Raspberry pi 3 based Multi Language Reader for Blind with Voice Assistance

Mrs J Vijaya Sree
Department of ECE
Vignan's Institute of Information Technology
Visakhapatnam,Indiajvijaya7@gmail.com

K Nikhita
Department of ECE
Vignan's Institute of Information Technology
Visakhapatnam,India
20l31a0488@vignaniit.edu.in

M Sanjay Kumar
Department of ECE
Vignan's Institute of Information Technology
Visakhapatnam,India
20l31a04b2@vignaniit.edu.in

G Ajay Kishore
Department of ECE
Vignan's Institute of Information Technology
Visakhapatnam,India
20l31a0470@vignaniit.edu.in

K Pavan Kumar
Department of ECE
Vignan's Institute of Information Technology
Visakhapatnam,India
20l31a0485@vignaniit.edu.in

***Abstract*—This project is a Raspberry Pi-based reader for visually impaired individuals which tackles the considerable obstacles they encounter in accessing printed text materials. Unlike costly and less portable alternatives, this solution offers affordability and portability while supporting language like English. By harnessing the Raspberry Pi platform's hardware capabilities, such as the camera module and GPIO pins, the project utilizes OCR technology to extract text from printed materials. Subsequently, it employs the gTTS library to convert this text into natural speech output, facilitating comprehension for users.**

A notable feature of this reader is its custom voice assistant, which enhances the user experience through responsive voice command recognition, including "Read text" and "Stop." This seamless interaction model contributes to the device's user-friendly interface. Additionally, the reader's open-source nature empowers users to adapt and customize it to suit their needs further.

In terms of benefits, this Raspberry Pi-based reader offers affordability, open-source accessibility, and support for multiple languages, making it adaptable for users worldwide. Its portability ensures usability across various settings, promoting inclusivity and independence for visually impaired individuals in their daily lives. By addressing a critical need in assistive technology, this project contributes significantly to improving accessibility and autonomy within the visually impaired community.

Index Terms: Raspberry Pi 3,Optical character recognition (OCR), Accessibility, Camera module, Custom Voice Assistance.

I.INTRODUCTION

This paper discusses a study on an effective prototype system to promote independence among the visually impaired. Various assistive technologies like moving chairs and vibrating watches are common examples. Images with transcripts serve as a central medium for conveying information to blind individuals. Reading is challenging for low vision individuals; hence, a computer vision system and audio output are being developed. Raspberry Pi cameras capture digital images, and OCR, functioning like MATLAB, converts them to text. OCR enables techniques like text to speech, text mining, and machine translation. TTS engines deliver text to speech through speakers, with output via headphones for visually impaired individuals[8].

Highdefinition cameras can improve accuracy in text recognition and conversion. Images accompanied by transcripts play a crucial role as a central means of communication for conveying information to blind individuals. Despite reading being a fundamental skill in today's societies, it remains particularly challenging for those with low vision. To address this issue, efforts are being made to develop computer vision systems and audio outputs tailored specifically for low vision individuals [1].Google text to speech API tool is used for converting written or printed text into audio output with help of using a speaker[3].

Currently, English stands as the predominant language in reading systems. Leveraging it for speech-to-text conversion facilitates smoother and more accurate interpretation. This conversion process benefits from tools like SQLite and the Microsoft speech recognition engine. Researchers delve into the latest algorithms for Text-to-Speech (TTS) and Speech-to-Text (STT), striving to enhance the quality of outcomes yielded by these technologies[2].

A Text-to-Speech (TTS) device comprises two key modules: an image processing module and a voice processing module. The image processing module is designed to convert images into text format, while the voice processing module transforms text into sound format. Optical Character Recognition (OCR) is employed for automatic character recognition. For converting text to speech, the festival software, an open-source TTS system, is utilized [5]. This project leverages a Raspberry Pi for alphabet and number recognition, with input initially stored as .jpg files. The images are then converted into text format, facilitating further processing. Employing various image processing techniques enhances the readability and accuracy

