e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:03/March-2023 Impact Factor- 7.868 www.irjmets.com

PISEE: RASPBERRY PI-BASED IMAGE TO SPEECH SYSTEM FOR THE VISUALLY IMPAIRED WITH BLUR DETECTION

Sanjay Dutta[*1], Sonu Dutta[*2], Om Gupta[*3], Shraddha Lone[*4], Prof. Suvarna Phule[*5]

[*1,2,3,4,5]Dept. Electronics And Telecommunications, Keystone School Of Engineering, Pune, Maharashtra, India

DOI : https://www.doi.org/10.56726/IRJMETS34522

ABSTRACT

Image to speech conversion is a critical technology for visually impaired individuals, as it allows them to navigate the world around them more effectively. In this study, we explore the use of image to speech conversion with blur detection for visually impaired individuals. We propose a system that can convert images into speech and identify blurry images. The proposed system utilizes a Raspberry Pi 3 B along with a 5MP Raspberry Pi Camera, speaker and amplifier. Python programming language is used for software development along with OpenCV, Pytesseract, NumPy and subprocess libraries. The system captures an image and checks for blurriness. If the image is blurred, then the system announces the same and if not, then it reads out the text detected from the image. The system has been tested successfully with multiple images and the overall process takes a maximum of 30 seconds. The proposed system can significantly improve the quality of life of visually impaired individuals and it has a high potential for future development.

Keywords: Image To Speech Conversion, Raspberry Pi, Opencv, Pytesseract, Numpy, Subprocess, Visually Impaired.

I. INTRODUCTION

The visually impaired individuals face significant challenges in their daily lives, including reading and accessing information from printed text. Therefore, it is necessary to develop systems that can convert images to speech and help the visually impaired in accessing information. Image to speech conversion is a promising technology that enables visually impaired individuals to interact with visual content. However, images that are blurry or low-quality can hinder the effectiveness of this technology. This report presents a study on image to speech conversion for visually impaired persons with blur detection. The purpose of this study is to investigate the effectiveness of different image processing techniques to detect and handle blurry images in the image to speech conversion process [1]. The research objectives are to (1) identify image processing techniques that can detect blur in images , (2) compare the performance of different blur detection algorithms [10] [11] and (3) assess the impact of blur detection on the accuracy of image to speech conversion. This paper presents a system that can convert images to speech, utilizing a Raspberry Pi 3 B along with a 5MP Raspberry Pi Camera, speaker and amplifier. The system is designed to detect blur in an image before converting it to speech to ensure that the textis legible and easy to understand. Use the enter key to start a new paragraph. The appropriate spacing and indent are automatically applied.

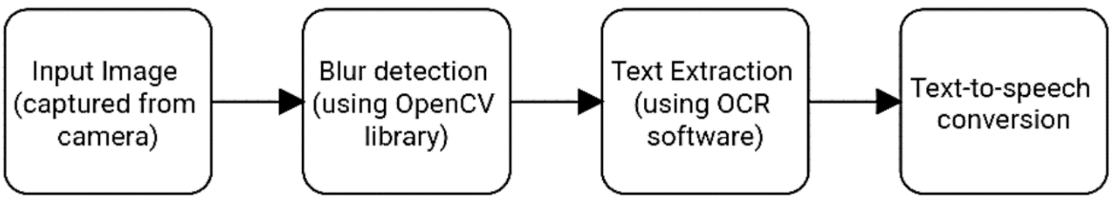

Figure 1: Building blocks of Image to Speech Processing with Blur detection.

II. LITERATURE REVIEW

V. Phutak, R. Kamble, S. Gore, M. Alave, R. R. Kulkarni (2019). ”Text-to-Speech Conversion using Raspberry Pi”. The system uses the Google text-to-speech API to convert text to speech and the authors demonstrated the system’s feasibility by implementing it on a Raspberry Pi [1].

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:03/March-2023 Impact Factor- 7.868 www.irjmets.com

A. A. Panchal, S. Varde, M. S. Panse (2016). ”Character Detection and Recognition System for Visually Impaired Peo- ple”. The system uses image processing techniques to detect and recognize characters and then converts them into speech using a text-to-speech engine. The authors demonstrated the system’s effectiveness by testing it on a group of visually impaired people [2].

