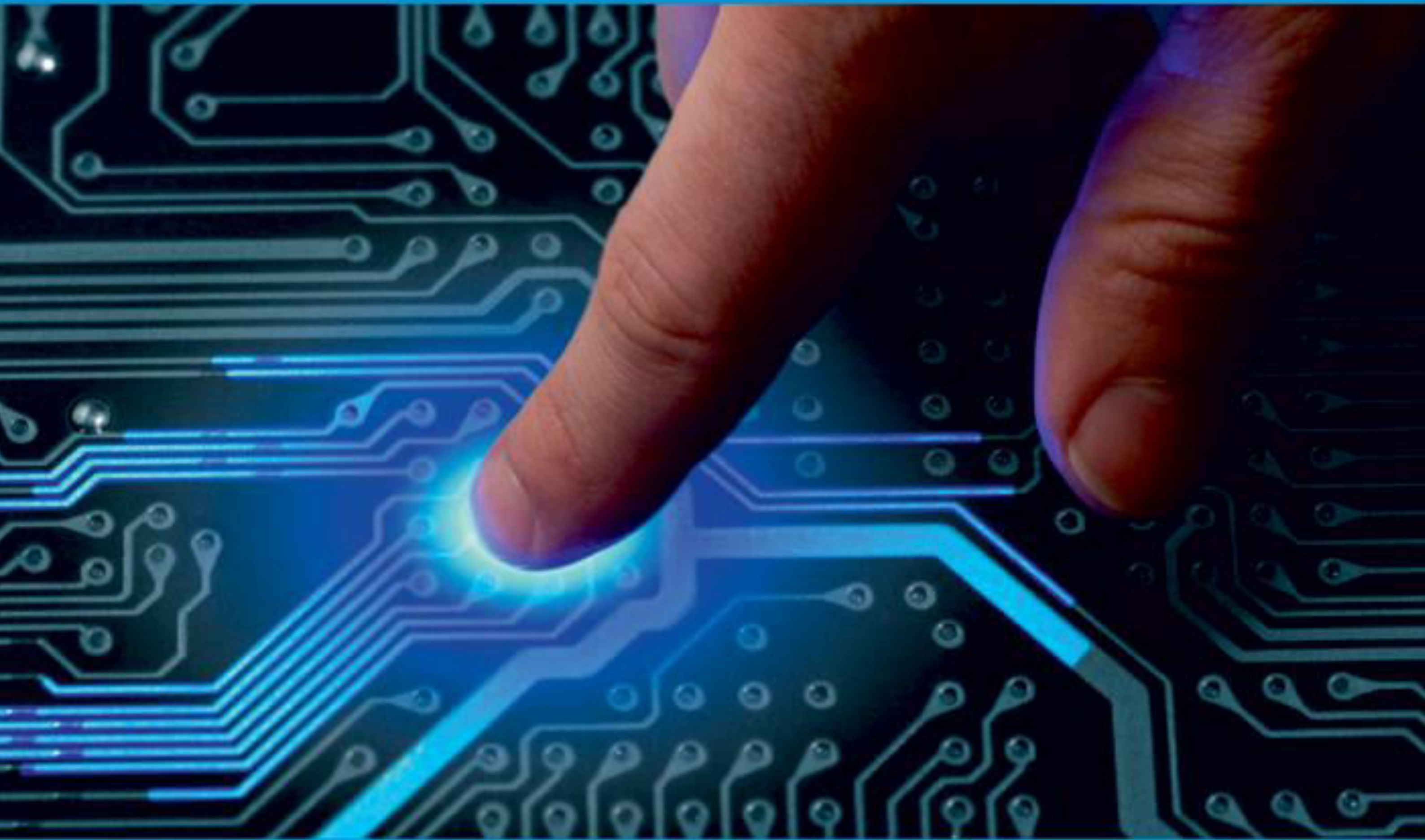
IJIRCCE

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



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 6, June 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

Implementation of a Portable Optophone System with Real-Time Object Recognition for Visually Impaired Person

[1] Dr. Smitha Shekar B., [2]Dr. Harish G., [3]Amrit Ranabhat, [4]Keshav Regmi, [5]Kushal Marahatta, [6]Suraj Belbase

