DESIGN AND DEVELOPMENT OF WEARABLE TEXT READING DEVICE FOR VISUALLY IMPAIRED PEOPLE

MWP Maduranga
Department of Computer Engineering,
General Sir John Kotelawala Defence University,
Ratmalana, Sri Lanka
pasanwellalage@kdu.ac.lk

NT Jayatilake
Department of Computer Engineering,
General Sir John Kotelawala Defence University,
Ratmalana, Sri Lanka
jayatikalan@kdu.ac.lk

PADV Pannala
The Centre for Open and Distance Learning,
University of Moratuwa, Katubedda
Sri Lanka
dvenuran@gmail.com

SPARS Jayathilake
Department of Information Technology
General Sir John Kotelawala Defence University,
Ratmalana, Sri Lanka
ashenw@kadu.ac.lk

***Abstract*— For the visually impaired, objects such as texts and symbols are difficult to be recognized and understood without additional help. Though the Braille method has been the solution for this for a long time, there are numerous instances and applications where it has proven to be slow or impossible since it incorporates the person's sense of touch. This project to design and develop a wearable text reader for visually impaired people is suggested as a solution to this problem. The system is intended to read any text using a camera-based Raspberry pi device to provide an audio output for the visually impaired person. The system, which will consist of mainly two parts, will first identify the text in a particular image using an OCR engine to provide an audio output of the text finally. The Optical Character Recognition (OCR) algorithms and the image processing function of the system were given priority in this project. To derive a suitable algorithm, many text recognition algorithms were analyzed and implemented on the python IDE against different kinds of images such as invoices, digital images, and handwritten characters. Pre-trained modelswere then used with these algorithms to increase the system's efficiency. Some image pre-processing techniques were adopted to improve the accuracy of the OCR. An image data set of 62992 images with 128x128px resolution containing characters, numbers, and symbols with four kinds of font styles was created to train a unique model for the OCR. The data set was trained using the Tensorflow machine learning framework, and technologies such as deep learning and image processing were mainly incorporated in the development of thesystem. The focus of this project is character recognition and conversion. The accuracy of the OCR will be effectively evaluated using proper and accepted testing methods. The Tesseract algorithm was tested at 94.93% accuracy, and the results are presented in this paper.**

Keywords- Machine Learning, Deep Learning, Image processing, Wearable devices, Optical Character Recognition

I. INTRODUCTION

Reading with visual impairment can significantly impact the quality of life of the visually impaired. Though an approximation of 135 million of the world's population is visually impaired, the consideration of the general society towards these people has been low. It has been brought forth how visually impaired people have different problems when accessing printed text and how reading with visual impairment has a significant impact on the quality of life of the visually impaired. In the literature, many electronic solutions have been developed[1-3]

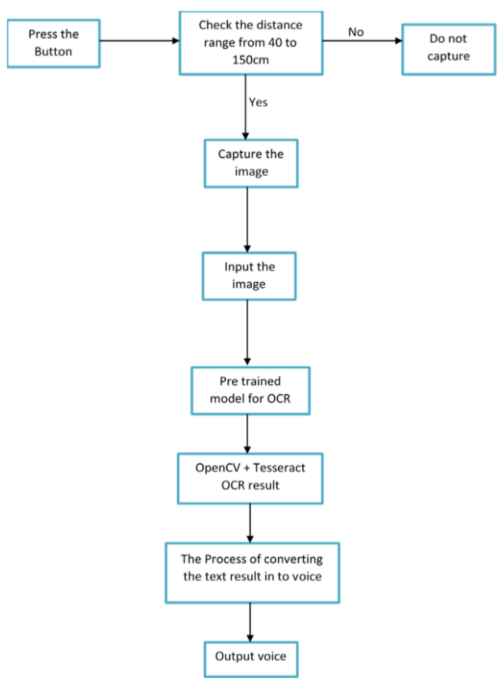
Despite the brail system being a perfect solution to support the visually impaired, the increasing amount of printed material and the need to read information and communicate through screens and other sources has posed a problem for them in the modern era. The busy day-to-day hassle in the outside society also demands efficiency when doing any task, and here too, using brail's method can be a problem due to the time it takes [4-5]. Even though the development of text-reading devices is still in its infancy, various devices have been designed to assist these people in viewing objects with alternate senses, such as sound and touch. But the inconvenience of having to use an additional device or the inability to use such a device in certain instances has posed the need for a portable device specific for the purpose. Wearable devices can be considered a perfect and suitable solution to remedy this problem.

