International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.047 ISSN: 2581-4621

Gesture Based Speaking System Using CNN

Devarakonda S S Siddhardha[1], Chata Nandu Kumar Yadav[2], Kaushik Chevendra[3], K. Narasimha Rao[4]

[1, 2, 3] UG- Electronics and Communication Engineering, Maturi Venkata Subba Rao Engineering College, Nadergul, Hyderabad-501510

[4]*Assistant Professor, Electronics and Communication Engineering, Maturi Venkata Subba Rao Engineering College, Nadergul, Hyderabad-501510*

***Abstract*—** In this work, we propose a system which helps the visually disabled to interact effectively with the world through hand gestures. Computer Vision and Deep Learning has opened new avenues for gesture recognition solutions. A gesture vocalizer is introduced in this work which is based on CNN and Gesture to Audio Converter. A lightweight CNN model has been trained on distinct hand gestures which are captured by Raspberry Pi camera. In Real time the model is used for inference on live camera feed from raspberry pi and recognizes these gestures. Based on the recognized gesture action such as Text or Audio Response associated with each gesture is transcribed.

Keywords— Raspberry pi, Convolutional Neural Networks (CNN), Open-Source Computer Vision (OpenCV), Text to speech Libraries (TTS), MediaPipe, eSpeak.

INTRODUCTION

This paper, the Hand Gesture-Based Speaking System for Disabled Individuals, is an innovative endeavor aimed at breaking down barriers to communication and control for individuals facing different impairments. This project harnesses the power of Raspberry Pi, a versatile and cost-effective computing platform, along with Convolutional Neural Networks (CNNs), a sophisticated technique, to create a user-friendly interface. We recognize that disabilities vary widely, and one-size-fits-all solutions are inadequate. Our system is designed in a way that it is adaptable, allowing users to tailor gestures and commands to suit their unique needs and preferences.

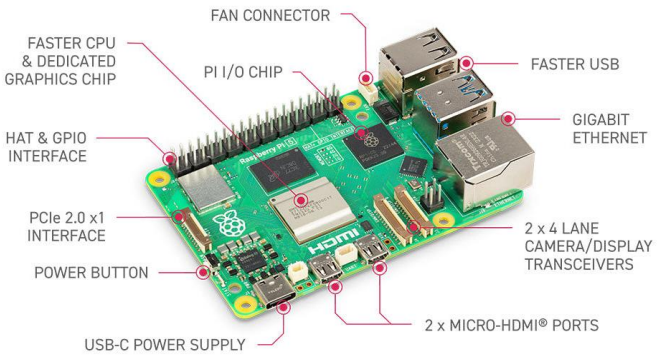

Figure 1: Raspberry Pi 4 Hardware

problem statement

People with disabilities often encounter barriers in accessing and interacting with their environment, hindering their independence and quality of life. Traditional assistive technologies are often

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.047 ISSN: 2581-4621

specialized for specific disabilities, limiting their reach and effectiveness across diverse user groups. To address this challenge, our project aims to develop a versatile assistive system leveraging computer vision and deep learning technologies to enhance accessibility and communication for individuals with various disabilities. By utilizing hand gestures as a universal interface, our system seeks to empower individuals with visual, auditory, motor, and cognitive impairments to navigate and interact with their surroundings more effectively.

existing system

Gesture recognition systems are pivotal in human-computer interaction, with applications ranging from robotics to accessibility solutions. This paper presents enhancements to an existing gesture recognition system deployed on the Raspberry Pi platform. The original configuration relied on a rudimentary setup with the Raspberry Pi's internal laptop camera. In contrast, our proposed system integrates a wired webcam, augmenting efficiency and accuracy in gesture detection. Additionally, we introduce text-to-speech (TTS) capabilities through libraries such as eSpeak, fostering improved user interaction and system adaptability. These enhancements collectively mitigate error susceptibility while affording greater flexibility in expanding gesture repertoire and corresponding feedback messages.

proposed system

This work presents an integration of hand gesture recognition with speech synthesis, enhancing human-computer interaction capabilities. Using the MediaPipe framework, real-time detection of hand gestures is achieved using pre-trained CNN models. These gestures are subsequently mapped to corresponding textual messages, facilitating seamless communication between users and systems. Through the utilization of text-to-speech (TTS) libraries, the textual messages are converted into spoken language, thereby augmenting the accessibility and usability of the interface. The system's architecture is carefully designed to ensure robustness, accuracy, and real-time performance. Extensive testing validates the effectiveness of the proposed approach, demonstrating its potential for practical applications across various domains, including assistive technology, gaming, and interactive interfaces.

In conclusion, the seamless integration of hand gesture recognition with speech synthesis offers a powerful and intuitive means of human-computer interaction, promising a future where technology adapts to user needs with unprecedented ease and inclusivity.