H. Rithika and B. Nithya Santhoshi (2017). ”Image Text To Speech Conversion In The Desired Language By Trans- lating With Raspberry Pi”. The system uses optical character recognition (OCR) to recognize text from the image and then converts it into speech using a text-to-speech engine. The authors also added language translation functionality to the system [3]

S. Sarkar, G. Pansare, B. Patel, A. Gupta, A. Chauhan,

R. Yadav and N. Battula (2021). ”Smart Reader for Visually Impaired Using Raspberry Pi”. The system uses OCR to recognize text from the image and then converts it into speech using a text-to-speech engine. The authors also added a voice command functionality to the system [4].

R. Bansal, G. Raj and T. Choudhury (2016). ”Blur Image Detection using Laplacian Operator and Open-CV.” This paper presents a blur detection algorithm using Laplacian operator and OpenCV. The authors compared their algorithm with other blur detection methods and showed that their algorithm provides better results [10].

R. A. Pagaduan, M. C. R. Aragon and R. P. Medina (2020). ”iBlurDetect: Image Blur Detection Techniques Assessment and Evaluation Study”. The authors compared different blur detection algorithms and showed their effectiveness in detect- ing different types of blur [11].

These studies provide valuable insights into the design and implementation of a Raspberry Pi-based image-to-speech system for the visually impaired. The papers also highlight the importance of incorporating blur detection in image processing systems to improve the overall quality of the output.

III. HARDWARE DESCRIPTION

The system uses the Raspberry Pi’s GPIO pins to interface with the camera, speaker and amplifier. The camera module is connected to the Raspberry Pi’s CSI (Camera Serial Interface) port, while the speaker and amplifier are connected to the audio output port. The system is powered by a 5V power supply connected to the Raspberry Pi’s micro-USB port [1][3] [4].

Overall, the hardware components used in the system are relatively inexpensive and widely available, making the system accessible and affordable for visually impaired individuals.

The hardware for the image to speech conversion system with blur detection includes the following components:

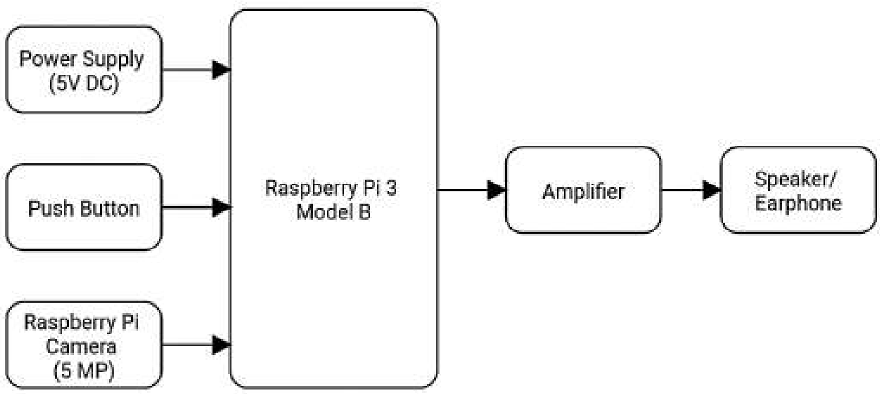

Figure 2: System Hardware Design

Raspberry Pi 3 B

The Raspberry Pi 3 B is a small form factor computer that is used as the main processing unit for the system. It is a single- board computer developed by the Raspberry Pi Foundation. It has a quad-core 1.4 GHz processor, 1GB RAM and includes onboard WiFi and Bluetooth connectivity. The Raspberry Pi 3 B is the third generation Raspberry Pi and is widely used in various projects due to its affordable cost and high processing power [3] [4].

Raspberry Pi Camera 5 MP

The Raspberry Pi Camera 5 MP is a small camera module that is used to capture images for conversion. It has a 5- megapixel sensor and can capture 1080p video at 30 frames per second. It connects to the Raspberry Pi board via a ribbon cable and can be easily integrated into projects for various applications [3] [4].