Associate Professor, Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India[1,2]

Students, Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India[3,4,5,6]

ABSTRACT*:* The visually impaired community faces significant challenges in perceiving and interacting with the world around them. In recent years, advances in technology have provided innovative solutions to improve the quality of life for visually impaired individuals. This paper presents the design and implementation of a portable optophone system with real-time object recognition, aimed at assisting visually impaired persons in identifying objects and navigating their environment. The proposed optophone system leverages computer vision techniques and state-of-the-art object recognition algorithms to recognize and classify objects in real time. A portable camera module captures images of the user's surroundings, which are then processed and analyzed by a robust object recognition model running on a handheld device. The system provides audio feedback to the user, relaying information about the recognized objects, their location, and additional contextual details.To ensure real-time performance, the object recognition model is optimized for efficiency and deployed on a high-performance mobile platform. The system incorporates advanced features such as depth estimation, enabling the user to perceive the spatial layout of the objects in their environment. Additionally, the optophone system can adapt to different environments and lighting conditions, ensuring accurate and reliable object recognition across various scenarios.User feedback and usability testing were conducted with visually impaired individuals to evaluate the effectiveness and user-friendliness of the system. The results demonstrated that the portable optophone system significantly improves the user's ability to identify objects independently, thereby enhancing their mobility and overall independence.This research contributes to the field of assistive technology for the visually impaired by providing a practical solution that combines object recognition and real-time feedback. The proposed optophone system has the potential to empower visually impaired individuals, allowing them to navigate their surroundings more confidently and engage in various activities with greater autonomy.

KEYWORDS: Optophone, Visually impaired individuals, Raspberry Pi, Real-time object identification, OCR technology.

I. INTRODUCTION

With the increasing emphasis on knowledge acquisition in various groups due to social and cultural development, there is a growing concern about improving the cultural quality of visually impaired individuals. Institutions focused on training visually impaired youth have emerged, but they face challenges in providing adequate reading materials. The existing customized books are limited in availability and expensive. Additionally, the production processes for braille texts and audio materials are complex and time-consuming, further limiting access to reading materials. Moreover, braille books are cumbersome and inconvenient to handle, particularly for young people and children. These limitations contribute to a significant imbalance in knowledge acquisition for visually impaired individuals, creating a pressing need for a solution.

To address these issues, the optophone is proposed as a potential solution. The optophone translates written words into sounds, allowing visually impaired individuals to utilize their hearing for reading. It aims to overcome the limitations of existing reading sources for the visually impaired, providing an efficient and affordable method to broaden their knowledge base. By translating written words into sounds, the optophone empowers sight-disabled individuals to utilize their auditory senses for reading, thereby bypassing the inherent limitations of traditional reading methods. It offers a relatively low-cost, fast, and easy-to-implement solution. Unlike previous optophone designs, which have not gained widespread adoption, this research leverages the potential of the Raspberry Pi module, a popular component in

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

the Internet of Things ecosystem. Although various optophone devices have been studied, they have not been widely adopted by the general public, leaving visually impaired individuals with limited access to knowledge. Previous studies conducted in Bangladesh and Pernambuco faced challenges such as a narrow scope of application, complex operation, and low efficiency, hindering widespread adoption. Therefore, developing an affordable, portable, and user-friendly reading method for visually impaired individuals to overcome limitations in reading materials is an urgent task.

This research paper introduces a new optophone design that utilizes the Raspberry Pi module, popular in the Internet of Things, and integrates with smartphones in a barrier-free mode developed for the blind. The optophone incorporates optical character recognition (OCR) and neural network technologies to create a reading device that can be easily adopted within the visually impaired community. Unlike previous optophone designs, this new approach utilizes Raspberry Pi, offering robustness, portability, and significantly expanding knowledge sources for visually impaired individuals.

The paper provides a technical overview of the optophone, highlighting its main modules, and presents a prototype that has been developed.