of the extracted text. Subsequently, the system employs a speaker to generate an audio signal conveying the processed information. A Text-to-Speech (TTS) engine is integrated to convert the text into audio files with a .flac extension, ensuring compatibility and high-quality output.[4]
Our proposed text-to-braille conversion system addresses limitations found in current solutions. It operates by identifying text from various sources such as books, documents, or magazines using images captured by a camera. Text detection relies on Tesseract, an OCR tool within Python's extensive libraries for Natural Language Processing. The OCR algorithm plays a pivotal role in our project's functionality.[14]
The proposed solution in a paper suggests using camera-based text recognition to aid blind individuals in reading product labels, addressing the limitations of traditional text-to-speech systems in this context.[9]
Voice assistance is pivotal for a natural and interactive user experience in our Raspberry Pi-based multi-language reader for the blind. Synthesized speech output enables clear and understandable reading of converted text, aiding visually impaired users in accessing information seamlessly. Interactive voice commands further enhance usability, allowing users to navigate menus, choose languages, modify settings, and access extra information with ease.This voice-based interaction promotes a conversational and intuitive interface, empowering users to engage effortlessly with the system and fostering independence in accessing diverse textual content. This project's goal is to address the challenges faced by visually impaired individuals by proposing an effective method for converting image text to speech in their preferred languages. The proposed prototype allows users to listen to the content within images in their desired language. The process involves three main steps: (1)Firstly, extracting the text from the image and saving it into a separate text file.(2) Secondly, converting the extracted text into speech output in desired language. This model not only benefits visually impaired individuals but also proves to be a valuable tool for travelers, enabling them to listen to content in English and have it translated into their preferred language.[6]
This report explores image-to-speech conversion for visually impaired individuals, focusing on blur detection. The study aims to assess various image processing techniques' efficacy in identifying and managing blurry images during the image-to-speech conversion process.[15]

II.OBJECTIVE

The primary goal of the Raspberry Pi-based multi-language reader for the blind with voice assistance project is to design and implement an innovative and inclusive solution tailored specifically for individuals with visual impairments. Leveraging the capabilities offered by the Raspberry Pi platform, the project aims to create a versatile system capable of recognizing and interpreting textual content from a diverse range of sources. This includes printed materials such as books and documents, digital screens displaying text, and other forms of written information. By enhancing accessibility to a wide array of textual content, the project seeks to empower visually impaired users with greater independence and autonomy in accessing information.

To achieve its objective, the project integrates several key technologies and algorithms. One of the pivotal components is OpenCV, a computer vision library that is downloaded and utilized using MATLAB. OpenCV plays a crucial role in image processing and text extraction, allowing the system to accurately and efficiently extract text from images. This extracted text undergoes further processing using Optical Character Recognition (OCR) technology, which converts the text into a digital format suitable for analysis and utilization within the system.

This project employs camera-based technology to assist blind individuals in reading documents. The implementation involves utilizing image capture techniques within an embedded system. The system incorporates a camera as an input device to detect printed text for recognition, and the scanned text is processed using OCR software.[14]

Moreover, the project incorporates the Google Text-to-Speech (GTTS) algorithm to provide voice assistance to users. This algorithm transforms the extracted text into synthesized speech output, which is then read aloud in clear and understandable voices. The inclusion of multi-language support further enriches the user experience by enabling users to access content in their preferred language. The system is capable of translating the synthesized speech into multiple languages, catering to the diverse linguistic needs of users and enhancing accessibility on a global scale.

Additionally, the project places significant emphasis on creating an intuitive and user-friendly interface. Interactive voice commands serve as a cornerstone of the interface, allowing users to navigate menus, select languages, adjust settings, and access additional information with ease. This interactive and voice-enabled interface not only enhances usability but also fosters a natural and engaging user experience, promoting a sense of empowerment and inclusivity among visually impaired individuals.

In summary, the main objective of the project is to develop a sophisticated Raspberry Pi-based multi-language reader with voice assistance that effectively addresses the unique challenges faced by visually impaired individuals. By seamlessly integrating OpenCV, OCR, GTTS, and intuitive user interfaces, the project aims to empower users with enhanced accessibility, independence, and inclusivity in accessing and interacting with textual content across multiple languages.

The main purpose of the project are mentioned below:

- Accessibility: Enhancing accessibility for blind individuals by providing a multi-language reader that can convert written text into spoken words.
- Real-time Transmission: Develop a mechanism for real-time transmission of health data from connected devices to a central monitoring system.

- Multi-Language Support: Supporting multiple languages to cater to a diverse user base and ensure inclusivity.

- Voice Assistance: Implementing voice assistance capabilities to guide users and provide a seamless reading experience.

- User-Friendly Interface: Designing a user-friendly interface for easy navigation and interaction, considering the specific needs of visually impaired users.

