ENHANCED VISION: A MACHINE LEARNING APPROACH FOR SMART GLASSES AIDING THE VISUALLY IMPAIRED

Project Reference No.: *47S_BE_0606*

College ***:*** *Atria Institute of Technology*
Branch ***:*** *Department of Electronics and Communication Engineering*
Guide(s) ***:*** *Mr. Somesh B S*
Student(S) ***:*** *Ms. Keerthana R*
Mr. Shaik Baba Hassan
Ms. Shwetha Satish
Ms. Spoorthy S

Keywords:

Raspberry Pi, Smart glasses, Visual aid, Raspberry Pi camera, Tesseract OCR, image to text, text to speech, image processing

Introduction:

Introducing our transformative project: pioneering smart glasses set to redefine the lives of visually impaired individuals. These glasses merge cutting-edge technology with unwavering dedication to accessibility, empowering users with newfound independence and enhanced navigation.
At the core, these glasses employ machine learning to extract English text from captured images, seamlessly converting it into spoken words. We recognize the limitations of Braille and aim to transcend them, offering a solution that prioritizes spoken output while acknowledging Braille's challenges.
Existing models like OrCam MyEye and eSight offer valuable functionalities but are limited by high costs and bulky designs. Our smart glasses bridge this gap, providing affordability and user-centric design. Continuous updates ensure our glasses remain at the forefront of assistive technology, constantly refining the user experience and functionality. Beyond technology, our project advocates for societal change and equal opportunities, fostering a more inclusive world for all.
By working hand in hand with stakeholders and driving positive change, we aspire to inspire a global movement towards a more accessible and equitable future.

Objectives:

1. Extract English text from writings in the image.
2. Transform the extracted English text into an audible voice message.

Methodology:

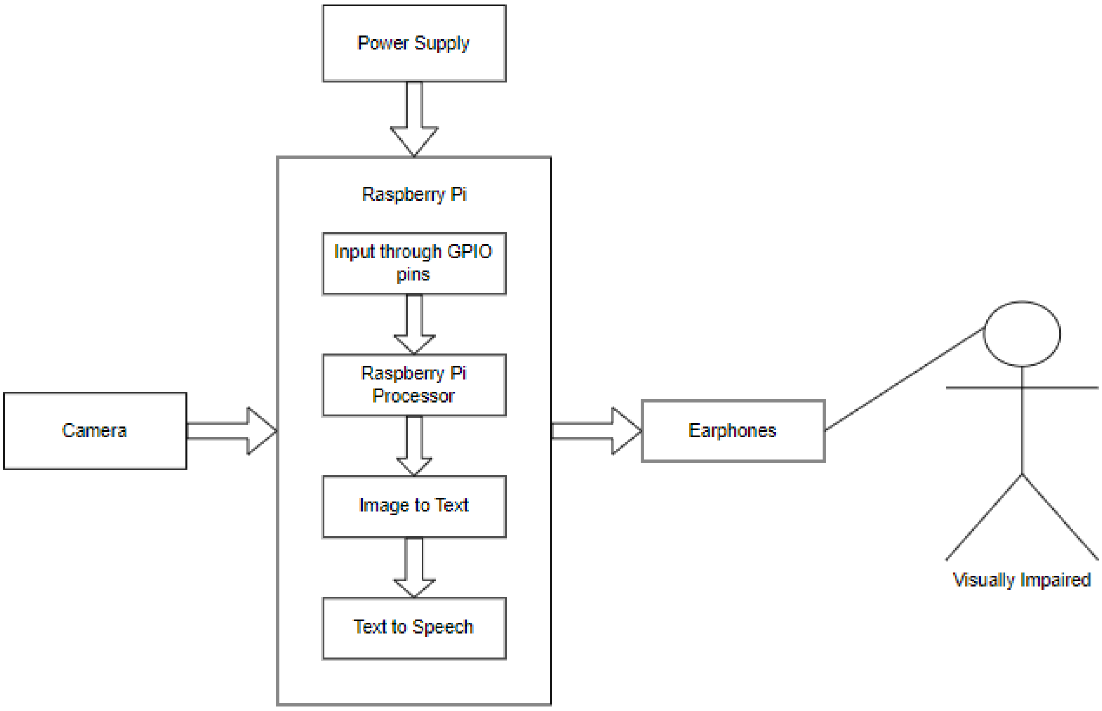

Fig1: Block Diagram of the Project

The Picamera2 module ensures real-time acquisition of visual data, contributing to the system's responsiveness and effectiveness. Through grayscale conversion and thresholding, preprocessing techniques enhance image quality and isolate text regions for accurate extraction. Tesseract OCR's robust text recognition capabilities enable precise extraction of textual information from images. In the second phase, Espeak synthesizes speech from extracted text, providing a clear and comprehensible output for visually impaired users. Parameters such as voice, speed and pitch offer flexibility and personalization in the synthesized speech output.

