WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Gesture Speak – Bridging the communication gap Through AI

Durgesh Bhusnale[1], Mahesh Marne[2], Prasad Das[3]

[1,2,3]Final Year Student, Department of IT, AISSMS Polytechnic, Pune, Maharashtra, India

ABSTRACT:

GestureSpeak represents a groundbreaking initiative conceived to address the pressing demand for comprehensive communication solutions for individuals without a voice. Leveraging sophisticated technologies including computer vision, machine learning, and web development, GestureSpeak operates with remarkable precision, swiftly identifying gestures and seamlessly translating them into articulate expressions. This pioneering endeavor endeavors to dismantle communication obstacles, enabling individuals to engage in profound dialogues and forge connections with the world at large. GestureSpeak is committed to inclusivity and accessibility, striving to ensure that every individual can actively participate in the global discourse. This endeavor stands as a testament to the transformative potential of technological advancements and underscores the profound impact of innovation in revolutionizing communication dynamics, thereby facilitating connections across diverse forms of human expression.

KEYWORDS: Computer Vision, Artificial Intelligence, Real Time, Gesture Recognition, Machine Learning, technology in education, User-Centric Design.

Introduction :

Communication stands as a fundamental human right, fostering understanding, collaboration, and connection. However, millions of individuals worldwide find themselves isolated due to their inability to express themselves in spoken language. This is especially true for those who rely on sign language as their primary mode of communication, facing significant challenges in conveying their thoughts and emotions to the broader community. GestureSpeak emerges as a beacon of hope and solidarity, offering innovative solutions to bridge the gap in non-verbal communication. At its core, GestureSpeak exemplifies the transformative power of technology in enhancing quality of life and fostering inclusivity. This vision drives the integration of cutting-edge technologies such as computer vision, machine learning, and web development to transform hand gestures into written and spoken language. The genesis of GestureSpeak lies in recognizing the lack of communication options for individuals with speech disabilities, particularly those dependent on languages often inaccessible to the wider public. Communication transcends borders, and the project's vision is to realize a web application that prioritizes user accessibility and real-time communication. GestureSpeak accurately interprets hand gestures through camera input, converting them into text messages displayed to both users and their conversation partners. Additionally, this text is seamlessly converted into speech, allowing the user's thoughts to be expressed and heard, ushering in a new era of interaction. Through meticulous planning and execution, GestureSpeak seamlessly integrates advanced technologies across data collection, preprocessing, machine learning model development, text-to-speech integration, web development, rigorous testing, and user experience refinement. Each phase of the project aims to deliver an accurate, user-friendly communication platform

Problem Statement :

The primary problem addressed by GestureSpeak is the limited communication options available to mute individuals, particularly those who use sign language. Current solutions are often limited in scope and accessibility, making it difficult for mute individuals to engage in meaningful conversations with the broader community. GestureSpeak aims to bridge this communication gap by providing a user-friendly, real-time hand gesture recognition system.There is a growing need for an online platform that goes beyond the norm, addressing the shortcomings of existing solutions. GestureSpeaks aims to bridge this gap by providing a space where users can foster authentic connections based on shared interests and proximity, setting a new standard for online relationship building.

Working :

Technologies Used:

- Python: For developing the recognition algorithm and backend server.
- Flask: For serving the webcam feed to a web interface.

- OpenCV: For capturing and processing video frames from the webcam.
- MediaPipe: For hand landmark detection and tracking.
- HTML/CSS/JavaScript: For designing and interacting with the web interface.

System Architecture:

The system consists of two main components: the backend recognition engine and the frontend web interface.

- **Backend Recognition Engine:**

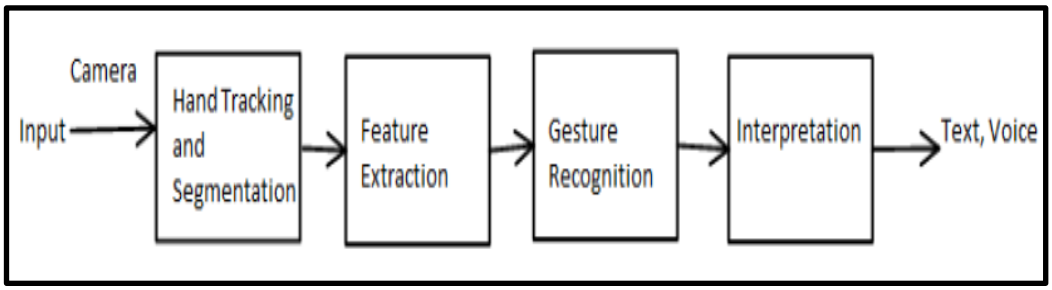
Implemented in Python, it uses OpenCV and MediaPipe libraries for capturing and processing video frames.
The hand landmarks are detected and tracked in real-time using the MediaPipe Hands model.
The detected hand gestures are fed into a pre-trained machine learning model (loaded from a pickle file) to recognize the corresponding alphabet.
The recognized alphabets are processed further to form words based on certain rules, such as a repeat gesture to re-enter the previous alphabet.