II.RELATED WORK

In [1], the writer has expressed his opinions about a character's experience as a blind poet, experiencing the written word through the voice of another. Similarly, the concept of "spectral disability" is discusssed, where the spectacles are interpreted as a prosthetic device or an assisitive technology that delays an inevitable specter of disability. In [2], a brialle terminal is constructed which contains 10 braille components and an eBook is translated into braille and stored in an USB. This USB is inserted into the Braille terminal and respective Braille component rises up. In [3], a smart device is introduced with a multimodal system that can convert any document to the interpreted form to a blind. A blind can read document only by tapping words which is then audibly presented through text to speech. In [4], a software is created which captures the contents of a digital textual document and converts its digit characters into Braille. After this conversion, the information is sent to the hardware of the system that reproduces them into an intelligent interface for tactile Braille signs. In [6], real-time implementation of optophone is carried out. The optophone is not an echo-locator; it is a device that converts images from a camera into audible sounds. In [7], Photo OCR is described. OCR stands for Optical Character Recognition. This is a system for text extraction from images. This paper incorporates modern data center-scale distributed language modelling which overcomes the drawbacks from tradional OCR systems, notably in the presence of substantial blur, low resolution, low contrast, high image noise and other distortions.

III. THE OPTOPHONE SYSTEM

The module design of the optophone system is illustrated clearly in Fig.1, showcasing its division into two key components: hardware and software. The hardware component encompasses the Raspberry Pi module, acquisition module, and playback module. On the other hand, the software component consists of the neural network image recognition module, Raspberry Pi module's communication and processing module, and the controller module.

In the hardware design, the Raspberry Pi module serves as the central component, while the acquisition module and playback module are connected to the Raspberry Pi through interfaces, forming a comprehensive reader entity. This integrated system facilitates functions such as book information collection and processing, as well as text content playback.

The software module design plays a pivotal role in realizing the reader function. The neural network image recognition module is responsible for extracting reading materials. The communication and processing module of the Raspberry Pi enables communication with the controller and facilitates the collection and subsequent processing of reading information. Meanwhile, the controller module handles communication with the reader topic and the transmission of control information. The integration of various technologies within these modules enables the resolution of technical challenges.
Ultimately, the combination of the software and hardware components culminates in achieving the intended design objective of the optophone system.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

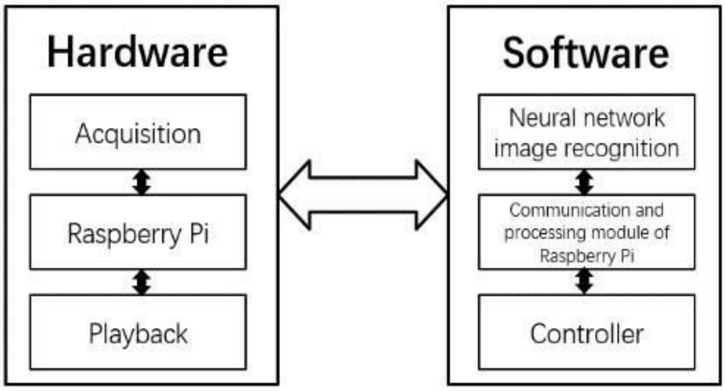


Figure 1: The module design of the Optophone system

III. THE SYSTEM IMPLEMENTATION

A. Hardware Part

Fig. 2 presents the hardware design, while Fig. 3 showcases the practical implementation of the optophone system. This section focuses on fulfilling the primary function of the optophone by incorporating the Raspberry Pi module, acquisition module, and playback module. Specifically, the acquisition module and playback module are integrated onto the Raspberry Pi module, forming an interconnected unit within the hardware design.

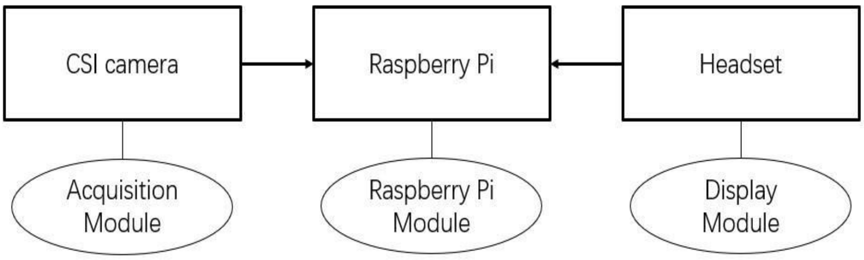


Fig 2 Hardware design

Fig. 3 Hardware part diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

1) Raspberry Pi Module

The central component of the Raspberry Pi module is highlighted in Fig. 4, utilizing the Raspberry Pi 3b+. This microcomputer motherboard, which operates on ARM architecture, is characterized by its compact and lightweight design, weighing only 50g and measuring 10cm × 6cm. This compactness enhances its portability. The development board features various components, including 40 input and output pins, a quad-core processor, 1GB RAM, four USB ports, Ethernet, built-in Wi-Fi, a CSI camera interface, and an audio interface. These features enable seamless connectivity with the acquisition module and playback module, facilitating their integration with the Raspberry Pi module.