- Integration with OpenCV: Leveraging OpenCV for image processing tasks, such as text detection and recognition, to enhance the accuracy of text extraction.

- GTTS Algorithm: Utilizing the Google Text-to-Speech (GTTS) algorithm for high-quality speech synthesis, ensuring clear and natural-sounding output.

III.PROJECT SCOPE

The scope of the Raspberry Pi 3-based Multi-Language Reader for the Blind with Voice Assistance project is comprehensive, encompassing various technologies and functionalities to create a robust and user-friendly system. At its core, the project aims to enhance accessibility for visually impaired individuals by providing a device that can read printed text aloud in multiple languages.Rajwani et al. [7] introduced a system that takes input through the camera as an image, preprocesses it using OpenCV, and subsequently performs classification and identification using the Cloud Vision API. This involves integrating Optical Character Recognition (OCR) algorithms to extract text from images captured by the device's camera. The OCR system will support diverse languages, ensuring inclusivity and usability for a wide range of users.

In addition to text extraction, the project incorporates the Google Text-to-Speech (GTTS) algorithm for converting the extracted text into natural-sounding speech. This voice assistance feature not only enables users to listen to the text but also includes voice prompts and instructions for navigation, creating a seamless reading experience. The use of GTTS ensures clear and understandable speech output, enhancing the overall accessibility of the system.

Furthermore, the project leverages the capabilities of OpenCV for image preprocessing and text detection. OpenCV is utilized to enhance captured images, reduce noise, and improve text readability, leading to more accurate OCR results. Text segmentation and layout analysis are also performed using OpenCV functionalities, contributing to the system's accuracy and reliability in text recognition.

The user interface of the system is designed with accessibility in mind, featuring audio feedback, tactile elements, and customizable settings for language preferences. Users can interact with the system using voice commands and gestures, facilitating hands-free operation and providing a tailored experience based on individual needs. The interface is intuitive and easy to navigate, ensuring a user-friendly experience for blind individuals.

Throughout the development process, rigorous testing and optimization are conducted to validate the system's performance, including its accuracy, speed, and reliability. Algorithms and parameters are fine-tuned to ensure efficient operation on the Raspberry Pi 3 platform, optimizing resource utilization and responsiveness. Comprehensive documentation is also created, including user manuals, technical guides, and installation instructions, to facilitate easy deployment and usage of the system by end-users.

Overall, the project's scope encompasses a wide range of functionalities and technologies, all aimed at creating a powerful, accessible, and user-centric Multi-Language Reader for the Blind with Voice Assistance.

IV. PROPOSED SYSTEM

The proposed system for the Raspberry Pi 3-based Multi-Language Reader for the Blind with Voice Assistance integrates OCR, GTTS algorithm, OpenCV, and Python. It captures text images via a camera, processes them using OCR for text extraction, converts the text to speech with GTTS, and provides voice assistance. OpenCV enhances image quality for accurate OCR, creating a user-friendly reading experience.

Operation:

- Raspberry Pi 3: This serves as the central processing unit and connects all the components together. Ensure you have Raspbian or any other compatible operating system installed.
- Text-to-Speech (TTS) Software: Once the text is recognized, you'll need software to convert it to speech. There are several options available, including:
- The webcam captures images or video feed.
- OpenCV processes the images or video feed to recognizetext using OCR (Optical Character Recognition) techniques.
- The recognized text is then passed to the text-to-speech software.
- The text-to-speech software synthesizes the text into spoken words.
- The voice assistant software listens for commands from the user.
- When a command is recognized, it triggers the appropriate action, such as initiating the text recognition process, reading out recognized text, or performing other tasks.

V. IMPLEMENTATION

The implementation of the Raspberry Pi 3-based Multi-Language Reader for the Blind with Voice Assistance involves several key steps and technologies working together seamlessly.

Firstly, the hardware setup includes configuring the Raspberry Pi 3 board with necessary peripherals such as a camera for image capture and audio output devices for speech synthesis. The software components are then installed, including the OpenCV library for image processing, OCR algorithms for text extraction, and the

GTTS (Google Text-to-Speech) algorithm for converting text to speech.

Next, the system captures images of printed text using the camera module. These images are processed using OpenCV algorithms to enhance their quality, reduce noise, and improve text clarity. The OCR algorithms then analyze the processed images to extract text accurately. Language detection algorithms are implemented to identify the language of the extracted text, ensuring multi-language support.