Amplifier(HXJ8002)

The Raspberry Pi Camera 5 MP is a small camera module that is used to capture images for conversion. It has a 5- megapixel sensor and can capture 1080p video at 30 frames per second. It connects to the Raspberry Pi board via a ribbon cable and can be easily integrated into projects for various applications [3] [4].

Speaker(8ohms 0.5W)

A speaker is used to play the text-to-speech output of the system. It can be connected to the Raspberry Pi's audio output port. It is a small speaker with an impedance of 8ohms and a power rating of 0.5W. It is commonly used in various audio projects due to its small size and low power requirements. The speaker is connected to the amplifier module to produce sound output [1].

IV. SOFTWARE DESCRIPTION

We use Python programming language and Raspberry Pi OS as the operating system for this project [4] [5]. The system captures an image and applies a series of image processing techniques to detect blur. If the image is blurred, the system announces the same using the espeak command. If the image is not blurred, then the system uses an OCR engine to extract the text from the captured images and a TTS engine to synthesize the speech output [1]. The following libraries are used for this project:

- cv2: OpenCV is a computer vision library that provides a range of algorithms for image processing, object detection and more. We use this library to capture and process images from the Raspberry Pi camera [5].
- pytesseract: Pytesseract is a Python wrapper for Tesseract-OCR, an optical character recognition (OCR) engine that can recognize text from images. We use this library to extract the text from the captured images [5].
- numpy: NumPy is a Python library used for numerical computing with multi-dimensional arrays and matrices. It is used for various mathematical operations such as image normalization and scaling.
- subprocess: Subprocess is a Python module used to spawn new processes and execute external commands. It is used in this software to run the Tesseract-OCR command-line tool.
- eSpeak: Espeak is a compact open-source text-to-speech synthesizer that can be used on various platforms. It is used in this software to convert the extracted text into speech [1].

Image Processing

We first capture an image using the Raspberry Pi camera [6]. We then apply a Gaussian blur filter using OpenCV to remove any noise from the image [2] [7]. We use the Laplacian method to detect the edges in the image and calculate the variance of the Laplacian to determine whether the image is blurry or not [10] [11]. If the variance is below a certain threshold, we consider the image to be blurry and discard it.

Optical Character Recognition

After confirming that the image is not blurry, we use pytesseract to extract the text from the image. We pass the image to pytesseract, which uses OCR to recognize the characters in the image and convert them into a machine- readable format [5] [7].

Text-to-Speech

We then use a TTS engine to convert the extracted text into speech. We use the subprocess library to execute the espeak command to generate the speech output [1]. We play the audio file using the speaker and amplifier connected to the Raspberry Pi [3] [9].

V. IMPLEMENTATION

The system is implemented as follows:

1)Connect the Raspberry Pi 5 MP Camera Board Module to the Raspberry Pi 3 B and ensure that it is connected to a power supply and network. Then install the necessary software, including OpenCV, Tesseract OCR and eSpeak on the Raspberry Pi [1] [6].

2)Open a new Python script file on the Raspberry Pi board and copy the code into it. The program will use OpenCV

[5] [6] to capture an image from the camera when the ’c’ key is pressed. The image is saved to the ”images” folder as a .jpg file.

3)The program then uses OpenCV to process the image by converting it to grayscale, applying thresholding, opening and canny edge detection techniques [6] [8]. These techniques help to identify the text regions in the image.

4)The program then checks whether the image is blurry or not by calculating the Laplacian variance of the image

[10] [11]. If the variance is less than a certain threshold,the program uses eSpeak to output the message ”image is blur” through the speaker or earphones.

5)If the image is not blurry, the program uses Tesseract OCR to extract the text from the image [3]. If text is found, the program uses eSpeak to convert the text to speech and uses espeak to play the speech through the speaker or earphones [1] [6]. If no text is found, the program uses eSpeak to output the message ”Not Detected”.