The integration of hardware and software components underscores the system's versatility and adaptability to diverse users. Smart glasses equipped with text recognition technology represent a significant advancement in accessibility for the visually impaired community. By converting visual information into synthesized speech, the system empowers users to access and comprehend textual content in their surroundings. The methodology's reliance on machine learning algorithms demonstrates the potential of AI technologies to address real-world challenges and improve quality of life. Raspberry Pi's affordability and accessibility make it an ideal platform for developing assistive technologies like smart glasses.

Python's simplicity and rich libraries enhance the ease of development and implementation of machine learning-based solutions. OpenCV's comprehensive functions and algorithms provide robust support for image processing tasks essential for text recognition. Tesseract OCR's accuracy and versatility make it a valuable tool for extracting text from images, particularly for applications like smart glasses.

Espeak's lightweight design and customizable parameters make it well-suited for generating synthesized speech from extracted text. The methodology's systematic approach to image preprocessing and text extraction ensures reliable performance and accuracy. By leveraging machine learning algorithms, the system can adapt and improve over time, enhancing its effectiveness and usability.

The integration of hardware and software components into wearable devices reflects a user-centric design approach focused on enhancing accessibility and usability. Smart glasses equipped with text recognition and speech synthesis capabilities represent a significant step towards empowering visually impaired individuals to navigate and interact with their environment independently.

Conclusion:

The testing of the smart glasses prototype indicates its capability to assist visually impaired users by converting text from images into audible voice. The successful extraction of text from diverse fonts confirms the reliability and versatility of the image processing and text-to-speech conversion algorithms. This functionality enhances accessibility and usability for visually impaired individuals, offering a practical solution for real-time text recognition and auditory.

Scope for future work:

This project can be further enhanced by focusing on performance optimization for resource-constrained devices like Raspberry Pi. Techniques like predefined graphics, parallel processing, or GPU utilization can be explored to achieve faster and more efficient image-to- speech conversion. Additionally, user experience can be improved through a user-friendly interface with features like voice commands, gesture recognition, or haptic feedback, especially for users with disabilities or those seeking a more interactive experience. Beyond its core functionality, the project holds promise for integration with assistive technologies and academic applications. Collaboration with stakeholders like disability organizations and educators can lead to adaptations that empower individuals with diverse needs. Opensourcing the project can foster a community of developers who can contribute to its ongoing development and explore its potential in educational settings and community outreach initiatives. In addition to its existing capabilities, the project can go deeper into understanding the quote and analyze its meaning and context. Integration with knowledge bases and other data can support information storage through reflection. Security and privacy are important. It requires strong data protection and user consent. Scalability can be explored by deploying cloud or edge applications to achieve broader and more efficient services. Advanced user interactions such as multiple input and feedback can improve user experience and performance. These future directions pave the way for more powerful and versatile picture-to-speech systems that have great potential across a wide range of users.

Reference:

[1] "Smart Glasses for Blind People" Dr. Sankit Kassa, Manisha Inchanalkar, Amisha Kamble, Mayuri Inchanalkar, Mumbai, May, vol. 6 in Proc. 12th Int. Conf. Cloud Comput. Data Sci. Eng., Jan. 2023, pp. 617622, doi: 10.1109/ Con uence52989.2023.9734204 pp. 2274- 2023.

[2] "A Glass Based Real-Time Scene Analysis for the Visually Impaired Masud for visually impaired people obstruction avoidance through object detection and classification, IEEE Access, vol. 10, pp. 1342813441, 2022, doi: 10.1109/ACCESS.2022.3146320.

[3] "A Wearable Smart System to Assist Visually Challenged", Kun Xia, Xueyong Li, Haiyang Liu, Mingli Zhou, and Kexin Zhu, IEEE, India, 2021., in Proc.Int. Conf. Syst., Comput., Automat. Netw. (ICSCAN), pp. 2274- 2021. doi: 10.1109/ICSCAN53069.2021.9526495.

[4] "Advanced Audio Aid for Blind People", Savera Sarwar, Muhammad Turab, Danish Channa, Aisha Chandio, M. Uzair Sohu, Vikram Kumar IEEE Nodia ,2021. Proc. IEEE 2nd Int. Conf. Electr. Power Energy Syst. (ICEPES), Dec. 2021, pp. 15, doi: 10.1109/ICEPES52894.2021.9699479.

[5] "Smart Glasses for Visually Impaired Person", Dr. Suvarndyal, Prajita R Udgiri, & Sakshi Sherikar, IEEE Trans. Consum. Electron., vol. 67, no. 1, pp. 311, Feb. 2022, doi: 10.1109/TCE.2022.3037065.