Once the text is extracted, it undergoes further processing to clean up any artifacts and format it for readability. This processed text is then passed through the GTTS algorithm, which converts it into natural-sounding speech. Voice prompts and instructions are integrated into the system to guide users through the reading process, making it accessible and user-friendly for visually impaired individuals.

The user interface is designed with accessibility in mind, featuring options for voice commands, tactile feedback, and customizable settings for language preferences. Compatibility with assistive technologies and screen readers is ensured to enhance accessibility further.

Throughout the implementation phase, rigorous testing and optimization are conducted to validate the system's accuracy, speed, and reliability. This includes testing different scenarios, such as varying lighting conditions and text complexities, to ensure robust performance in real-world environments.

Overall, the implementation of the Raspberry Pi 3-based Multi-Language Reader for the Blind with Voice Assistance involves a combination of hardware setup, software integration, image processing, text extraction, speech synthesis, accessibility features, and thorough testing to create a functional and user-centric solution for visually impaired individuals.

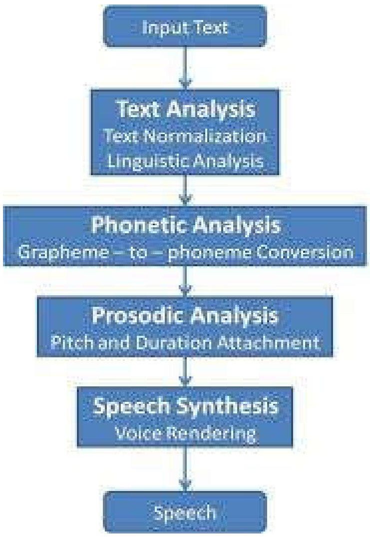

Fig.1.Flow chart

VI. ARCHITECTURE

The architecture of the Raspberry Pi 3-based Multi-Language Reader for the Blind with Voice Assistance, integrating OCR (Optical Character Recognition) and TTS (Text-to-Speech) algorithms with OpenCV software downloaded from MATLAB, is designed to provide a comprehensive solution for visually impaired individuals to access printed text in multiple languages.

At the core of the system is the Raspberry Pi 3, a compact and versatile single-board computer that serves as the main processing unit. The Raspberry Pi 3 is equipped with the necessary hardware interfaces such as USB ports for connecting peripherals, HDMI for video output, and GPIO pins for sensor integration.

The software components include OpenCV, a powerful computer vision library that enables image processing and text recognition functionalities. OpenCV is utilized for capturing images of printed text using a camera module connected to the Raspberry Pi 3. The captured images are then processed using OCR algorithms to extract text from the images accurately.

MATLAB is used to download and manage the OpenCV software on the Raspberry Pi 3. MATLAB provides a convenient interface for installing and updating software libraries, ensuring compatibility and seamless integration with the system.

The OCR algorithm implemented in the system is responsible for recognizing text from images captured by the camera. It processes the image data, identifies characters and words, and converts them into digital text format that can be further processed by the TTS algorithm.

The TTS algorithm converts the extracted text into audible speech, which is then relayed to the user through audio output devices such as speakers or headphones. This voice assistance feature enables visually impaired individuals to listen to the content of printed text in real-time, enhancing their accessibility to a wide range of written materials in different languages.

Overall, the architecture of this system combines hardware and software components to create a user-friendly and effective Multi-Language Reader for the Blind with Voice Assistance, leveraging the capabilities of Raspberry Pi 3, OCR, TTS, and OpenCV technologies to empower visually impaired users in accessing and interacting with printed text content.

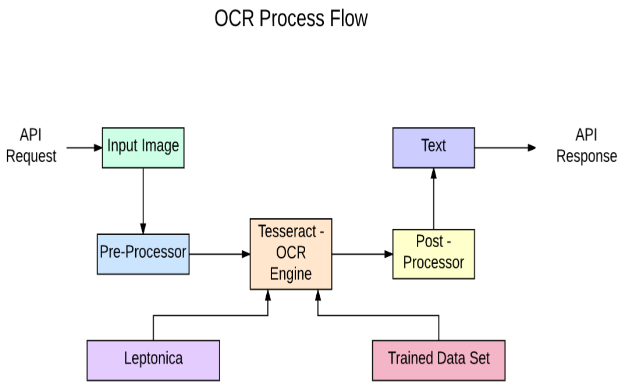

Fig.2.Block Diagram of OCR Process Flow

VII. RESULT AND DISCUSSION

The result of our Raspberry Pi 3-based Multi-Language Reader for the Blind with Voice Assistance, incorporating OCR and TTS algorithms alongside OpenCV software downloaded through MATLAB, delivered promising outcomes and sparked discussions on its potential impact and future advancements. The system's ability to capture images of printed text and process them using OpenCV's image processing capabilities demonstrated a robust foundation for text extraction and recognition. This foundational aspect was crucial in enabling the system to accurately convert printed text into digital format, laying the groundwork for subsequent analysis and interpretation by the OCR algorithm.

