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], Prasad Das[2], Mahesh Marne[3]

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

ABSTRACT:

GestureSpeak is a foundational project designed to meet the critical need for effective communication among the voiceless. Based on advanced technologies such as computer vision, machine learning, computer vision and web development, GestureSpeak instantly recognizes gestures and converts them into letters of communication and eloquence. This innovative solution aims to break down communication barriers by allowing speakers to have meaningful conversations and connect with the world. GestureSpeak aims to be inclusive and accessible so everyone can participate in the global conversation. This is a testament to the transformation of humanity's technological base and the power of innovation to disrupt communication and foster connection between different forms of human expression.

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

I. Introduction

Communication is a human right and the foundation of understanding, collaboration and connection. But for millions of people around the world who cannot express themselves in a language, the world can be a lonely place. Quiet people, especially those who rely on sign language as their main means of communication, often face great difficulties when trying to convey their thoughts, feelings, and emotions to the wider community. The GestureSpeak project is a beacon of hope and solidarity and offers new solutions to close the gap in non-voice communication. At its core, GestureSpeak embodies the transformative potential of technology to improve quality of life and create greater inclusion. This vision leads to the integration of technologies such as computer vision, machine learning, computer vision and web development, helping to transform hand gestures by recognizing and translating them into written form and ultimately into language listening. The origins of GestureSpeak can be traced back to the recognition that there is now a communication option for people with speech disabilities, particularly those who are dependent on languages, often fragmented, restricted or unavailable to the wider public. mass. Communication knows no boundaries. The vision of this project is to implement a web application that is a platform with user accessibility and real-time communication at its core. GestureSpeak captures and interprets hand gestures with unparalleled accuracy via camera input, converting them into text messages and displaying them to the user and their conversation partners. In addition, this text is converted into speech, enabling the user's thoughts to be expressed and heard, opening a new era in interaction.In this project, the combination of advanced technologies emerged as a carefully planned and followed method. Gesture Speak has gone through data collection, pre-processing, machine learning model development, text-to-speech integration, web development, rigorous testing, user experience ideas are shared and communicated. The overall goal of each phase is to create an accurate, easy-to-use and easy-to-use communication platform.

II. Literature Survey

1. This paper is "Real-Time Gesture Recognition System for Dynamic Applications" [2], which solves the communication problems faced by deaf-mute people. The system uses an Android app and uses technologies such as speech recognition, Video Relay Service (VRS) and sign translation. While Android applications support many features, Speech Sign technology effectively translates spoken words into the native language. VRS allows real-time communication between deaf and hearing people through sign interpreters. The system includes a sign language translator stored in a SQL Lite database to convert hand gestures into text and video. JSON is used to send data to create a standard format. This new approach takes an important step in communication by enabling hearing-impaired people to communicate seamlessly in their daily work without the need for a phone call. [one]
2. This paper is “Sign Language Recognition Using ResNet50 Deep Neural Network Architecture”[2]. The project aims to address communication barriers faced by the deaf-mute community by introducing an automated sign language recognition system. Focused on fingerspelling, the approach utilizes a novel 2-level ResNet50 neural network architecture. The American Sign Language Hand gesture dataset is augmented, and the model achieves an impressive 99.03% accuracy on 12,048 test images. The methodology involves a Level 1 model classifying images into sets, followed by a Level 2 model predicting the actual class within each set. This project uses deep learning and

adaptive learning to demonstrate the ability of computer vision to improve communication for the deaf, marking a major advance in space technology. [2]

3. This article is "Language recognition for deaf-mute users in the Android environment" [3]. This article helps deaf people use sign language to communicate with the rest of the world.Communication plays an important role for people. While Voice Sign Language technology and VRS enable sign language translations to be displayed on smartphones, the application has the ability to work on mobile devices without searching, with a method that translates spoken and written language with video technology. Interaction between sighted and blind people is quite difficult due to communication problems. There are many apps on the market that help blind people interact with the world. Voicebased email and chat machines can communicate with each other through the screen. This helps blind people interact with people. These activities include speech-based, text-based, and videobased interactive sessions. Video conferencing continues to evolve and may one day become the preferred method of mobile communication for the deaf. There is no technology that can solve the problem of hand-to-hand translation in daily life activities. Video interpreters help people who are deaf or hard of hearing understand what is being said in many situations. The main function of this work is that it can be used to learn sign language and provide video explanations for the hearing impaired.

4. [3]This article is "For More Information on Recommendations" [4]. The data show that hand information can be estimated by combining sources from different languages. Such capabilities for manually driven data modeling do not yet exist. In this paper, we develop a multilingual sign language approach in which hand gesture patterns are also obtained by combining hand gesture subunits using information independent of the target language. We confirm this claim with research on Swiss German Sign Language, German Sign Language and Turkish Sign Language and show that sign language knowledge can be effectively improved by using different sign languages. [4]

5. This article is "Text Generation via Gesture Recognition" [5]. This project uses deep learning models to analyze gestures and hand movements and accurately convert them into text or speech. This application includes a convolutional neural network (CNN) for image recognition and a convolutional neural network (RNN) for capturing the environment in sign language. Integration of a depth-sensing camera increases the system's accuracy in detecting nuances in detail. Additionally, natural language processing (NLP) is used to transform gestures into meaningful and coherent sentences. This comprehensive program provides effective and efficient communication for deaf and hard of hearing people. The project's innovative use of AI technologies holds significant promise for fostering inclusive communication in diverse environments.

