VOICE ASSISTANT SIGN LANGUAGE DETECTION USING MACHINE LEARNING

Project Reference No.: *47S_BE_1355*

College *: Jain Institute of Technology, Davanagere*
Branch *: Department of Electronics and Communication Engineering*
Guide(s) *: Ms. Deepthi G B*
Student(S) *: Mr. Mallikarjun M*
Ms. Prathibha M H
Ms. Sahana T
Mr. Varun K

Keywords:

Sign language, Hand gestures, Spectrum disorder, Deep learning technique, Convolutional Neural Network (CNN).

Introduction:

Sign language is crucial for communication within the deaf and hard of hearing community, facilitating interactions both within the community and with others. The language relies on hand gestures, categorized into static and dynamic gestures. While static gestures are easier to recognize, dynamic gestures present a greater challenge. To address this, researchers propose using deep learning techniques in computer vision for gesture recognition. Complex neural network architectures, designed to mimic the human brain, learn from vast datasets to emulate human cognitive processes. Among these architectures, Convolutional Neural Networks (CNNs) are particularly effective. With their convolutional layers and three-dimensional neuron arrangement, CNNs excel in image processing, making them adept at identifying objects and distinguishing between elements within images. This capability extends to tasks like computer vision and speech recognition. Utilizing CNNs for sign language recognition holds significant promise for advancing this technology, enhancing communication for the deaf and hard of hearing community.

Objectives:

- Enabling seamless communication between sign language users and voice-operated devices in real-time.
- Utilizing machine learning to integrate sign language education and awareness into voice assistants.
- Advocating for inclusive technology solutions for individuals with disabilities for effective communication.
- Provides wireless communication via extended Wi-Fi version for people with disabilities.
- Striving for high accuracy and precision in sign language detection to ensure dependable and effective communication between users and voice assistant systems.

Methodology:

Figure.1 depicts the block diagram of the proposed system architecture, comprises of several units like web camera, a Raspberry Pi 4 B+ model powered by a 5V DC power supply, a display unit (such as smartphones or laptops), a speaker, a relay driver circuit, and home appliances like lights and fans. The system operates using Real-time hand gestures from the individual with disabilities serve as input to the web camera. The camera captures these gestures, which are then transmitted to the Raspberry Pi 4 B+ model for the detection and interpretation of sign language gestures.

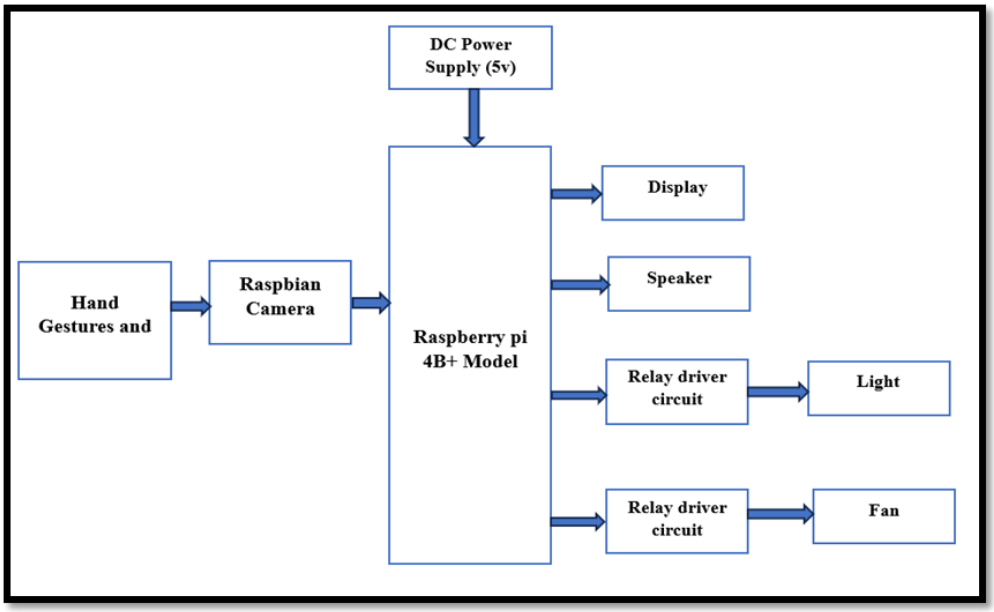
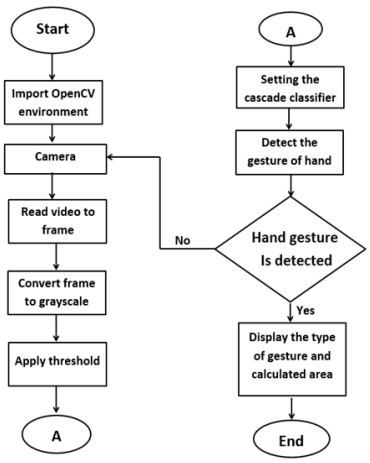


