Wearable assistive Device : OCR-Guided Wearable Solution for Visual Impairment Support

Abstract*:Advances in technology have led to innovative solutions to help the blind. Among these solutions, Braille stands out as a traditional method for tactical reading. However, the costs associated with Braille printing are a challenge. To solve this problem, a wearable assistive device for the blind is proposed that converts text into acoustic output, allowing users to read any type of text. The device uses a self-contained Raspberry Pi system and a finger-pointing camera to assist speech-impaired people by capturing finger-pointing text data. The system includes a light-dependent resistor (LDR) sensor to detect ambient light conditions and adjust the brightness. The webcam captures enhanced images and a new methodology is used to extract the words indicated by the finger. The extracted text is then processed by an optical character recognition (OCR) engine. Finally, the text output is converted to speech using a Text to Speech (TTS) converter that can be accessed through a sound output device such as headphones or speakers. This Raspberry Pi-enabled assistive device offers a low-cost and affordable solution to improve the reading experience for the visually impaired and ultimately improve access to textual information.

Keyword:*blind; blind reader; visually impaired; smart reader ; word based text extraction; text to speech; assistive wearable device;*

I. INTRODUCTION

Visually impaired individuals often face challenges in accessing reading resources independently, especially in public spaces such as shops, restaurants and hospitals. Traditionally, efforts have focused on generating Braille equivalents for printed text. However, early implementations of Braille-based technologies were cumbersome, time-consuming, and expensive. As a result, alternative approaches to providing various forms of feedback for visually impaired readers have emerged.

One notable advance came in the form of the Digital Accessible Information System (Daisy) standard, which combined text and audio components in accessible books, improving the reading experience for the visually impaired. Although audio versions of books were previously available on cassettes or tapes, they presented challenges in terms of navigation and proper placement of audio content.

The advent of optical character recognition (OCR) technology brought about a significant innovation in this area. OCR enables the extraction of text from images, paving the way for mobile applications such as KNFB Reader and Blindsights Text Detective. These apps use OCR and speech synthesis to capture and read text from images, making them more accessible to the visually impaired. However, they still require careful alignment, lighting and focus for optimal results and often lack the ability to provide detailed text information.

Specialized devices such as Eye-Pal, OrCam, iCare and Finger Reader have been developed to address the challenges faced by visually impaired individuals in reading and analyzing text. These devices use cameras or headsets to recognize objects and read printed text in real time, offering auditory or tactile feedback. Despite their functionality, these devices tend to be bulky, lack real-time effects, and may require additional external components or complex hardware setups.

There is a growing need for computationally inexpensive algorithms and cost-effective methodologies to overcome the limitations of existing technologies. This paper aims to present a simple algorithm that offers satisfactory results and provides a more accessible and affordable solution for visually impaired individuals to access reading resources independently. By addressing these challenges, the goal is to improve the overall accessibility and independence of visually impaired people in their reading efforts.

II. RELATED WORKS

This project proposes an efficient and portable camera-based text-to-speech converter for blind individuals, using a convolutional recurrent neural network to train words independently. Two possible text extraction methods are proposed: motion or direct identification of text parts. The solution uses Python libraries such as Tesseract and Sphinx for speech processing and OpenCV for image processing using OCR software. Detection and recognition of individual characters is performed using DCNN models. The system compares the text in real time with the dataset using a convolutional recurrent neural network. A combination of OCR and CRNN is used to recognize the text present in any object. The image is captured, pre-processed with OCR and passed through the CRNN to train the text part. A rectangular bounding box is applied using the sorted outline method.

The OCR Based Facilitator for the Visually Challenged is a camera-based system that helps visually challenged individuals read printed text using a portable and affordable device. The system is built on the Raspberry Pi and integrates image processing algorithms, OCR, and a text-to-speech synthesis module. The OCR process includes binarization, de-noising, deskewing, segmentation, and feature extraction to convert the image into machine-readable text that is processed by the device. Google Tesseract is the open-source OCR engine used in this project, and Pico is the lightweight TTS engine used to generate speech in multiple languages. The Raspberry Pi provides the necessary processing power and storage for the device, which has a simple interface that allows the user to capture an image of the text and convert it into speech. The system can also save the machine-readable text for later use or translate it into different languages. The OCR Based Facilitator for the Visually Challenged is an innovative and customizable solution that can make a significant impact on the lives of visually challenged individuals by providing them with the ability to read printed text independently.