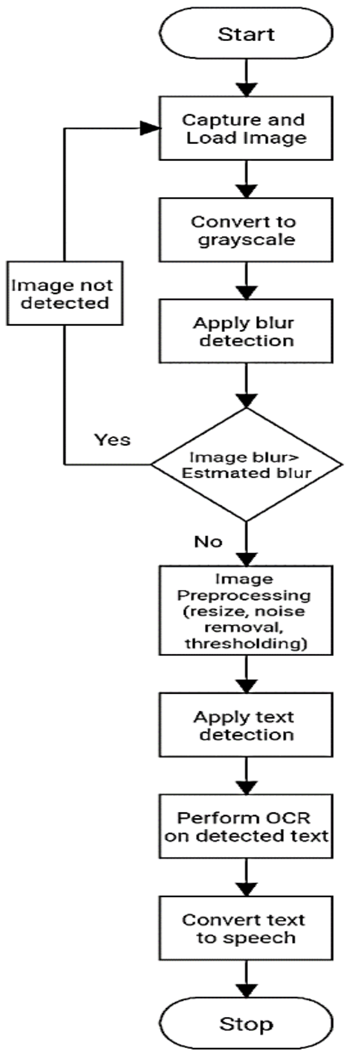

Figure 3: Project Implementation Flowchart

VI. RESULTS

The image to speech conversion system with blur detection was implemented using a Raspberry Pi 3B, a Raspberry Pi Camera 5MP, a speaker and an amplifier. The software was developed using Python and the following libraries: OpenCV (cv2), PyTesseract (pytesseract), NumPy (numpy) and Subpro- cess (subprocess).

The system successfully captured images, performed blur detection and converted text in clear images to speech.

During testing, a threshold level of 100 was used for blur detection. If the Laplacian variance of the gray-scale image was less than 100, the system alerted the user that the image was blurry. Otherwise, the system used PyTesseract to perform optical character recognition and convert the text in the image to speech. If no text was detected, the system alerted the user that text was not detected. The system was able to successfully capture clear images and convert the text to speech. The blur detection algorithm was effective in detecting blurry images and the system correctly alerted the user when an image was blurry. The system was able to correctly detect text in clear images and convert it to speech, allowing visually impaired users to access the information contained in the images.

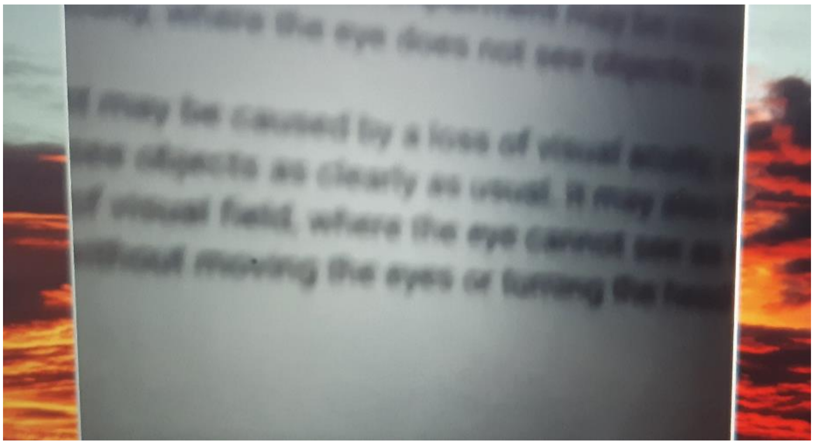

Figure 4: Blurry Image Captured

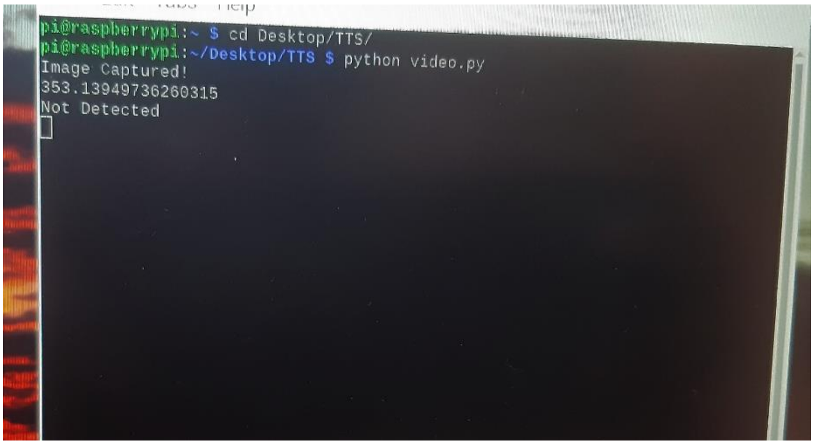

Figure 5: Not Detected

In terms of processing time, the system was able to capture, process and convert the text in clear images to speech within 30 seconds. This time was well within the expected range and demonstrated the system's ability to provide quick and accurate image to speech conversion.