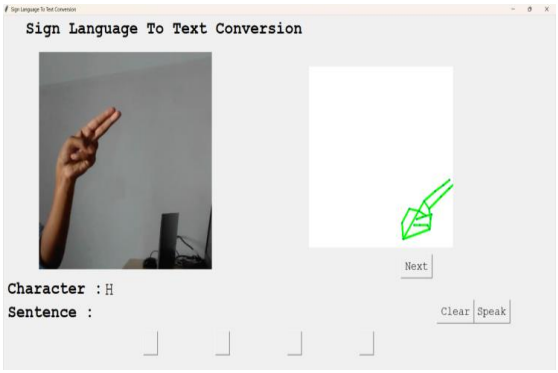
Figure.1. Block Diagram of Proposed System

Deep learning concepts such as computer vision and convolutional neural networks (CNN) techniques are employed for gesture recognition. Upon recognition, the interpreted hand gestures are displayed on the smartphone or laptop screen, and the system provides voice messages to the user. Additionally, specified hand gestures trigger the relay driver circuit, activating corresponding home appliances such as lights and fans, enhancing accessibility and convenience for individuals with disabilities.

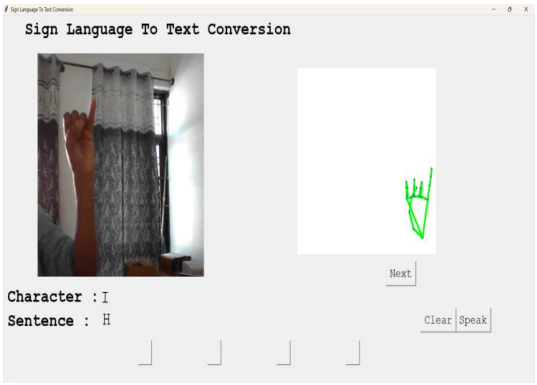


Step 1: The OpenCV environment is imported to access its functions and methods. OpenCV (Open-Source Computer Vision Library) is widely used for image processing, computer vision, and machine learning tasks.
Step 2: The camera is used to capture live video feed of the user's hand gestures. Individual frames are extracted from the live video stream. Each frame represents a snapshot of the scene captured by the camera.
Step 3: Each video frame is converted into grayscale to reduce the color data and simplify the image processing.
Step 4: Thresholding is a technique used to segment an image based on pixel intensity. It helps distinguish the hand gesture from the background. By setting a threshold value, pixels are classified as either foreground (hand gesture) or background.
Step 5: A cascade classifier is a machine learning-based object detection method used to set up to detect the hand gestures based on specific patterns or features.
Step 6: When a hand gesture is detected, it checks whether their exists an overlap hand gestures around the rectangle box or not. If no gesture is detected the process loops back to reading another video frame (point "A"), otherwise it displays the type of gesture and area of region of interest.
Step 7: The process ends by displaying the gesture on laptop or smartphone and through speaker via voice message to the user.