Implementing a word-based text extraction algorithm in a wearable assistive device for the blind is a cost-effective and portable solution for blind people to read text. The algorithm uses a finger-worn camera to acquire focused images at a fixed distance, allowing the user to use touch to scan the surface of the document. The captured image is reformatted to 800x600 and goes through image enhancement procedures such as RGB format, histogram equalization, brightness and contrast adjustment, and gamma correction to produce a uniform size image for further processing. The algorithm is implemented on a Raspberry Pi 3 Model A+ and takes about 5 seconds to process and deliver the audio output to the headphones. The same algorithm is also implemented on a virtual machine provided by Azure Cloud, which takes about 0.5 seconds to process and deliver the audio output. The approximate accuracy of word-to-audio conversion is 75% for a black and white document, and special characters obtained in a text file after image-to-text conversion are removed to avoid TTS conversion errors. The device is portable, computationally efficient and affordable compared to existing technologies.

them as disabled. Haptic technology in BLIND READER provides an additional sense to the BVI people to help them interact with the reading material in a more comprehensive way. This technology allows the user to feel the layout of the text, such as the start or end of a line, new lines, or other text features. It also alerts the user if they move away from the baseline, helping them maintain a straight scanning motion. By providing haptic feedback, BLIND READER improves the user's reading experience and makes it more efficient.

This product appears to be a text reader built on a Raspberry Pi module with a camera used to capture input images. The captured image is enhanced using image processing techniques and the Tesseract OCR engine searches for text in the image and converts it into a digital document. The digital document is then analyzed using a semantic review engine and the text is converted to speech using a Python-based TTS conversion unit built into the Raspberry Pi. Finally, the audio output is passed to the audio amplifier to be loaded. The input image is captured using a web camera with a GUI and passed to OCR for text conversion. The eSpeak Speech synthesizer is used to convert text to speech. OCR converts the text present in the image into binary coded text, which is later provided for semantic review.

III. PROPOSED METHODOLOGY

OCR based facilitator for the visually challenged

This project is a camera-based framework built on the Raspberry Pi and integrated with image processing algorithms, OCR, and a text-to-speech (TTS) module. The main goal of this framework is to capture an image of text and convert it into audible speech.

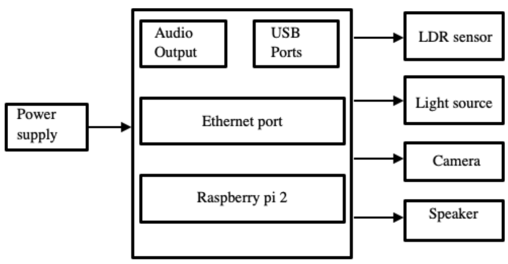

Fig. 1 System hardware design

The preprocessing phase of this framework includes several steps such as binarization, denoising, slope equalization, segmentation, and feature extraction. Binarization is the process of converting an input image into a binary image where pixels are either black or white. Denoising is done to remove all the noise present in the image. Skew correction is used to correct the skew of the text in the image. Segmentation is the process of dividing an image into small areas for further processing. Feature extraction is used to extract important features from segmented regions.

For OCR, this project uses Google Tesseract, which is an open-source OCR engine that can recognize text in different languages. Tesseract is used to search for text in an image and convert it into a digital document. OCR converts the text present in the image into binary coded text, which is later provided for semantic review.

For TTS, this project uses Pico, a text-to-speech engine developed by the Raspberry Pi Foundation. Pico is used to convert extracted text into audible speech. The audio output is then passed to the audio amplifier to be loaded.

Overall, this camera-based framework is a useful tool for converting printed text into audible speech, making it accessible to people with visual impairments or those who prefer to listen rather than read.

The system consists of a LDR sensor which detect the ambient light condition and turns on the light accordingly, webcam that captures images which are enhanced. Following this the word pointed by the finger is extracted using a novel methodology and given to an Optical Character Recognition (OCR) engine. Subsequently, the textual output is given to a Text to Speech (TTS) converter to obtain audio via an audio output device such as earphones or speaker

IV. SYSTEM OVERVIEW

The wearable device consists of a camera module and headphones/speakers connected to a Raspberry Pi. The camera module captures an image of the text to be read and sends it to the Raspberry Pi. The Raspberry Pi then uses optical character recognition (OCR) software to convert the image to text and a text-to-speech (TTS) engine to read the text aloud through headphones/speakers.