Fig. 4 Raspberry Pi Module

2) Acquisition Module

The implementation of the Acquisition Module is depicted in Fig. 5. In this module, the CSI camera of the Raspberry Pi is utilized, possessing 5 megapixels and compact dimensions of 32mm × 32mm. By connecting to the interface of the Raspberry Pi, the CSI camera facilitates the acquisition of images from reading materials. It is essential to maintain stability during image capture [5]. The combination of high pixel resolution and small size ensures the extraction of accurate information from reading materials, enhancing the convenience of use.

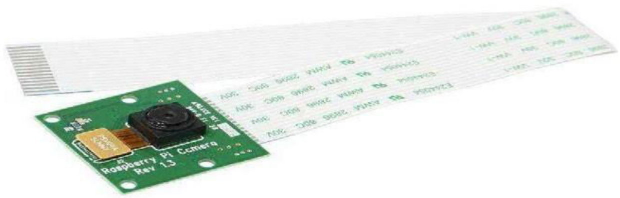


Fig 5. Acquisition Module

3) Playback Module

The implementation of the playback module is depicted in Fig. 6. To achieve the transmission of reading material information to users, the playback module is established by connecting the playback device, primarily a wired headset with an audio interface, to the Raspberry Pi audio interface. This setup enables the playback module to fulfill its intended function effectively.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

Fig 6. Playback Module

4) CSI Camera

The incorporation of a CSI (Camera Serial Interface) camera within an optophone system holds significant advantages for visually impaired individuals. CSI cameras provide a high-resolution image capture capability, allowing for detailed and accurate object recognition. With their fast data transfer rates, CSI cameras enable real-time processing of captured images, enabling prompt and immediate feedback to the user. Their compact form factor makes them well-suited for integration into portable optophone systems, enhancing the system's mobility and convenience. By seamlessly integrating with handheld devices, CSI cameras contribute to a user-friendly experience, facilitating easy operation and interaction. Furthermore, the advanced features offered by CSI cameras, such as low-light sensitivity and adjustable focus, enhance the overall functionality of the optophone system, enabling visually impaired individuals to navigate their surroundings and identify objects with improved efficiency and accuracy.

Fig 7: CSI Camera

B. Software Part

The design diagram of the software part is shown in Fig.7. The software part completes the background processing of the optophone, the communication between the Raspberry Pi and the idea of an optophone raises two questions. (1) Can information in an image be converted into information in sounds? (2) Can humans learn to interpret these sounds?[6] The extraction of text from uncontrolled images is a challenging problem with many practical applications.[7] The neural network image recognition technology can solve these problems relatively well. Deep learning algorithm is a multilevel feature learning method. It uses simple non-linear components to transform the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

features of each layer into higher-order and more abstract level features. It combines multiple feature transformations to fit more complex functions.