Among different kinds of existing text reading wearable devices already developed globally, the finger reader and the OrCam are the most effective and famous devices. The finger reader is a wearable device capable of reading severalwords and whole lines of texts[6-8]. It is also associated witha laptop or smartphone. But the finger reader's inability to read text smaller than 12-point typeface and its incapability to read handwritten characters can be seen as significant drawbacks of this prototype-level device. The OrCam, one ofthe best wearable text reader devices for the visually impaired, uses a camera attached to a pair of spectacles and a small computer to convert text to audio. But despite its ease to use and efficiency, the high cost of the device, its tendency to scrap at different light levels, and its inability to read handwritten characters have proven to be the OrCam's disadvantages.

After careful consideration and intense research on these systems, their drawbacks, and the needs of the visually impaired, a system was designed to meet the requirements of these people. Hence, this project suggests an accurate and efficient, low-cost wearable IoT device that can automatically find and read visually impaired people's printed or handwritten text. The project uses a camera-based Raspberry Pi device to read printed texts. Using Optical Character Recognition, the captured image will be converted to text output. The blind user can listen to the text of that image by using the Text to Speech method. A newly trained data model will be used for Optical Character Recognition and a few image processing techniques to increase the text recognition accuracy. The text-to-speech method will convert the text into clear sound. The final product will be a camera-based small-scale wearable device that operates automatically by pressing a button.

II. DESIGN AND IMPLEMENTATION

The process aimed to develop a user-friendly, accurate, and efficient, low-cost text reading wearable device for the visually impaired. The device's functions were decided such that the device identifies any text image within the range of 40cm to 150cm when a button is pressed. The image will then be processed and converted to text form for the text-to- voice



output, as shown below in "Fig.1".

Fig. 1. Process of OCR

The existence of OCR systems such as Google Docs OCR, Tesseract, ABBYY FineReader, Transym, OCRFeeder, etc. and the ability to use them with different kinds of text such as Handwritten characters, multi-Oriented text strings, Multilingual text images, Machine written characters, Characters in noisy images, Characters in digital images, etc. provided a great base of knowledge and resource for OCR analysis and the improvement the system. To improve the accuracy of the OCR, a dataset was trained using Tensorflow and Keras machine learning frameworks. The dataset consisted of 62992 black and white images of 128x128px resolution that included numbers from (0-9), capital, and simple letters from (A-Z) [9-11].

Python IDE with the OCR systems focusing on the Tessaract OCR algorithm (Fig. 2 below) and OpenCV OCR algorithms. [12-16]

Initialization
1. Read image path from command line
2. Define config parameters using LSTM OCR engine.
3. Read image from the disk
4. Run Teseract OCR on the image
5. Print recognized text

End

The device's hardware design was done using a Raspberry pi 4 model B microprocessor with 1GB RAM, along with a Raspberry pi camera (module V2.1) with 8MP camera quality for the images, an HC-SR04 Ultrasonic Range Sensor to identify distance, and additional components such as a Transcend Micro SD 32GB Chip, a5V Buzzer, a Speaker, Flex pi camera cable and a Push- button as shown in fig.3. to finally develop a successful prototype (fig. 4). The Raspbian operating system, a Linux- based Operating system with a Graphical User Interface, wasused for the development. The primary programing languageof the device is Python. The Python IDE with OpenCV3, Tesseract, pip3, Python Imaging Library (PIL), NumPy, andText to speech (TTS) libraries were used to develop the result. The device's hardware design was done using a Raspberry pi 4 model B microprocessor with 1GB RAM, along with a Raspberry pi camera (module V2.1) with 8MP camera quality for the images, an HC-SR04 Ultrasonic Range Sensor to identify distance, and additional components such as a Transcend Micro SD 32GB Chip, a5V Buzzer, a Speaker, Flex pi camera cable and a Push- button as shown in fig.3. to finally develop a successful

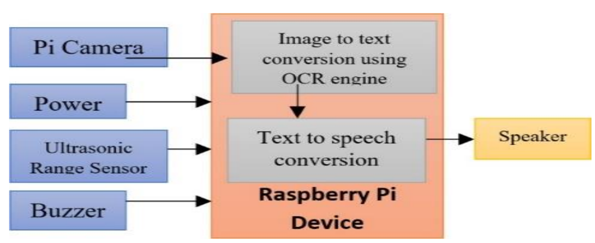


Fig. 3. The hardware Design

prototype (fig.4).

The Raspbian operating system, a Linux-based Operating system with a Graphical User Interface, was used for the development [17]. The primary programing language of the device is Python. The Python IDE with OpenCV3, Tesseract, pip3, Python Imaging Library (PIL), NumPy, and Text to speech (TTS) libraries were used to develop the result.