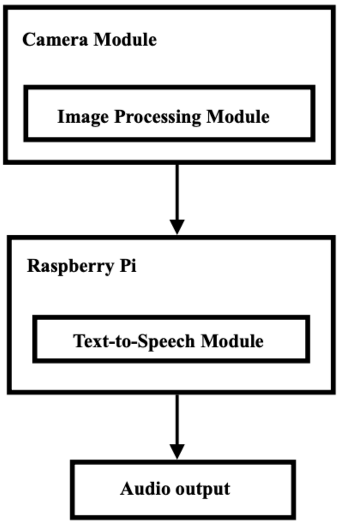

Fig. 2: Basic Structure of the Device

The system consists of a LDR sensor which detect the ambient light condition and turns on the light accordingly, webcam that captures images which are enhanced. Following this the word pointed by the finger is extracted using a novel methodology and given to an Optical Character Recognition (OCR) engine. Subsequently, the textual output is given to a Text to Speech (TTS) converter to obtain audio via an audio output device such as earphones or speaker .

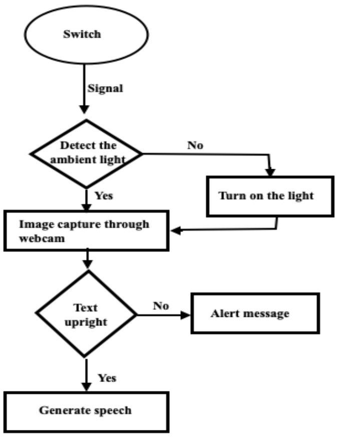

Fig. 3 Architecture of the Device

V. SYSTEM FUNCTIONALITIES

A. HARDWARE INTERFACES

All of these interface options are great for a wearable Raspberry Pi device for the blind. The Voice Command Interface would be very useful for those who have difficulty using buttons or touchpads. The Button Interface would be very helpful for those who prefer tactile buttons that they can feel and control. The Touchpad Interface could work well for those who are comfortable using a touchpad, as it is similar to using a laptop touchpad.

The Smartphone App Interface would be a good option for those who prefer to use their smartphones for control and those who already have experience using smartphones. The Braille Interface would be very helpful for those who are proficient in braille and prefer to read the text directly. Lastly, the Gestures Interface could be a good option for those who prefer to use hand gestures for control, as it is an intuitive way of interacting with devices.

In general, providing multiple interface options would be ideal for accommodating different users' needs and preferences. It is also important to ensure that the interface is accessible, user-friendly, and easy to use, especially for people who are blind or visually impaired.

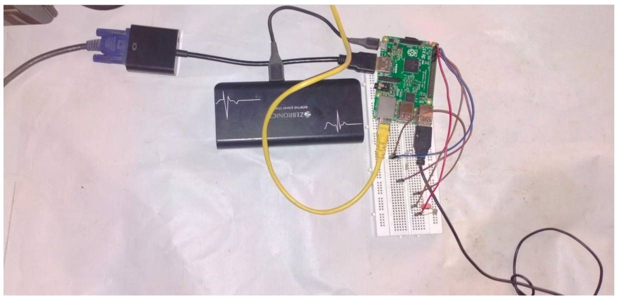

Fig. 4 Model setup

The wearable device consists of a camera module and headphones/speakers connected to a Raspberry Pi. The camera module captures an image of the text to be read and sends it to the Raspberry Pi. The Raspberry Pi then uses optical character recognition (OCR) software to convert the image to text and a text-to-speech (TTS) engine to read the text aloud through headphones/speakers.

B. SYSTEM REQUIERMENT SPECIFICATION

The system requirements specification given above outlines the hardware, software, performance, usability and security requirements required for the development of a wearable assistive device for the blind. The device aims to improve the quality of life for visually impaired individuals by providing a reliable and easy-to-use tool to navigate their surroundings. The device will use a Raspberry Pi microcontroller as its primary computing platform, and will have an array of sensors that will be worn by the user. These sensors will include ultrasonic sensors for obstacle detection and cameras for object recognition. Data from these sensors will be processed using machine learning algorithms to determine the user's location and orientation in their environment. The device will be designed with accessibility in mind, and will include large buttons and a simple user interface for easy use. Customizable features such as audio output volume and pitch and sensor sensitivity will also be included to cater to individual user preferences. The device will be powered by a rechargeable battery and designed to be easily recharged using a standard USB charging cable. Security and privacy measures will also be implemented to protect user data and prevent unauthorized access and use. Overall, the system requirements specification aims to provide a broad outline of the requirements needed to develop a wearable assistive device for the blind that is reliable and easy to use.

