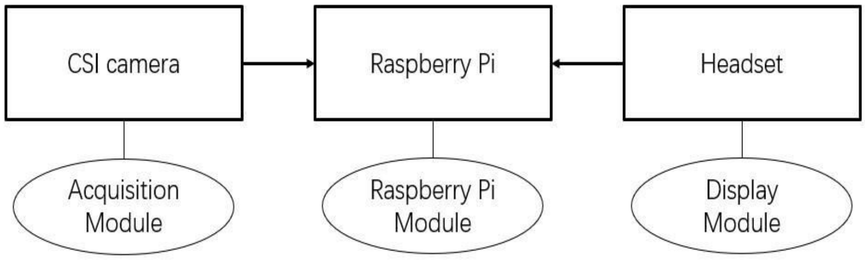


Fig. 2: Hardware design

Fig. 3 Hardware part diagram

1) *Raspberry-Pi capsule*

In the fig.4, the focal point is the Raspberry-Pi 3b+ which serves as the central element of the Raspberry-Pi capsule. The microcomputer motherboard which operates on the ARM architecture stands out for it compact and lightweight nature weighing only 50g and measures 10cm x 6cm. It is miniature which thereby enhances its portability. It also has a development board that has a component of 40 input and output pins, a quad processor, 1GB RAM, four USB ports, Ethernet, built-in Wi-Fi, a CSI camera interface and an audio interface. These attributes provide a seamless connectivity with the acquisition module and playback capsule, streamlining their integration with the Raspberry-Pi capsule.

Fig. 4 Raspberry-Pi Capsule

2) *Acquisition Module*

The fig 5 depicts the prosecution of the acquisition component, where the Raspberry-Pi's CSI camera is employed. It has a 5-megapixel resolution and a compact dimension of 32mm x 32mm, the CSI camera interfaces with Raspberry-Pi enabling the capture of images from printed materials. The amalgamation of the high pixel density and miniature guarantees the precise extraction of information from the reading materials thereby enhancing the user convenience.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

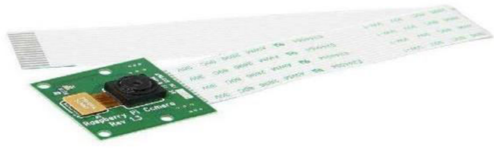


Fig 5. Acquisition Module

3) Playback Component

The below figure, fig 6. shows the playback component. The playback item transmits the reading material information to the users and it is established through the speakers. These speakers have an audio interface which is bridged to the Raspberry-Pi's audio interface. This configuration ensures the effective functioning of the playback component and it delivering its intended purpose.

Fig 6. Playback Module

4) CSI Camera

Implementing the CSI camera into the optophone system has major advantages for the vision impaired individual. These CSI cameras gives high-resolution image capture capabilities facilitating a precise object recognition. Seamless integration with the handheld devices provides a user-friendly experience, simplifying the operation and the interaction. However, the advanced features like low-light sensitivity and adjustable focus enhances the functionality of the optophone which helps the destitute of vision individuals to navigate their surroundings and identify objects with an increased efficiency and accuracy.

Fig 7: CSI Camera

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

B) Software Requirements

The software design diagram is shown below in fig 7. It shows the backcloth technique of the optophone. The two fundamental inquiries that arises related to the transmission between the Raspberry-Pi and the optophone are: (1) Can the visual information be translated into auditory signals? (2) Can the individuals gain the ability to interpret these auditory cues? The extrication of the text from instructed images poses a challenge to numerous practical implications. The neural network image recognition technology gives a propitious solution to these challenges.

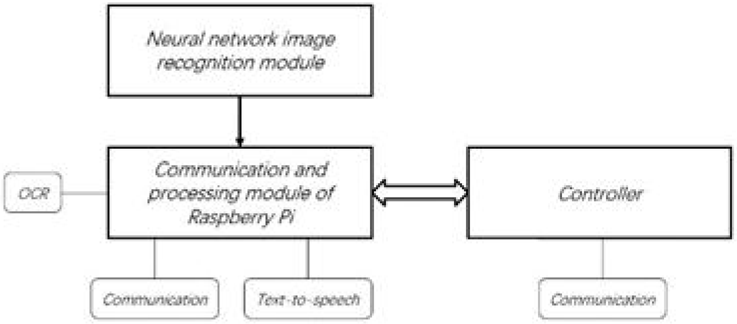


Figure 7: The design diagram of the software requirement part

1) Neural network image recognition module

The process of generating text for images establishes a connection between visual content and language. This process involves two main components: image feature extraction and language generation. Image feature extraction represents the visual content of an image as a vector or a matrix. High-quality image features are crucial for accurately reflecting the visual content and ensuing the effectiveness and accuracy of generation of text. Language generation focuses on linking the image geatures with semantics to generate to the point descriptive text.