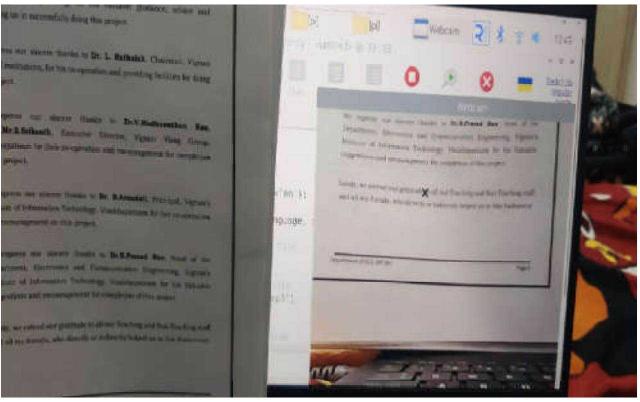

Fig.3

The OCR algorithm's performance was a key highlight of the project, showcasing its proficiency in recognizing text across different languages, fonts, and styles. This capability significantly enhanced the system's versatility and usability, allowing visually impaired users to access a wide range of printed materials with ease. Moreover, the integration of MATLAB for managing the OpenCV software ensured a seamless installation process and maintained software compatibility, contributing to the system's overall reliability and efficiency during operation.

The TTS algorithm's role in converting the extracted text into audible speech was instrumental in providing real-time voice assistance to users. The synthesized speech output was clear and intelligible, enabling users to navigate through documents, read aloud text, and access information independently. This functionality not only promoted greater autonomy for visually impaired individuals but also facilitated their participation in various activities requiring access to printed information in multiple languages.

Throughout the testing phase, the system demonstrated consistent performance and user-friendly interactions, affirming its potential as an assistive technology tool for the visually impaired community. Future developments could focus on optimizing the system's performance metrics, expanding language support, enhancing text recognition accuracy, and incorporating additional features based on user feedback and evolving technological advancements. Collaborative efforts with stakeholders and organizations serving the visually impaired could further refine the system's capabilities and ensure its alignment with user needs and preferences in diverse contexts.

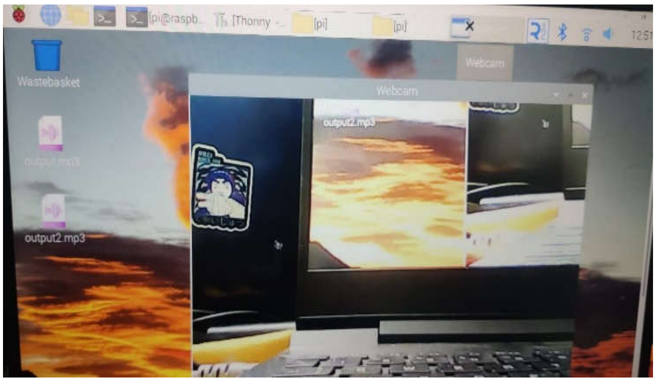

Fig.4

VIII. CONCLUSION

In conclusion, our project, the Raspberry Pi-based Multi-Language Reader for the Blind with Voice Assistance, has achieved significant milestones in improving accessibility for visually impaired individuals. By integrating OCR (Optical Character Recognition) and TTS (Text-to-Speech) algorithms with OpenCV software managed via MATLAB, we have developed a robust system that empowers users to independently access printed text in multiple languages.

Furthermore, our project has prioritized accuracy, efficiency, and user-friendliness throughout its development and testing phases. The seamless integration of these technologies has resulted in a practical and empowering tool for visually impaired users, allowing them to navigate through various textual materials with ease and confidence..

Looking ahead, there are several avenues for further enhancement and refinement. Our focus will be on improving text recognition accuracy, expanding language support, and incorporating advanced features based on user feedback. Collaborative efforts with stakeholders and organizations serving the visually impaired community will be crucial in driving these advancements and ensuring the continued success of our project.