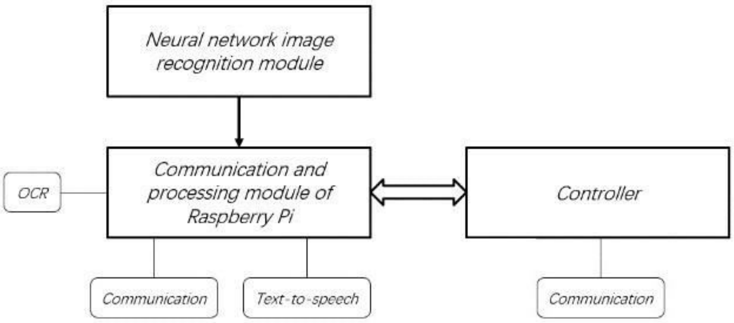


Figure 7: The design diagram of the software part

According to the report on basic information of Internet users with visual impairment in China issued by CAPA in 2016, 92% of the sight-disabled people in China are using smart phones. 33% of the sight-disabled people have 11-20 applications installed on their mobile phones, 83% of them can fully rely on the screen reading function when they operate their mobile phones and computers. Due to the lack of vision, the hearing of Sight-Disabled People is often more acute than that of ordinary people. In the process of the design of the blind mobile phone, the advantage of the sight-disabled people should be used reasonably. The operation of the mobile phone should be converted into the voice reminder mode. According to the voice reminder and voice feedback, various operations should be designed. The sound interaction interface suitable for the sight-disabled people operation must greatly improve the convenience of the sight-disabled people in the operation of the mobile phone. With the birth of the barrier free auxiliary function of the intelligent machine, the sight-disabled users can enjoy the same opportunity to use the intelligent machine app as the sight-disabled user group. It conveys the visual interface information of App to sight-disabled users through voice prompt, so as to help them better use App to obtain information. Using the existing open source OCR interface and text to speech interface at Raspberry Pi to complete the speech conversion of reading materials. The communication between Raspberry Pi and controller is completed by TCP / IP protocol stack. TCP/UDP/IP and Ethernet have become undisputed leaders of network. Communication since the last decade, defined by IETF's RFC793.And the two programs on the network through a two way communication link for data exchange, one end of the connection is called a Socket.

1) Neural network image recognition module

The key of image descriptive text generation is to connect images with language. The model of generating image descriptive text can be divided into two parts, extraction of image features as well as language modeling and generation. Image feature extraction is to express the image through a certain vector or matrix. Extracting enough good image features can reflect the visual content information of the image to the greatest extent, which also directly affects the effect and accuracy of subsequent description text generation. Language modeling and generation is the key part of image descriptive text generation, its main task is to directly link image features with semantics. How to use image features to generate appropriate descriptive text is the focus of this part.

The basic principle of the neural network image recognition module is image title generation with an attention mechanism. Introducing an attention-based model, which can automatically learn and describe the content of images. The following describes how to train the model in a deterministic way using standard back propagation techniques, and how to randomly train the model by maximizing a changing lower bound. It shows how the model can generate the corresponding words in the output sequence, at the same time, it can automatically learn how to focus on the highlighted objects, and validates the use of attention and real-time performance state on the benchmark data set(flickr8k). The specific implementation process is as follows

a) Train the deep learning model in the Windows environment. The first step is to configure the required environment: TensorFlow, NumPy, NLTK etc.
b) Generate title of image with attention mechanism. The form of attention is combined with two variants of "hard" attention and "soft" attention to extract a set of eigenvectors to form a word sequence by using a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

convolutional characteristic encoder (convolutional neural network). Then, use the decoder of Long Short-Term Memory (LSTM) to generate one word at each time step according to context. vectors, the foregoing hidden state and the words generated previously, thus generate the title.

c) Training procedure. The two variables of the attention model are trained by the adaptive learning rate algorithm. For flickr8k dataset, RMSprop [8] works best. In order to create the annotation AI which would be used by decoder, VGGNet of Oxford, an image network pre-training without fine tuning, is used.

d) Use the data in flickr8k dataset. There are 8000 images, and each image is accompanied by 5 reference sentences. In all experiments, the fixed vocabulary size used was 6000.

e) Run test the model. After completing the deep learning model training in the Windows environment, configure the required environment in Raspberry Pi, and copy the trained model to SD card to run the test.py.

2) Communication and processing module of Raspberry Pi This module is divided into three parts: communication part, processing part, acquisition and playback part. The whole module runs in RASPBAIN system and is developed by Python 3.5. The specific implementation is as follows:

a) In the communication part, Raspberry Pi as the server side of the TCP/IP protocol stack to communicate with the controller. After receiving the control information, the communication part makes a simple judgment and submits it to the processing part for processing. After one processing, it continues to monitor until the whole reader is closed.

b) In the processing part, according to the different information received by the communication part, the processing part can be divided into three categories: collection of reading materials, processing picture information, processing text information and display. The collection of reading materials is completed by using the built-in instructions of the system, the processing of picture information is completed by using neural network image recognition and the open-source OCR character recognition interface, and processing text information and display are completed by using the open-source text to speech interface.

c) The part of acquisition and playback is assisted by the acquisition module and play module of hardware module and the built-in instruction of the system.

