International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 04 | Apr 2024 www.irjet.net p-ISSN: 2395-0072

DEVELOPING AN INTELLIGENT SYSTEM FOR MULTILINGUAL VISUAL TO AUDIO CONVERSION FOR PERSONS WITH VISUAL IMPAIRMENTS USING ARTIFICIAL INTELLIGENCE

Dr. S RAJAN[1], Ms. P SUBHAVARSHINI[2], Ms. A SOUNDARYA[3], Mr. S SAKTHIVEL[4], Mr. J VIJAYAKUMAR[5]

[1] *Professor, Dept. of ECE, Velalar College of Engineering and Technology, Thindal, Erode, Tamilnadu, India.*
[2,3,4,5] *B.E Final Year Students, Dept. of ECE, Velalar College of Engineering and Technology, Thindal, Erode, Tamilnadu, India.*

ABSTRACT

According to the World Health Organisation, around 40 million people in the world are blind, while another 250 million have some visual impairments. Reading poses a significant challenge for them. To address this, an automatic reader for Visually Impaired People is developed. It works by capturing text from documents using a webcam and converting it into digital format through Optical Character Recognition (OCR). The text is extracted from the visual image and converted to audio output, in which one can hear through the speaker. In this, it proposed an intelligent system for multilingual visual-to-audio conversion to facilitate audio accessibility for persons with visual impairments. It aims to serve both visually impaired people and illiterate people. The system can process visual images in multiple languages and generate audio descriptions in the user's preferred language. The audio descriptions are designed to be concise, descriptive, and accurate, providing users with a detailed understanding of the visual image content.

KEYWORDS

Raspberry Pi 4 Model B, Optical Character Recognition, OpenCV, Raspbian, Python, Pytesseract, Real VNC Viewer App

1. INTRODUCTION

Visual impairment is a condition that affects a significant percentage of the global population. People with visual impairments face significant challenges in accessing and interpreting visual information, particularly in a world that relies heavily on visual communication. To address this challenge, intelligent systems for multilingual visual-to-audio natural language processing technologies analyze visual media and provide audio descriptions of their content. By converting visual information into audio, these systems enable visually impaired people to access and understand visual media content that would otherwise be inaccessible to them. Many individuals with visual impairments are proficient in multiple languages, and they deserve equal access to information in all of them. Moreover, the internet has enabled global connectivity, making it imperative for assistive technologies to break down language barriers. This project aims to provide an overview of intelligent systems for multilingual visual-to-audio conversion for persons with visual impairments.

2. PROPOSED SYSTEM

An intelligent system for visual-to-audio conversion for persons with visual impairments is a technology that converts visual information, such as images, graphs, and videos, into audio signals, so that individuals with visual impairments can comprehend the visual information through sound. The system uses computer vision and machine learning algorithms to analyze the visual content and generate a corresponding audio description that describes the relevant details of the visual information.

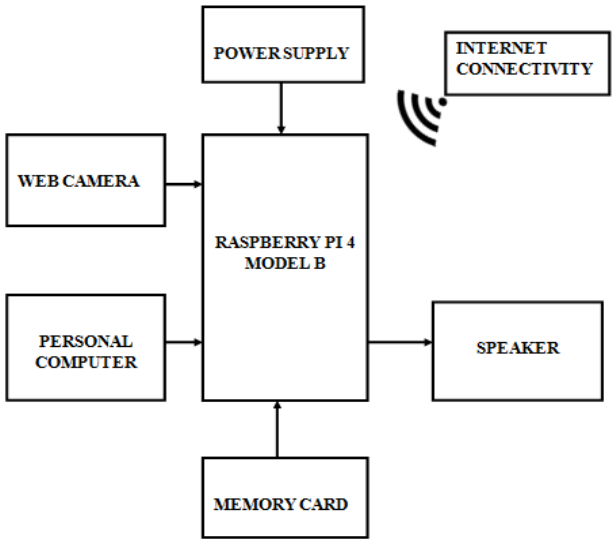

Figure 1 Block Diagram of Proposed System

The above Figure 1 shows the Block Diagram of the system, the image input is given to the web camera the

Captured Image is stored in the memory card and the PC is used to run the output using the Real VNC Viewer App to avoid such hardware components like keyboard, mouse, etc., There is a Power Supply for the Raspberry pi to engage with the kit. There is an Internet Connection for Text to Speech Conversion through wireless connection. The conversion of text to speech is observed through the speakers or headphones.