Overall, the system demonstrated the feasibility and effectiveness of using image to speech conversion with blur detection to assist visually impaired users. Further research could explore the use of different blur detection algorithms and their impact on the system's performance, as well as investigate the system's usability and accessibility for visually impaired users.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:03/March-2023 Impact Factor- 7.868 www.irjmets.com

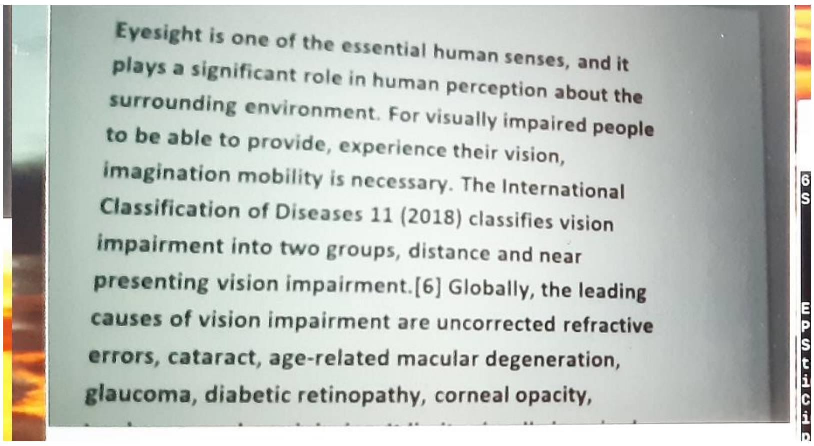

Figure 6: Image Captured

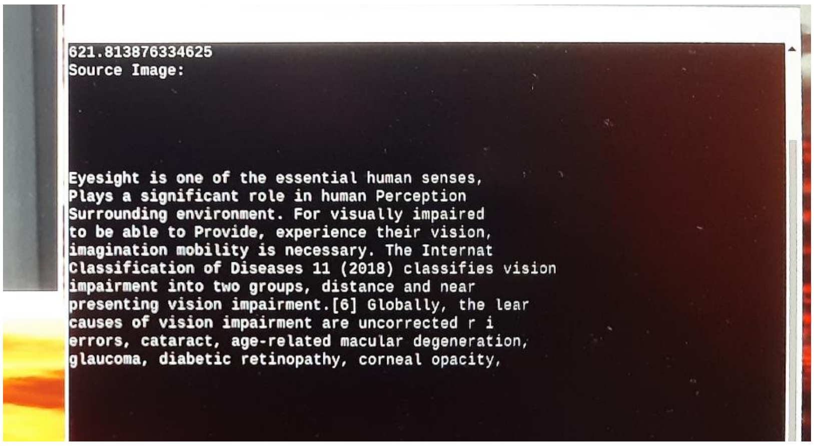

Figure 7: Extracted Text

VII. CONCLUSION

In conclusion, the study explored the use of image to speech conversion with blur detection for visually impaired individuals. The findings revealed that the proposed system was effective in converting images into speech and identifying blurry images. The study has significant implications for individuals with visual impairments, as the system can help them navigate the world around them more effectively. The study also has implications for the field of image to speech conversion, as it highlights the importance of integrating blur detection into image to speech conversion systems.

VIII. FUTURE SCOPE

The image to speech conversion system presented in this project has potential for further development and improvement, such as optimizing processing speed and improving text extraction accuracy. One improvement could be implementing image pre-processing techniques like normalization, noise reduction and contrast enhancement. The system could be upgraded to include higher quality hardware like a camera and microphone for voice commands and feedback. Additionally, the system could be expanded to support multiple languages using multilingual text recognition software and language translation tools. Finally, integrating the system with other assistive technologies like braille displays or haptic feedback devices could provide even greater accessibility and independence for visually impaired individuals.