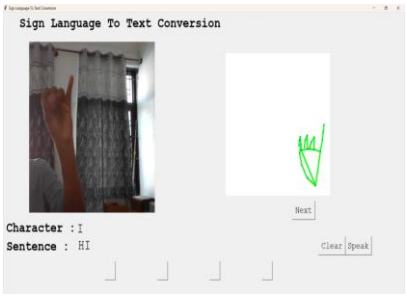
Conclusion:



Output window predicted Character "H"

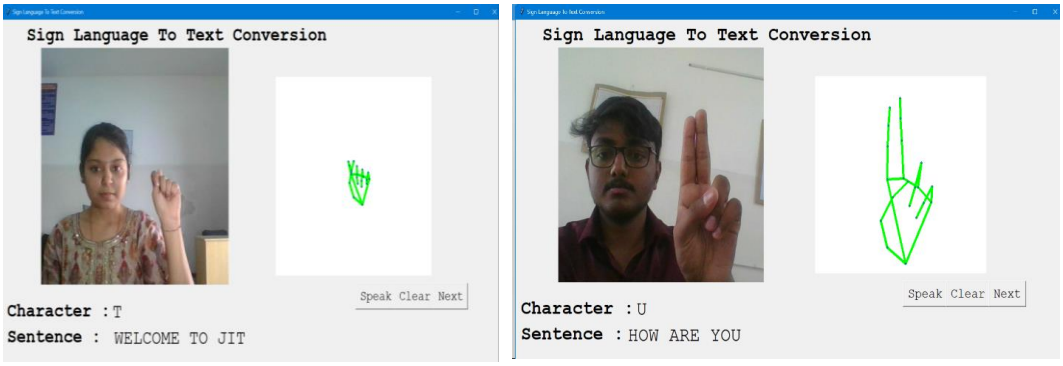


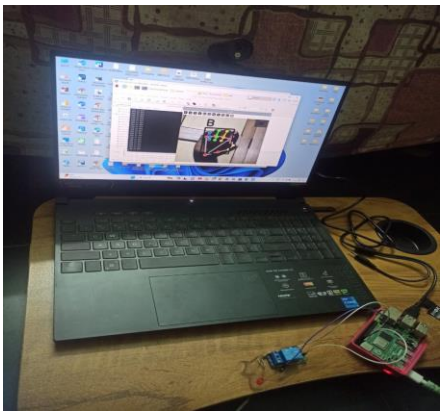
Next Character is predicted



Sentence Formation from the predicted Characters

Figure 3 illustrates the output window predicting the character "H" from the real time video frames captured by the user. After predicting "H," the user presses the "NEXT" button, causing the character "H" to appear in the sentence string. Subsequently, in Figure 4, the next character, "I," is predicted. Upon predicting "I," it is added to the sentence, forming a complete string, as demonstrated in Figure 5. After formation of sentence the “SPEAK” button is pressed the sentence formed can be heard through speaker via voice message.





In conclusion, sign language detection using machine learning, particularly Convolutional Neural Networks (CNNs) offers a powerful solution for bridging communication gaps by capturing and interpreting hand gestures from real-time video frames, this technology enables effective communication for individuals with hearing or speech disabilities. As these models continue to improve, voice assistants can play a pivotal role in making sign language more accessible and user-friendly. The predicted gestures after successful model training and deployment, includes accurate sign language recognition. Recognised hand gestures are displayed on smartphone or laptop also through the speaker via voice message to the user and includes real-time responsiveness and adaptability to user diversity.

Scope for future work:

1. The scope of the proposed system is to build a real time hand gesture classification system in complex backgrounds by various background subtraction algorithms. The accuracy of the system can be enhanced using extended features of Raspberry pi 5 B+ model. The advanced pre-processing

methods can be implemented to predict gestures in low light conditions. The Efficiency of the current project can be enhanced by using extended version of raspberry pi model to process huge datasets of real time hand gestures.

2. Further, the smartphone camera can be used to capture the hand gestures with supportable android version. The AI integrated machine learning algorithms can be implemented to ease the interface between the user and the device. The proposed work could be transformed into a compact product, facilitating remote accessibility for individuals with disabilities.

DESCRIPTION OF THE INNOVATION IN THE PROJECT

Voice Assistant Sign Language Detection Using Machine Learning achieves the detection of Real time hand gestures of disabled person. The detected hand gestures are converted into characters and then into the sentence formation using convolutional neural network. These gestures will be displayed using tkinter python interface model in the form of text messages and audio output via Laptop/Smartphone speaker and also controls the home appliances such as lights and fans using specified hand gestures.