A. Raspberry Pi 4 Model B:

The Raspberry Pi 4 Model B is a credit card-sized single-board computer developed by the Raspberry Pi Foundation. It features a quad-core ARM cortex-A72 processor, up to 8GB of RAM, Multiple USB ports, HDMI output, Ethernet connectivity, GPIO pins, Wi-Fi, and Bluetooth capabilities. It supports Operating Systems like Raspberry Pi OS, Ubuntu, Linux distros, and third-party systems like Windows 10 IoT core. The below Figure 2 shows the Raspberry Pi 4 Model B,

Figure 2 Raspberry Pi 4 Model B

B. Web Camera:

The below Figure 3 below shows the Web Camera, it is also known as a webcam, it is a digital camera typically connected to a computer via USB or built directly into a device, such as a laptop or smartphone. It is used for capturing images, videos, live streaming, etc., It consists of a lens, image sensor, and electronics for processing and transmitting video data. Webcams vary in resolution, typically ranging from low-definition (480p) to high-definition (720p, 1080p or even 4k). Higher-resolution webcams offer better image quality.

Figure 3 Web Camera

C. Power Supply:

The Figure 4 below shows the USB cable to connect to Power. The Raspberry Pi 4 Model B typically requires a 5V DC power supply. The recommended power supply for the Raspberry Pi 4 Model B provides a minimum of 3amps (3000 mA) of current. The most common power connector for the Raspberry Pi 4 Model B is USB Type-C. The features built-in power management circuitry to protect against over-voltage, over-current and over-temperature conditions.

Figure 4 USB Cable

D. Memory Card:

The Raspberry Pi 4 Model B supports microSD cards, including 32GB variants. It provides ample storage space for operating system files, applications, and data on the Raspberry Pi 4 Model B. Below Figure 5 shows the 32 GB memory card,

Figure 5 Memory Card

E. Personal Computer:

A personal computer or PC is a computing device designed for general purposes, intended for individual use. It consists of a Central Processing Unit (CPU), Memory, Storage, Input / Output devices (such as a keyboard, mouse, and monitor), and various peripheral devices. Here, the PC is used to run the output using the Real VNC Viewer App to avoid such hardware components like keyboard, mouse, etc.,

F. Speaker:

The Figure 6 Shows the mini speaker, which is a device that converts electrical signals into sound waves. It consists of one or more transducers that vibrate in

response to the electrical signals received from an audio source. These vibrations generate sound waves that propagate through the air and can be heard by humans.

Figure 6 Speaker

G. Internet Connection:

To establish an internet connection with a Raspberry Pi 4 Model B we have several options, in that we use Wi-Fi Connection. It has built-in Wi-Fi connectivity, which allows you to connect it to a wireless network.

3. RESULTS AND DISCUSSIONS

The results of our project are shown below: First, we type a code - (sudo python3 code.py) in the page of Real VNC Viewer App, then the code works and it asks us to choose a specified language in the given 3 languages (English, French, Russian) in a specified number given for the three languages. Then the capturing page is opened and we click a picture like a language we selected before through a web camera, then it converts to Tamil language in text format and finally, we observe through a speaker or headphones.

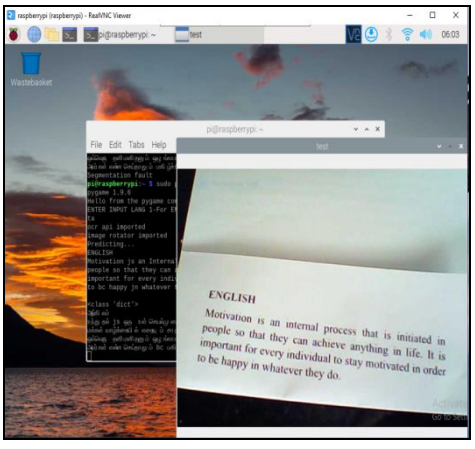

Figure 7 Translating English language to Tamil language

The above Figure 7 shows the translation of the English language to the Tamil language, for the English language the specified number is 1, first of all, the image is successfully captured by the web camera and it saves the image. The Captured image is converted to text format and the text is converted to speech by gTTS. The Speech cannot be displayed so it is not attached here.

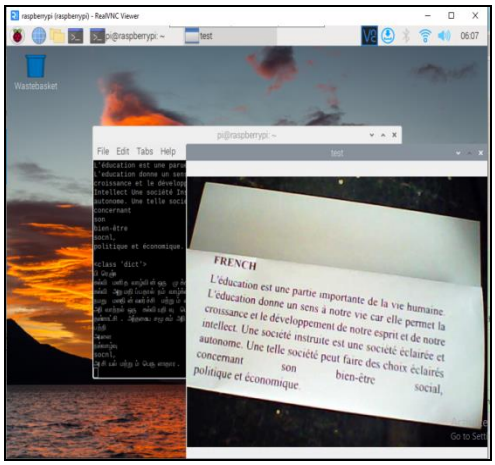

Figure 8 Translating French language to Tamil language

The above Figure 8 shows the translation of the French language to the Tamil language, for the French language the specified number is 2, first of all, the image is successfully captured by the web camera and it saves the image. The Captured image is converted to text format and the text is converted to speech by gTTS. The Speech cannot be displayed so it is not attached here.