C. SOFTWARE INTERFACES

A wearable Raspberry Pi device designed to help visually impaired people read text requires several software interfaces to function efficiently. First, a text-to-speech (TTS) interface is required to convert the captured text into speech. This can be achieved by including software libraries such as eSpeak, Festival or the Google Text-to-Speech API. Second, optical character recognition (OCR) interfaces are important for recognizing printed text in images and converting them to editable text. This can be accomplished using software libraries such as Tesseract OCR, OpenCV or the Google Cloud Vision API.

Thirdly, a user interface is required to control the device, adjust settings and view information. This can be achieved through a graphical user interface (GUI) or a command-line interface (CLI) that can be controlled by voice commands or physical inputs. Fourth, an operating system interface is required to manage the hardware and software components of the device, and this can be achieved using the Raspberry Pi's default operating system, Raspbian, or another Linux-based operating system such as Ubuntu or Debian.

Fifth, accessibility interfaces may be needed to connect to other assistive devices, such as braille displays or switches. This can be achieved through software libraries or APIs that provide support for accessibility devices. Sixth, the device may require a wireless connectivity interface to connect to the Internet or other devices. This can be achieved using software interfaces such as Wi-Fi, Bluetooth or cellular connectivity.

Finally, the user needs an audio interface to output TTS audio, which can be achieved using software libraries or APIs that provide support for audio output, such as ALSA or PulseAudio. In summary, incorporating all these software interfaces into a Raspberry Pi wearable device for the blind can significantly improve the user experience and functionality of the device.

D. SYSTEM INTERACTION

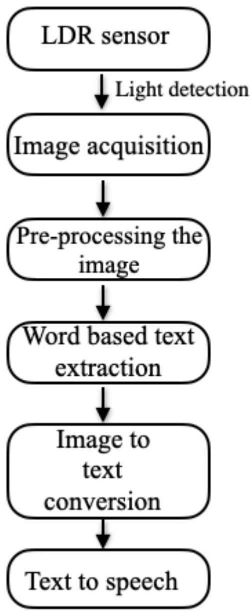

Fig. 5 Preprocessing Flow Chart

A wearable Raspberry Pi device designed to read text to the visually impaired may require several communication interfaces for various purposes, including data transfer, wireless connectivity, and audio output. Some of the possible communication interfaces that can be used in such a device include USB port for connecting external devices like keyboard, mouse or USB storage drive, CSI camera port for camera module to capture images of text for OCR, HDMI. port for connecting to external displays for visual output, built-in Wi-Fi connectivity for wireless Internet access and to connect to other devices on a local network, built-in Bluetooth connectivity for wireless communication with smartphones, headphones or other devices Braille display , 3.5mm audio jack for audio output to speakers or headphones, GPIO pins for connecting external hardware components such as sensors or switches, and cellular connectivity for remote data access and communication.

The selection of the appropriate communication interface for the device should take into account the specific requirements of the device and ensure that it is compatible with the hardware and software components. Additionally, when selecting and implementing a communication interface, it is important to consider accessibility requirements and user needs and preferences to ensure that the device is user-friendly and practical. Therefore, careful consideration should be given to ensure that the communication interface is properly selected and integrated into the device, providing the functionality and accessibility required by the visually impaired user

VI. TOOLS AND TECHNOLOGIES

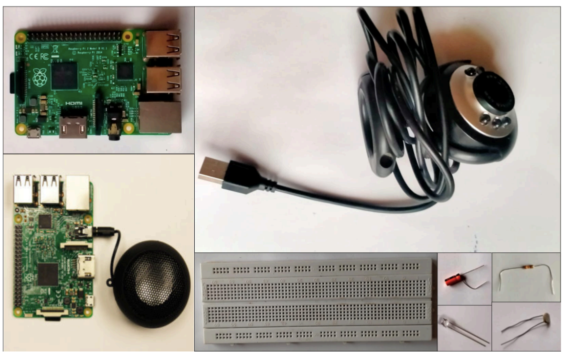

Fig. 6 hardware components used

A wearable Raspberry Pi device for the blind to read text requires a combination of hardware and software components to function effectively. The tools and techniques that can be used in the development of such a device are numerous, and the selection of the appropriate tools and techniques depends on the specific requirements of the device.