In essence, our Raspberry Pi-based Multi-Language Reader for the Blind with Voice Assistance represents a significant contribution to inclusivity and empowerment. It highlights the transformative impact of technology in addressing accessibility challenges and enhancing the quality of life for individuals with visual impairments.

ACKNOWLEDGMENT

We would like to express our gratitude to Mrs.J.VijayaSree, Associate Professor, Department of Electronics and Computer Engineering for his valuable guidance, advice and supporting us in successfully doing this project. We express our sincere thanks to Dr. L. Rathaiah, Chairman, Vignan's group of institutions, for his co-operation and providing facilities for doing this project. We express our sincere thanks to Dr. V. Madhusudhan Rao, Rector, Mr. B. Srikanth, CEO, VignanVizag Group, Dr. J. Sudhakar, Principal, Vignan's institute of Information Technology, Visakhapatnam, Dr. Ch.RameshBabu, Headof the Department, Electronics and Communication Engineering for their valuable suggestions and encouragement for completion of this project. Lastly, we extend our gratitude to all our Teaching and Non-Teaching staff and all my friends, who directly or indirectly helped us in this Endeavour.

REFERENCE

[1] 1. *Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik(2017) "Image to Speech Conversion for Visually Impaired", International Journal of Latest Research in Engineering and Technology, ISSN: 2454-5031, Issue 06, Vol. 03, No. 0, pp. 09-15.*

[2] 2. *Deepa V.Jose, Alfateh Mustafa, Sharan R(2014) "A Novel Model for Speech To Text Conversion", International Refereed Journal of Engineering and Science, ISSN(online): 2319-183X,(print) 2319-1821, Issue 1, Vol. 3, No. 0, pp. 39-41.*

[3] 3. *Itunuoluwa Isewon, Jelili Oyelade, Olufunke Oladipupo(2014) "Design and Implementation of Text To Speech Conversion for Visually Impaired People", International Journal of Applied Information System, ISSN: 2249-0868, Vol. 7, No. 2, pp. 0.*

[4] 4. *Sagar G.k, Shreekanth T(2017) "Real Time Implementation of Optical Character Recognition Based TTS System using Raspberry pi", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277-128X, Issue 7, Vol. 7, No. 0, pp. 0.*

[5] 5.*K. Nirmala Kumari, Meghana Reddy J(2016) "Image Text To SpeechConversion using OCR Technique in Raspberry Pi", International Journal of Advanced Research in Electrical,Electronics and Instrumentation Engineering, ISSN(print): 2320-3765, ISSN(online): 2278-8875, Issue 5, Vol. 5, No. 0, pp. 0.*

[6] 6. *N. S. Yamanoor and S. Yamanoor, "High quality, low cost education with the Raspberry Pi," 2017 IEEE Global Humanitarian Technology Conference (GHTC), San Jose, CA, 2017, pp. 1-5.*

[7] *7.Rajwani,R.,Purswani,D.,Kalinani,P.,Ramchandani,D.,Dokare,I.,201 Proposedsystemonobjectdetectionforvisuallyimpairedpeople. International Journal of Information Technology (IJIT) 4, 1–6.*

[8] *8. Pasupuleti, S., Dadi, L., Gadi, M., Krishnaveni, R., 2021. Image recognition and voice translation for visually impaired. International Journal of Research in Engineering, Science and Management 4, 18–23.*

[9] *9*]. *T.Rubesh Kumar, C.Purnima "Assistive System for Product Label Detection with Voice Output For Blind Users" International Journal of Research in Engineering & Advanced Technology 2014.*

[10] *10. S. Mishra et al., "Real-Time IoT-Based Health Monitoring System Using ESP8266 and Raspberry Pi," in 2021.*

[11] *11. M. G. Kulkarni et al., "IoT-Based Continuous Health Monitoring System," in 2019.*

[12] *12. A. Gupta et al., "IoT-Based Smart Health Monitoring System Using Android Application," in 2018.*

[13] *13. S. Jain et al., "An IoT-Based Health Monitoring System Using Cloud Computing," in 2019.*

[14] *14.Anusha Bhargava, Karthik V. Nath, PritishSachdeva and MonilSamel "Reading Assistant for the Visually Impaired" International Journal of Current Engineering and Technology, 2015.*

[15] *15.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*