- **Frontend Web Interface:**

Developed using HTML, CSS, and JavaScript, it provides a user-friendly interface for displaying the webcam feed and recognized words.
Utilizes Flask, a Python web framework, to serve the webcam feed and communicate with the backend recognition engine.
Allows users to view the live webcam feed, see the recognized alphabets forming words, and interact with the system.

Working of the Project:

- Upon running the Flask server, the web interface becomes accessible through a web browser.
- The webcam feed is displayed on the webpage, allowing users to make hand gestures in front of the camera.
- The backend recognition engine processes each frame of the webcam feed, detecting and tracking hand landmarks.
- Hand gestures are translated into alphabets using a machine learning model, and recognized words are displayed on the webpage in real-time.
- Users can interact with the system by making different hand gestures, with the recognized words dynamically updating on the webpage.
- Certain gestures, such as repeating the previous alphabet, are supported to enhance user experience and input flexibility.



Outputs And Result

Website UI:

Gesture Recognition:

Outcome:

The project's output showcases a high level of accuracy in recognizing hand gesture alphabets, achieving an impressive accuracy rate of 98.03%. Leveraging machine learning algorithms and computer vision techniques, the system accurately detects and interprets hand movements, translating them into corresponding alphabetic characters in real-time. Users can seamlessly form sentences by sequentially gesturing individual alphabets, with the system dynamically updating the displayed text as each gesture is recognized. Additionally, the system offers an audio output feature, enabling users to hear the recognized text spoken aloud, enhancing accessibility and usability for individuals with visual impairments or those preferring auditory feedback. Through these capabilities, the project delivers a robust and user-friendly solution for hand gesture recognition and sentence formation, catering to a diverse range of users and use cases.

Conclusion :

There are several avenues for further enhancement and development of the project. One potential area of improvement is the expansion of the recognition capabilities to support more complex gestures and sign language. Additionally, integrating user authentication and personalization features could enhance the user experience and make the system more adaptable to individual preferences. Furthermore, optimizing the performance of the recognition algorithm and exploring real-world applications, such as assistive technology for individuals with disabilities, could broaden the project's impact and relevance. Continued research and development in the field of gesture recognition could unlock additional opportunities for innovation and refinement of the system.

Future Scope :

The project successfully demonstrates real-time hand gesture recognition using machine learning and provides a user-friendly interface for users to interact with the system. By integrating backend processing with a frontend web interface, the project enables intuitive interaction and feedback for users. Leveraging technologies such as Python, Flask, OpenCV, MediaPipe, and HTML/CSS/JavaScript, the system captures and processes video frames from a webcam, detects hand landmarks, and translates gestures into corresponding alphabets. Users can dynamically view recognized words as they form, enhancing the usability and effectiveness of the application.

REFERENCES :

1. Rautaray, Siddharth S., Real Time Hand Gesture Recognition System for Dynamic Applications (2012). International Journal of UbiComp (IJU), Vol.3, No.1, January 2012, Available at SSRN: https://ssrn.com/abstract=3702844
2. Rathi, Pulkit and Kuwar Gupta, Raj and Agarwal, Soumya and Shukla, Anupam, Sign Language RecognitionUsing ResNet50 Deep Neural Network Architecture (February 27, 2020). 5th International Conference on Next Generation Computing Technologies (NGCT-2019), Available at SSRN: https://ssrn.com/abstract=3545064 or http://dx.doi.org/10.2139/ssrn.3545064
3. Gayathri, A. and Sasi Kumar, A. and Sasi Kumar, A., Sign Language Recognition for Deaf and Dumb People Using Android Environment (January 1, 2017). International Journal Of Current Engineering And Scientific Research (IJCESR), Volume-4 Issue-5 2017, Available at SSRN: https://ssrn.com/abstract=3374759
4. S. Tornay, M. Razavi and M. Magimai.-Doss, "Towards Multilingual Sign Language Recognition," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 6309-6313, doi: 10.1109/ICASSP40776.2020.9054631. https://ieeexplore.ieee.org/document/9054631
5. Khati, Uday and Singh, Prajitesh and Shankar, Achyut, Text Generation through Hand
6. Gesture Recognition (June 23, 2020). Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020, Available at SSRN: https://ssrn.com/abstract=3633633 or http://dx.doi.org/10.2139/ssrn.3633633
7. Saini, Prashant and Bamane, Nikhil and Bhanpurwala, Mohammad and Hirlekar, Vaishali, Smart Gesture Recognition for Disabled People (June 26, 2020). Available at SSRN: https://ssrn.com/abstract=3696438 http://dx.doi.org/10.2139/ssrn.3696438
8. Saxena, Rohan and Garg, Romy and Gupta, Bhoomi and Kaur, Narinder, Hand Gesture Recognition (February 4, 2022). Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2022, Available at SSRN: https://ssrn.com/abstract=4027056 or http://dx.doi.org/10.2139/ssrn.4027056