One of the most important components of the device is the Raspberry Pi itself, which provides the necessary computing power and interfaces to other hardware components. Additionally, OCR requires a camera module to capture images of the text.

The device also requires text-to-speech software libraries, which can convert valid text to speech output. Popular text-to-speech software libraries include eSpeak, Festival, or the Google Text-to-Speech API.

OCR software libraries are also important in recognizing printed text in images and converting them into editable text. Popular OCR software libraries include Tesseract OCR, OpenCV or Google Cloud Vision API.

To create an intuitive user interface to control the device, programming languages such as Python or C++ and GUI frameworks such as Qt, GTK or Tkinter can be used. Linux-based operating systems such as Raspbian, Ubuntu or Debian provide a stable and flexible platform for the device's software components.

Accessibility software libraries such as Assistive Technology Service Provider Interface (AT-SPI) or GOK provide support for accessibility devices such as braille displays or switches. Wireless connectivity via Wi-Fi, Bluetooth or cellular connectivity is also required for wireless communication with other devices or internet connectivity.

Finally, audio output interface support is required for devices such as ALSA or PulseAudio, which provide support for audio output to a 3.5mm audio jack or Bluetooth speaker. The device also needs a rechargeable battery or a power bank that can power the device for a long time.

Using the right tools and technology, it is possible to develop an effective wearable Raspberry Pi device for the blind that meets the specific needs of the user.

VII.REFERENCE

[1] D. Keating, Assistive technology for visually impaired and blind people, M. A. Hersh and M. A. Johnson, Eds. Berlin: Springer London, 2008.

[2] "Visual impairment and blindness 2010," 2013. [Online]. Available: http://www.who.int/blindness/data_maps/VIFACTSHEETGLODAT201 0full.pdf. Accessed: Aug. 30, 2016.

[3] J. A. M and K. Omar, “Quranic Braille System,” International Journal of Humanities and Social Sciences, vol. 3, no. 4, pp. 313–319, 2009.

[4] P. G. Anuradha, "A Refreshable Braille Display for the Interaction with Deafblind People," in Middle-East Journal of Scientific Research, IDOSI Publications, 2016, pp. 96–100.

[5] C. Asakawa, H. Takagi, S. Ino, and T. Ifukube, "Auditory and tactile interfaces for representing the visual effects on the web," Proceedings of the fifth international ACM conference on Assistive technologies - Assets ’02, 2002.

[6] P. Parente, "Clique," ACM SIGACCESS Accessibility and Computing, no. 84, pp. 34–37, Jan. 2006.

[7] R. Shilkrot, J. Huber, W. Meng Ee, P. Maes, and S. C. Nanayakkara, "FingerReader," Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15, 2015.

[8] X. Zhang, C. Ortega-Sanchez, and I. Murray, "A hardware based Braille note taker," 2007 3rd Southern Conference on Programmable Logic, Feb. 2007.

[9] R. Sarkar, S. Das, and S. Roy, "SPARSHA: A low cost Refreshable Braille for deaf-blind people for communication with deaf-blind and non- disabled persons," in Distributed Computing and Internet Technology. Springer Science + Business Media, 2013, pp. 465–475.

[10] P.Blenkhorn,G.Evans,A.King,S.HastutiKurniawan,andA.Sutcliffe, "Screen magnifiers: Evolution and evaluation," IEEE Computer Graphics and Applications, vol. 23, no. 5, pp. 54–61, Sep. 2003.

[11] Y.N.El-Glaly,F.Quek,T.L.Smith-Jackson,andG.Dhillon,"Itisnota talking book;," Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility - ASSETS ’12, 2012.

[12] P. Biswas and P. Robinson, "Evaluating interface layout for visually impaired and mobility-impaired users through simulation," Universal Access in the Information Society, vol. 12, no. 1, pp. 55–72, Dec. 2011.

[13]F.Alonso,J.L.Fuertes,Á.L.González,andL.Martínez,"User-interface Modelling for blind users," in Computers Helping People with Special Needs. Springer Science + Business Media, pp. 789–796.

[14] HART, G. and STAVELAND, L.E., "Development of a Multidimensional Workload Rating Scale: Results of Empirical and