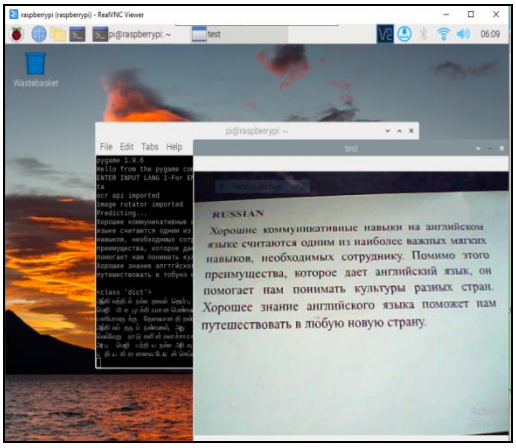

Figure 9 Translating Russian language to Tamil language

The Figure 9 shows the translation of the Russian language to the Tamil language, first of all, the image is successfully captured by the web camera and it saves the image. The Captured image is converted to text format and the text is converted to speech by gTTS. The Speech cannot be displayed so it is not attached here.

4. CONCLUSION

In conclusion, the development of intelligent systems for visual-to-audio conversion for persons with visual impairments has significant potential to improve accessibility and inclusivity for individuals with visual disabilities. Such systems can provide a more immersive and engaging experience for visually impaired individuals by converting visual content into audio formats that can be more easily understood and consumed.

REFERENCES

1. Chetana Thaokar, Hema Amlani, Nutan Rathod, Pooja Bodkhe, Pradnya Kolhe, Ujwalla Gawande, "Novel Machine Learning based Text-To- Speech Device for Visually Impaired People", 2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), (2023), pp.1-5.

2. Abbineni Charishma, Alla Amrutha Vaishnavi, D.Rajeswara Rao, Tirumalasetti Teja Sri, "Smart Reader for Visually Impaired", 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS),(2023),vol.1,pp.349-352.

3. B. J. Komathi, V. Kiruthika, C. Nithiya, J. Pavithra, N. G. Praveena, P. Samuda and "Arduino based Customized Smart Glasses for the Blind People", 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), (2022), pp. 1136-1141.

4. N.Kasthuri, S.Naveen Kumar, A.Nethra Krupa, R.Madhavan, "Finger Vision for Visually Impaired", 2021 Innovations in Power and Advanced Computing Technologies (i-PACT),(2021),pp.1-6.

5. T.Asha, T. N. S. Joshna, P.G. Jyothirmai, S.Khasimbee and Ravi ,"Raspberry pi based Smart Reader for Blind People", 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC),(2020),pp. 445-450.

6. Akshay Akole, Pallavi Angnani, Sangeeta Kumari, Yash Bhamare, Zaid Naikwadi, "Enhanced Braille Display Use of OCR and Solenoid to Improve Text to Braille Conversion", 2020 International Conference for Emerging Technology (INCET), (2020), pp.1- 5.

7. F. Al-Muqbali, K. Al- Kiyumi, N. Al-Tourshi and F. Hajmohideen, "Smart Technologies for Visually Impaired: Assisting and conquering infirmity of blind people using Al Technologies", 2020 12th Annual Undergraduate Research Conference on Applied Computing (URC), (2020), +pp. 1-4.

8. Rajan S, Abinaya S, Dharanidevi V, Dharanitharan V, Dharani B, “Intelligent Shopping Cart with Online Payment for Futuristic Shopping Experience”, International Research Journal of Engineering and Technology, Vol.7, Issue 5, May 2020, pp:405-409.

9. K.Kaushik, S.Kumar, K.Rao, S.Rao, S.Santhosh, "Smart Reader for Visually Impaired", 2020 Fourth International Conference on I-SMAC (IoT in Social Mobile Analytics and Cloud) (I-SMAC),(2020),pp. 846-850.

10. V.Ghiye and S.Musale "Smart reader for visually impaired", 2018 2nd International Conference on Inventive Systems and Control (ICISC),(2018),pp. 339-342.

11. M.H.Ashmafee and S.A.Sabab, "Blind Reader: An intelligent assistant for blind", 2016 19th International Conference on Computer and Information Technology (ICCIT),(2016),pp. 229234.