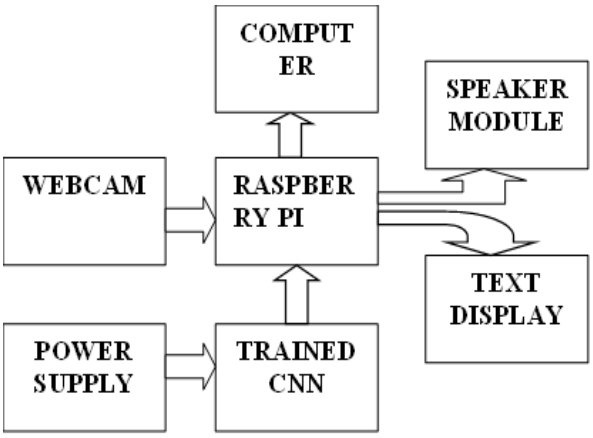

Figure 2: Block diagram of Hardware

DOI:10.46647/ijetms.2024.v08i03.047 ISSN: 2581-4621

hardware requirements

- Camera Module
- Raspberry Pi.
- Power Supply

A. Camera

Webcams for hand gesture recognition rely on computer vision algorithms to detect and track hand movements in real-time video streams. Techniques like background subtraction and machine learning classifiers are commonly used for accurate detection. The Camera Module for hand gesture recognition utilizes specialized algorithms and hardware to interpret hand movements accurately. It enables applications in gaming, virtual reality, and human-computer interaction by detecting gestures with high precision. Developers can leverage its SDKs and APIs to integrate gesture recognition functionalities seamlessly into their systems.

B. Raspberry Pi

Raspberry Pi is a series of small, affordable, single-board computers developed by the Raspberry Pi Foundation, primarily intended for educational and hobbyist purposes. It features a low-cost ARM-based processor, RAM, USB ports, HDMI output, GPIO pins for interfacing with external devices, and support for various operating systems, including Linux-based distributions. Raspberry Pi can be used for hand gesture recognition by integrating with cameras and machine learning libraries like OpenCV and TensorFlow. With its low-cost and compact size, Raspberry Pi offers a versatile platform for developing gesture recognition applications. Developers can utilize pre-trained models or train custom models to recognize specific hand gestures.

C. Power Supply

Raspberry Pi requires a stable 5V DC power supply to function properly. It's recommended to use a power supply with a minimum current rating of 2.5A to ensure sufficient power delivery. Official Raspberry Pi power supplies or third-party supplies from reputable manufacturers are advisable for reliability. Ensure the power supply has the correct connector type (micro-USB or USB-C) for your Raspberry Pi model. Due to limited supply of voltage the raspberry pi is less susceptible to overheating. The Raspberry Pi power supply is a crucial component for ensuring stable and reliable operation of the Raspberry Pi board. It typically provides 5 volts of DC power and sufficient current to meet the requirements of the Raspberry Pi model being used. The power supply should have a micro-USB connector compatible with the Raspberry Pi's power input.

software requirements

- Convolutional Neural Networks (CNNs)
- Python
- OpenCV
- Text-to-Speech (TTS) libraries
- VNC Viewer
- Putty

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are pivotal in hand gesture recognition, enabling machines to interpret gestures from images or video frames. CNNs extract intricate spatial features from hand images through convolutional layers, capturing gesture nuances. By training on labeled hand gesture datasets, CNNs learn to classify gestures with high accuracy. This technology finds applications in sign language translation, human-computer interaction, and virtual reality interfaces. Leveraging CNNs for hand gesture recognition empowers seamless and intuitive human-machine communication. There are different datasets in CNN but the google media pipe does not use any

DOI:10.46647/ijetms.2024.v08i03.047 ISSN: 2581-4621

particular dataset but rather a framework for building various perception pipelines. It provides tools and pre-trained models for tasks such as hand tracking, pose estimation, face detection, and more. While MediaPipe itself doesn't come with a dedicated dataset, it can be trained and used with various datasets depending on the task at hand. For instance, for hand gesture recognition using MediaPipe, researchers anddevelopers might utilize datasets like the American Sign Language (ASL) dataset or the ChaLearn Looking at People (CLAP) dataset to train their models.

Python

It is a Programming language for development. Python is commonly used for recognizing hand gestures because it offers libraries like OpenCV. These libraries help in capturing, processing, and analysing images or video streams to detect hand movements. Python's simplicity allows for quick prototyping and integration with hardware interfaces, making it suitable for real-time applications. Hand gesture recognition in Python finds applications in various fields such as human-computer interaction and virtual reality interfaces. Its open-source nature encourages collaboration and innovation in gesture recognition research and development.

OpenCV

OpenCV stands for open-source computer vision. It is a machine learning based library and provides functionalities like image processing, object detection and feature extraction. OpenCV on Raspberry Pi leverages the Pi's hardware acceleration, optimizing performance for image processing tasks. Its compatibility with the Raspberry Pi camera module enables real-time computer vision applications. With GPIO integration, OpenCV can interact directly with sensors and actuators for IoT projects. Utilizing OpenCV on Raspberry Pi presents several advantages for computer vision projects. Firstly, OpenCV's extensive library provides a comprehensive suite of tools for tasks like image processing, object detection, and facial recognition, enabling developers to implement complex algorithms with ease. Secondly, Raspberry Pi's low-cost and compact form factor make it an accessible platform for experimenting with computer vision applications, democratizing the field for hobbyists and educators..