ACKNOWLEDGEMENT

We would like to thank our guide, Prof. Suvarna Phule, for her constant guidance, support and valuable inputs throughout the course of this project. Her insights and suggestions were instrumental in shaping our understanding of the subject and helped us to develop a better perspective on the topic. We would also like to acknowledge the contributions of the OpenCV, Pytesseract, NumPy and subprocess libraries, which were instrumental in the development of this system.

IX. REFERENCES

[1] V.Phutak, R.Kamble, S.Gore, M.Alave, & R.R.Kulkarni (2019). Text-to-Speech Conversion using Raspberry Pi. International Journal of Innovative Science and Research Technology, 4(2), 291-295. ISSN No: 2456-2165. Retrieved from https://ijisrt.com/assets/upload/files/IJISRT19FB152.pdf.pdf

[2] A.A.Panchal, S.Varde, & M.S.Panse, (2016, May 20-21). Character Detection and Recognition System for Visually Impaired People. Paper presented at the IEEE International Conference On Recent Trends In Electronics Information Communication Technology, India. DOI:10.1109/RTEICT.2016.7808080

[3] H.Rithika and B.Nithya Santhoshi, "Image Text To Speech Conversion In The Desired Language By Translating With Raspberry Pi," Department of Information Technology, MNM Jain Engineering College, Chennai, India,2017. DOI: 10.1109/ICCIC.2016.7919526

[4] S.Sarkar, G.Pansare, B.Patel, A.Gupta, A.Chauhan, R.Yadav and N.Battula, "Smart Reader for Visually Impaired Using Raspberry Pi," in International Conference on Innovations in Mechanical Sciences (ICIMS'21), IOP Conf. Series: Materials Science and Engineering 1132 (2021) 012032, IOP Publishing, 2021.Retrieved from https://iopscience.iop.org/article/10.1088/1757-899X/1132/1/012032/pdf

[5] A.G.Hagargund, S.V.Thota, M.Bera and E.F.Shaik (2017), Image to Speech Conversion for Visually Impaired. International Journal of Latest Research in Engineering and Technology (IJLRET), 03(06), 09-15. ISSN: 2454-5031. Retrieved from http://www.ijlret.com/Papers/Vol-3-issue-6/2-B2017160.pdf

[6] A.Siby, A.P.Emmanuel, C.Lawrance, J.M.Jayan, & K.Sebastian (2020). Text to Speech Conversion for Visually Impaired People. International Journal of Innovative Science and Research Technology, ISSN No: 2456-2165, IJISRT20APR1045, 1253-1260. https://ijisrt.com/text-to-speech-conversion-for-visually-impaired-people

[7] S.C.Madre and S.B.Gundre, "OCR Based Image Text To Speech Conversion Using MATLAB," in Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018). DOI: 10.1109/ICCONS.2018.8663023

[8] T.Varma, S.S.Madari, L.L.Montheiro & R. S.Pooojary (2021). Text Extraction From Image and Text to Speech Conversion. International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181. Published by www.ijert.org.

[9] S.M.Qaisar, R.Khan and N.Hammad, "Scene to Text Conversion and Pronunciation for Visually Impaired People," 2019 Advances in Science and Engineering Technology International Conferences (ASET), Electrical and Computer Engineering Department, Effat University, Jeddah, KSA.
DOI: 10.1109/ICASET.2019.8714269

[10] R.Bansal, G.Raj and T.Choudhury, "Blur Image Detection using Laplacian Operator and Open-CV," in Proceedings of the 5th International Conference on System Modeling & Advancement in Research Trends (SMART-2016), IEEE Conference ID: 39669, Nov. 25-27, 2016, College of Computing Sciences & Information Technology, Teerthanker Mahaveer University, Moradabad, India.
DOI: 10.1109/SYSMART.2016.7894491

[11] R.A.Pagaduan, M.C.R.Aragon and R.P.Medina. iBlurDetect: Image Blur Detection Techniques Assessment and Evaluation Study. Technological Institute of the Philippines, Information Technology Department Quezon City, Philippines. https://www.scitepress.org/Papers/2020/103077/103077.pdf