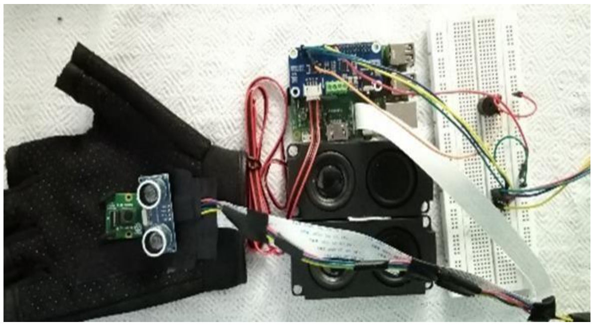


Fig. 4. A working prototype of the device

III. TESTING AND RESULTS

TABLE 1. COMPARISON OF TESSERACT ALGORITHM

Sample	Algorithm	Total Characters	Characters Extracted	Accuracy
01	Tesseract OCR Algorithm	54	45	83.333%
02	Tesseract algorithm after pre-processing	79	75	94.9367%

Evaluation and testing are the most critical and essential stages for the adequate performance of any device or software application. To ensure that the whole system is

Fig. 5. Test sample 1 and results

working as expected with the least possible deviations, pre-defined test cases were used in the testing process.

Among the few algorithms that were executed before training the model, the main algorithm chosen to be used was tested with a set of images to derive a comparison of the algorithm's accuracy.

After using the selected algorithm. The system was built step by step, and the function of the device was defined in three stages real-time image capturing, image-to-text conversion, and text-to-speech conversion to execute unit and performance tests with pre-defined test cases for each stage.

The system was finally integrated to do the integrated testing stage. (Fig.5 and Fig.6.) All the expected functions were tested on the complete system following black box testing techniques. The device was also tested for user acceptance. Different scenarios were tested. The final device worked without any error, met a minimum of problems, and was given a pass.

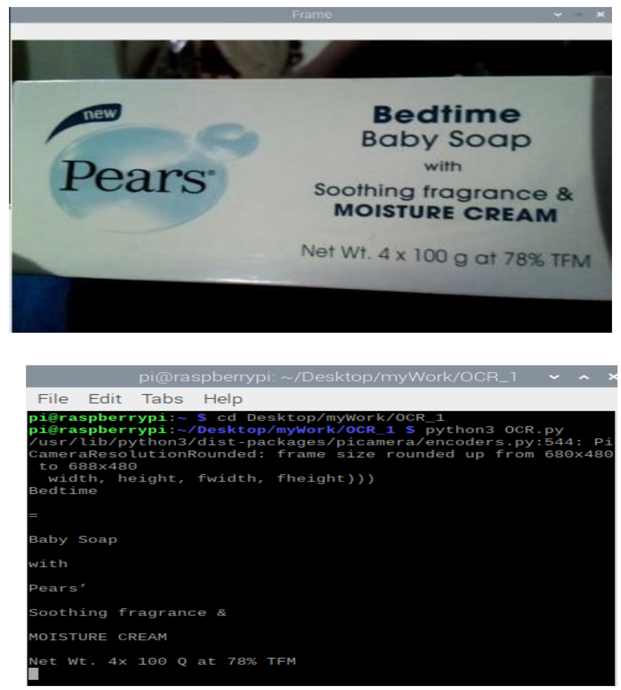


Fig. 6. Test sample 2 and results

IV. CONCLUSION

This device was designed to transform an image into the text to convert the text to audio then. These two functions were then combined with Optical Character Recognition techniques to build this beneficial solution for visually impaired people. As proven by the tests carried out, the pre-processing image technique permits extracting the essential text from a multifaceted background to provide quality input for OCR. The extracted text is then directed to the text-to-speech engine for the output audio, proven accurate. This paper describes the testing samples, testing results, and overall system accuracy. This system was successfully established on the OpenCV platform. The device mainly concentrates on helping the visually impaired handle their difficulty reading books or any script in the text form. It is intended to improve this system further to read, identify and read other languages. It can be improved in the future to recognize and predict even the objects and the various colors used in an image. But given the requirements for which the system was built and by observing the outcomes of the system, it can be concluded that the device was successfully designed. However, it is still at its prototype level and requires further improvement. Hence, the project results show that clusters are more efficient than single boards when the workload is high. This proves the hypothesis of this research project.

REFERENCES

[1] M. P. Agrawal and A. R. Gupta, "Smart Stick for the Blind and Visually Impaired People," 2018 Second International Conference on Inventive Communication and Computational Technologies(ICICCT), 2018, pp. 542-545, DOI: 10.1109/ICICCT.2018.8473344.