The implementation process involves several steps:

a) We configure the environment in the Windows including TensorFlow, NumPy and NLTK.

b) The image title is generates with an attention mechanism combining the variants "hard" and "soft" to extract the feature vectors using convolutional characteristic encoder. To generate words in a sequential order based on context vectors, hidden states and previously generated words to form titles we use the Long Short-Term Memory (LSTM).

c) Using the conform learning rate algorithm we train the attention model with the RMS prop which performs well on the flickr8k dataset. To create annotations for the decoder we use a pre-trained image network without fine-tuning that is The VGGNet of Oxford.

d) Utilizing the flickr8k dataset, which consists of 8000 images, each co-existed by 5 reference sentences. A fixed glossary size of 6000 is used in all experiments.

e) After the training is done in the Windows environment, we test it on the model by configuring the required environment on Raspberry-Pi and then transferring this trained model to an SD card to run test.py.

Raspberry-Pi consists of three parts for divulgation and processing the module: communication, processing anf acquisition/playback. The various information which is received from the transmission part is processed, including collecting reading materials, citing image and text information using the neural network image recognition. The OCR interfaces and displays the processed text information using a text-to-speech interface.

An implemented App page is involved in the controller to facilitate the interdependence between the Raspberry-Pi and the controller, file transfer and other hardware controls. It is divided into four functional sections - Connect Raspberry-Pi, play/pause, select files and control photo taking —this module is developed as Android client mobile phone software in barrier-free mode, named Optophone controller. The interface design prioritizes ease of operation for individuals with vision impairment, with all button options featuring voice prompts.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

Figure 8: Optophone Controller

a) Initially, initializing the inter-relatedness through the connection button is imperative. It makes sure that the mobile phone is linked either to the same LAN as the Raspberry Pi or to the open hotspot of the Raspberry Pi. The confirmation for this is provided through the voice prompt.

b) In the post-connection establishment process, the users can control the initiating or pausing of the current reading playback through the play/pause button. Upon resuming, the playback resumes from the last paused position for the user's convenience.

c) To choose files on their mobile phones, the user can use the system's built-in file selector and these files can be in formats such as txt or pdf. On transmitting, the Raspberry-Pi receiver downloads the files to the specified storage location allowing for reading or performing other operations.

d) The photo capture feature allows to capture the physical books, magazines or photos. Once it is connected, the camera becomes operational and the Raspberry Pi reads the information. Subsequently, Raspberry-Pi can provide users with reading materials from pre-existing storage and allow to add reading materials via mobile phones and enable scanning and inputting reading materials. This offers an informative array of reading materials and is done via the controller.

V. OPTOPHONE RESULTS

Object Recognition and Audio Description: Using the advanced technology, the system will identify the objects in its view and will announce the nomenclature of the object aloud in the language that the user prefers. This feature will help the users to move around and navigate in their surroundings with more confidence and much more awareness.
Printed Text Translation and Reading: The interpreted printed text is scanned by the system irrespective of the original language. With the help of powerful translation tolls, the system translates the text into the preferred language. This functionality will eradicate language barriers thereby allowing users and understand the readily printed information.
Facial Detection and Recognition: The optophone designed is capable of detecting and recognizing the faces thereby providing the users auditory cues about the familiar individual such as family members and hence improving social interaction and recognition for the visually impaired.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 4, April 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1204318 |