Text-to-Speech (TTS) libraries

Used to convert interpreted gestures into spoken words. Text-to-speech (TTS) libraries are essential tools that convert written text into spoken words, facilitating accessibility and enhancing user experiences. Leveraging advanced algorithms, these libraries generate natural-sounding speech across various languages and voices. Widely used in virtual assistants, navigation systems, and educational platforms, TTS technology enables seamless communication between humans and machines.

VNC Viewer

VNC Viewer is a widely-used remote desktop application that facilitates graphical access to remote systems. It's commonly employed with Raspberry Pi for remote desktop functionality. After installing VNC Server on the Raspberry Pi, users can connect via VNC Viewer by entering the Pi's IP address. This enables seamless control and interaction with the Raspberry Pi's graphical interface from a separate device. VNC Viewer simplifies tasks and troubleshooting by providing remote access to the Raspberry Pi's desktop environment.

Putty

Putty, a versatile open-source terminal emulator, is an essential tool for remote access and management, including Raspberry Pi. It provides a terminal interface for managing the Raspberry Pi OS from a Windows computer. Simply enter the Raspberry Pi's IP address and SSH port (default is 22) in Putty. Connect using SSH protocol, then log in with your Raspberry Pi's username and password. Putty streamlines remote administration and configuration tasks for Raspberry Pi projects.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.047 ISSN: 2581-4621

result

In this section, users can see their hand gestures translated into text output on screen. As users gesture, the system converts their actions into meaningful text feedback or commands. This feature enables applications like sign language recognition or gesture-based text input, offering users clear insight into how their gestures are interpreted.

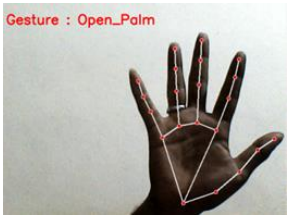

Figure 3: Open Palm

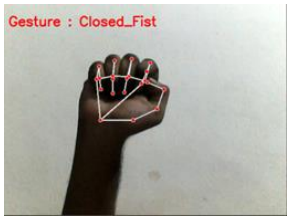

Figure 4: Closed Fist

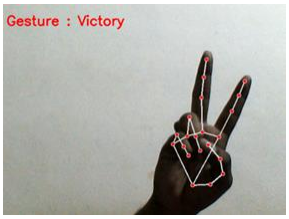

Figure 5: Victory

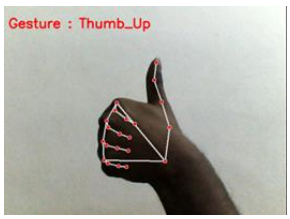

Figure 6: Thumbs Up

Figure 7: Thumb Down

conclusion

The gesture-based speaking system utilizing OpenCV and Raspberry Pi provides a novel approach to human-computer interaction. Through precise gesture recognition, it enables users to communicate effectively without conventional input devices. With ongoing refinement, this system holds promise for enhancing accessibility and convenience in diverse settings. Its affordability and adaptability make it a valuable tool for improving communication, particularly for individuals with disabilities. Continued development will unlock its full potential in revolutionizing how we interact with technology.

References

[1] S. Muthamilselvan, Himanshu Kumar, Vineela Kalluru, and Poorna Sabari Sundar.S,” Gesture Recognition using OpenCV in Raspberry Pi”, International Journal of Future Generation Communication and Networking Vol. 13, No. 3, (2020), pp. 486 – 489.

[2] Abdul Haris Rangkuti, Varyl HasbiAthala, Farrel Haridhi Indallah, Fajar Febriansyah,” Optimizing Hand Gesture Recognition Using CNN Model Supported by Raspberry pi for Self-Service Technology”**,** JOIV: Int. J. Inform. Visualization, 7(1) - March 2023 58-69.

[3] Dr. S. KoteswaraRao, B. Ravi Kumar, Manjunath B E,” Raspberry Pi Based Gesture Vocalizer for Dumb People Using OpenCV Python”**,** JOURNAL OF CRITICAL REVIEWS, VOL 9, ISSUE 05, 2022.

DOI:10.46647/ijetms.2024.v08i03.047 ISSN: 2581-4621

[4] A. Selvarani, Saaiyokesh.P, Muhilan.I, Jayakrishna.D," Hand Gesture Recognition Using Raspberry Pi and Open Cv", International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:06/June-2022.

[5] Aishwarya Pissay, Mahajan Saikumar, Kaushik Bandhu, M. Krishna Chaitanya**,** "Raspberry PI - Camera Based Digitzed Art Through Gesture Recognition", International Journal of Innovative Science and Research Technology, ISSN No: -2456-2165, Volume 4, Issue 4, April – 2019.

[6] Surya Narayan Sharma, Dr. A Rengarajan, "Hand Gesture Recognition using OpenCV and Python"**,** International Journal of Trend in Scientific Research and Development (IJTSRD), Volume 5 Issue 2, January-February 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470.