3) Controller Fig.8 shows the implemented App page of this module. This App realizes the connection between Raspberry Pi and controller, transfer files and connect other hardware control. This part is divided into four functional sections: Connect Raspberry Pi, play and pause, select files, control photo taking. This module is realized by developing the Android client mobile phone software in barrier free mode. This App takes Android studio as the development environment, named Optophone controller. The interface design is based on the convenient operation of people with visual impairment. The button options all have voice prompts. The specific implementation is as follows:

Figure 8: Optophone Controller

a) During initial use, it is necessary to initialize the connection through the connection button. To ensure that the mobile phone has connected to the open hotspot of Raspberry Pi or in the same LAN as Raspberry Pi, the connection between the controller and Raspberry Pi will be confirmed in the form of a voice prompt.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106009 |

b) In the normal use process after the connection, the user can control to start or pause the play of the current reading through the play/pause button. After the pause, the play will start from the last paused position for easy use.

c) The file selecting section is used to realize file sharing with raspberry pie memory. The user can call the system self set file selector on the mobile phone to select files. The file format can be txt, pdf etc. After sending, the Raspberry Pi receiver completes receiving and downloads the files to the designated storage location and can read or operate other operations.

d) The photo taking section is used to take physical books (books, magazines or photos). After connecting, the camera can work. Raspberry Pi will read the information. Therefore, through the controller, Raspberry Pi can provide users with reading materials from pre-storage, mobile phone adds reading materials, scan input reading materials, and provide users with a comprehensive range of reading materials to the greatest extent.

V. CONCLUSION

In conclusion, this research paper has presented the implementation of a portable optophone system with real-time object recognition, aimed at enhancing the independence and mobility of visually impaired individuals. The system leverages computer vision techniques and state-of-the-art object recognition algorithms to recognize and classify objects in the user's surroundings. By utilizing a CSI camera and a high-performance mobile platform, the system provides fast and accurate feedback, enabling real-time interaction. Usability testing with visually impaired individuals has demonstrated the effectiveness and user-friendliness of the optophone system, highlighting its potential to empower users and improve their quality of life. This research contributes to the field of assistive technology for the visually impaired, offering a practical solution that combines object recognition and portable functionality. Future work can focus on expanding the system's capabilities, integrating additional features, and exploring opportunities for wider adoption in real-world scenarios. Ultimately, the portable optophone system holds promise in enabling visually impaired individuals to navigate their environment more confidently and independently.

REFERENCES

[1] Hsy J. Close listening: Talking books, blind readers and medieval worldbuilding. J. postmedieval: a journal of medieval cultural studies, 2016, 7 (2), p 181-192

[2] Velázquez R., Hernández H., Preza E. A portable eBook reader for the blind. C. Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE,2010, p 2107-2110

[3] Sabab S A., Ashmafee M H.: Blind Reader: An intelligent assistant for blind. C. International Conference on Computer & Information Technology. IEEE, 2017, p 229-234.

[4] Braga DDS., Neto FBDL., Oliveira SC.: An Intelligent Human Machine Interface for Reading Digital Texts for The Blind. C. Brazilian Congress of Computational Intelligence. 2016, p 1-7.

[5] Elias M., Eltner A., Liebold F. et al.: Assessing the Influence of Temperature Changes on the Geometric Stability of Smartphone- and Raspberry Pi Cameras. J. Sensors 2020, 20(3), p 643

[6] Picton P.D.: Real-time implementation of an optophone. C Colloquium on Computers in the Service of Mankind: Helping the Disabled, IEEE,1997, p 6

[7] Bissacco A., Cummins M., Netzer Y., et al.: PhotoOCR: Reading Text in Uncontrolled Conditions. C 2013 IEEE International Conference on Computer Vision (ICCV). IEEE, 2013, p 785-792.

[8] Lecun Y., Bengio Y., Hinton G.: Deep Learning. J. Nature, 2015, 521(7553):436.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379

doi crossref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details