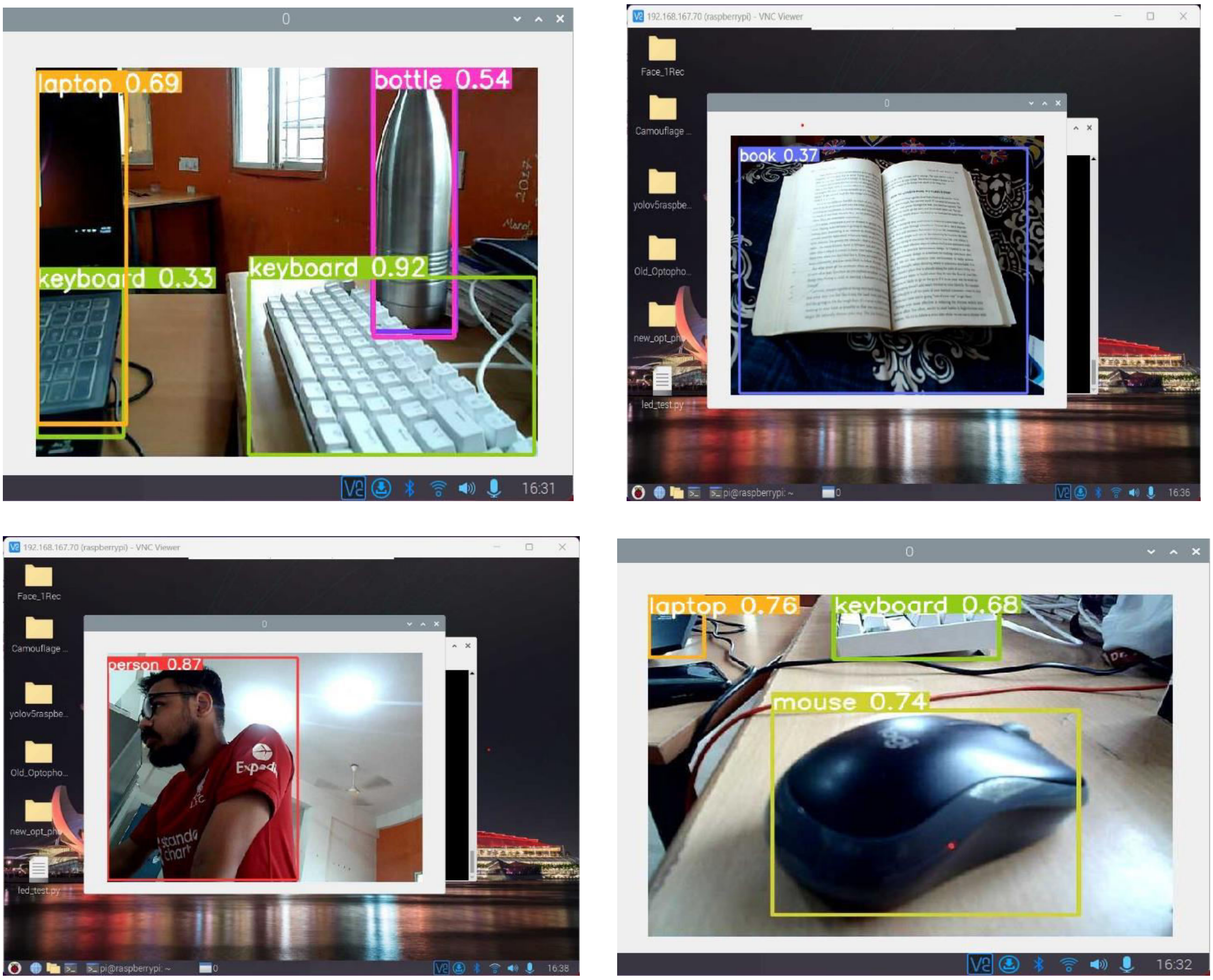


Fig 9: Object detection in the surroundings

VI. CONCLUSION AND FUTURE WORK

In conclusion, this paper has ushered in the development of a portable optophone which enables real time object recognition with the aim of improving the mobility of individuals with visual impairments. By using a CSI camera and a high-performance mobile platform, the system provides precise and fast feedback providing seamless interaction in real time. This research contributes to the assistive technology for the visually impaired but providing a solution that offers object recognition with portable functionality. For the future enhancements it could explore future endeavors to the system which includes the integration of additional features and helps for broader adoption in the practical settings. Finally, and ultimately, the portable optophone has a significant promise in boosting greater confidence and independence among the visually impaired individual so that they can navigate in the surroundings confidently.

REFERENCES

[1] Hsy J. explores the theme of "Close listening: Talking books, blind readers and medieval worldbuilding" in the Journal of postmedieval: a journal of medieval cultural studies (2016, 7(2), pp. 181-192).
[2] Velázquez R., Hernández H., and Preza E. present a study on "A portable eBook reader for the blind" in the Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2010 (pp. 2107-2110).
[3] Sabab S. A., and Ashmafee M. H. discuss "Blind Reader: An intelligent assistant for the blind" in the Proceedings of the International Conference on Computer & Information Technology, IEEE, 2017 (pp. 229-234).

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204318 |

[4] Braga DDS., Neto FBDL., and Oliveira SC. present research on "An Intelligent Human Machine Interface for Reading Digital Texts for The Blind" in the Proceedings of the Brazilian Congress of Computational Intelligence, 2016 (pp. 1-7).
[5] Elias M., Eltner A., Liebold F. et al. assess "The Influence of Temperature Changes on the Geometric Stability of Smartphone- and Raspberry Pi Cameras" in the Journal of Sensors, 2020, 20(3), pp. 643.
[6] Picton P.D. discusses "Real-time implementation of an optophone" in the Proceedings of the Colloquium on Computers in the Service of Mankind: Helping the Disabled, IEEE, 1997 (p. 6).
[7] Bissacco A., Cummins M., Netzer Y., et al. present "PhotoOCR: Reading Text in Uncontrolled Conditions" in the Proceedings of the 2013 IEEE International Conference on Computer Vision (ICCV), IEEE, 2013 (pp. 785-792).
[8] Lecun Y., Bengio Y., and Hinton G. review "Deep Learning" in the journal Nature, 2015, 521(7553):436.