[2] B. Muthusenthil, J. Joshuva, S. Kishore, and K. Narendiran, "Smart assistance for blind people using Raspberry Pi," International Journal Of Advance Research, Ideas And Innovations In Technology, vol. 4,no. 2, pp. 884-891, 2018.

[3] A. Waghmare, B. Sawale, M. Pangare, U. Shinde, and A. Jadhav, "Raspberry Pi Based Reader for Blind People,", InternationalResearch Journal of Engineering and Technology (IRJET), vol. 6, no. 4, pp. 3040–3041, April 2019. Available: doi: 10.26769/IJIERE.2018.5.1.191540

[4] Govinda Dumane, "Image Augmentation using Python, Numpy, Opencv and Skimage", Towards data science, Dec. 24, 2019. Accessed on: Oct. 12, 2020. [Online] Available: https://towardsdatascience.com/image-augmentation-using-python-numpy-opencv-and-skimage-ef027e9898da.

[5] S. Karunaratna and P. Maduranga, "Artificial Intelligence on Single Board Computers: An Experiment on Sound Event Classification," 2021 5th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), 2021, pp. 1-5, doi: 10.1109/SLAAI-ICAI54477.2021.9664746.

[6] J. Köhler, T. Nöll, G. Reis and D. Stricker, "A full-spherical device for simultaneous geometry and reflectance acquisition," 2013 IEEE Workshop on Applications of Computer Vision (WACV), 2013, pp. 355-362, doi: 10.1109/WACV.2013.6475040.

[7] P. Romulus, Y. Maraden, P. D. Purnamasari and A. A. P. Ratna, "An analysis of optical character recognition implementation for ancient Batak characters using K-nearest neighbors principle," 2015 International Conference on Quality in Research (QiR), 2015, pp. 47-50, doi: 10.1109/QiR.2015.7374893.

[8] M. Brisinello, R. Grbić, D. Stefanovič and R. Pečkai-Kovač, "Optical Character Recognition on images with colorful background," 2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), 2018, pp. 1-6, doi: 10.1109/ICCE-Berlin.2018.8576202.

[9] T. C. Wei, U. U. Sheikh and A. A. -H. A. Rahman, "Improved optical character recognition with deep neural network," 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 2018, pp. 245-249, doi: 10.1109/CSPA.2018.8368720.

[10] D. Berchmans and S. S. Kumar, "Optical character recognition: An overview and an insight," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014, pp. 1361-1365, doi: 10.1109/ICCICCT.2014.6993174

[11] P. Mishra, P. Pai, M. Patel and R. Sonkusare, "Extraction of Information from Handwriting using Optical Character recognition and Neural Networks," 2020 4th International Conference onElectronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1328-1333, doi: 10.1109/ICECA49313.2020.9297418.

[12] P. Imsamer, V. Boonyaphon and S. Tiacharoen, "The Comparison of Deep Learning Driven Optical Character Recognition for Hard Disk Head Slider Serial Number," 2020 International Conference on Power, Energy and Innovations (ICPEI), 2020, pp. 217-220, doi: 10.1109/ICPEI49860.2020.9431431

[13] A. Chakraborty, "Image Processing and Image Pattern Recognition a Programming Tutorial," 2018 First International Conference on Artificial Intelligence for Industries (AI4I), 2018, pp. 122-123, doi: 10.1109/AI4I.2018.8665702.

[14] S. Tajrean and M. Abu Yousuf, "Handwritten Bengali Number Detection using Region Proposal Network," 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), 2019, pp. 1-6, doi: 10.1109/ICBSLP47725.2019.202049.

[15] Ç. Gider and S. V. Albayrak, "Identifing of Alphanumerical Codes in Promotional products by Using of Deep Neural Network," 2018 3rd International Conference on Computer Science and Engineering (UBMK), 2018, pp. 458-462, doi: 10.1109/UBMK.2018.8566404.

[16] A. M. Lokeshwar, A. Hardikar, G. C. Srivarsha, R. S. Priyanka and K. J. Bhanushree, "An App to Assist Differently Abled People," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021, pp. 1609- 1613, doi: 10.1109/ICAC3N53548.2021.9725637.

[17] H. Huda, M. Ariful Islam Fahad, M. Islam and A. K. Das, "Bangla Handwritten Character and Digit Recognition Using Deep Convolutional Neural Network on Augmented Dataset and Its Applications," 2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM), 2022, pp. 1-7, doi: 10.1109/IMCOM53663.2022.9721634.