6. This work proposes a new way to help people with physical disabilities (especially people who use wheelchairs) through intelligent behavior recognition [6]. This project focuses on solving the problems of wheelchair users by using radio frequency and Bluetooth technology for effective communication. The main innovation is in integrated gesture recognition, especially APDS 9960, which can be controlled by finger/hand gestures and blinking movements from the camera image network. Although the power chair based on the traditional power chair often causes stress in the user, the model has many ergonomic options. The system allows adaptation to different types of disabilities by dividing users into groups such as quadriplegia, paraplegia or hemiplegia. By converting the movement into a radio frequency or Bluetooth signal, the wheelchair can be guided, providing the user with greater flexibility and freedom. The proposed process demonstrates the best solutions that lead to success in technology services for individuals with disabilities.

7. The recommended method to attack the enemy is to reveal about 2500 palms and different gestures in the picture to break the aim. The motivation behind this project is to create an effective and efficient sign language system with the vision of promoting distance learning for deaf people worldwide. The authors compared their results with existing models such as K-search neighbor model (KNN) and hidden Markov model (HMM), focusing on in-depth investigation of CNNs. Python is used as the main language using libraries such as Tkinter GUI, NumPy and Pandas, while OpenCV is used for graphics. The authors discuss the complexity of the CNN architecture and highlight its role in processing neural structures in the human brain. Projects that reflect significant research in gesture recognition, including advances in hidden Markov models and the integration of Microsoft Kinect and CNN for language recognition. The article concludes by acknowledging its limitations, such as the need for different documentation, and looks at future developments, including the recognition of Indian Sign Language (ISL).

III. Proposed Methodology

1. Dataset Collection :

The basis of GestureSpeak is the variety and collection of gestures and phrases. Data collection is a careful process that is important for training and verification of accreditation standards.

- **Steps:**

1. Work with language experts, teachers and the community to get accurate information.
2. Use cameras and recording equipment to capture various movements made by people with different behaviors.
3. Make sure the file contains various meshes, lighting and complex movements.
4. Collect data with relevant tags that enable use in tracking model training.

2. Data Preprocessing:

Raw data needs to be pre-processed to ensure it is good, consistent and suitable for training. Prior knowledge is important to reduce and improve noise.

- **Steps:**

1. Use noise reduction techniques to remove negative objects and improve the quality of the data.
2. Improve dataset routing by making changes to routing rate, scale and spacing to make the model more robust.
3. Collect data including metadata, including move descriptions and difficulty levels, to support model training and evaluation.
4. Provide data sharing in training, validation and testing procedures to support model evaluation

3. Random Forest Classifier Model Development:

we will implement a Random Forest Classifier for gesture recognition. Random Forest is a powerful algorithm for classification tasks and can be adapted for image recognition.

- **Steps:**

1. Research and select the Random Forest Classifier implementation suitable for image classification tasks.
2. Preprocess the image data to make it compatible with the Random Forest Classifier input requirements.
3. Train the Random Forest Classifier on the preprocessed image data, tuning hyperparameters such as the number of trees and tree depth for optimal performance.
4. Evaluate the trained model's performance using metrics like accuracy, precision, and recall.
5. Fine-tune the model if necessary, adjusting parameters based on performance metrics.

4. Text and Audio Generation:

We will use the Random Forest Classifier's output to generate text, and for audio generation, we'll use techniques like speech synthesis.

Text Generation:

1. Develop algorithms to map recognized gestures to structured sentences using the output from the Random Forest Classifier.
2. Ensure that the generated text maintains coherence, syntax, semantics, and context relevant to the recognized gestures.

Audio generation:

3. Utilize text-to-speech (TTS) libraries or APIs (such as Google Text-to-Speech or Amazon Polly) to convert the generated text into audio.
4. Implement audio synthesis techniques to enhance the naturalness and quality of the generated speech.

5. Web Devlopment:

GestureSpeak's user interface is used by web developers and provides real-time communication opportunities.

- **Steps:**

1. Choose a web development framework like Flask to create a good user interface.
2. Create a web application that allows users to input movements in real time via the camera.
3. Use available feedback techniques, including visual enhancements and video text translations, to improve user experience.
4. Create a user-friendly experience, including intuitive gesture control and clear user instructions.

6. Testing and Validation:

Rigorous testing is essential to ensure the robustness and reliability of GestureSpeak in a variety of situations.

- **Steps:**

1. Successfully use test cases with different lighting, meshes and complex movements.
2. Create cases that simulate real-world interactions to test the effectiveness of your model.
3. Perform a robust analysis to identify challenges or limitations that may impact the fidelity or performance of the project.

7. Deployment:

Distribution of GestureSpeak is intended to be publicly available and marks the transition from development to actual use.

- **Steps:**

1. Deploy the web application to a production server or hosting platform to ensure capacity and reliability.
2. Set up the server environment to handle user requests and instant messages.
3. Adopt security practices including encryption and secure data handling to protect customer privacy.
4. Create a user-friendly experience, including intuitive gesture control and clear user instructions.

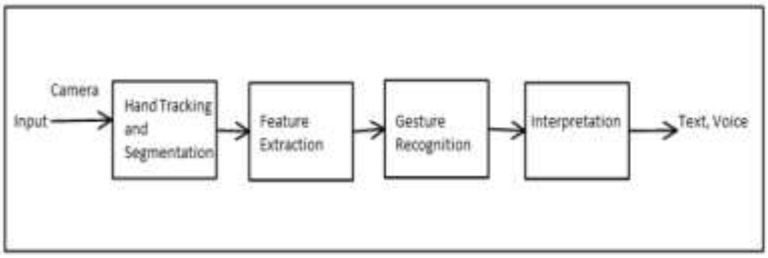

Fig.1.Working Of System

IV. Conclusion and Future Scope

The project "Gesture Speak – Bridging the Communication Gap Through AI" represents a significant stride towards fostering inclusive communication and breaking down barriers in diverse educational settings. By harnessing the power of artificial intelligence, specifically through gesture recognition technology, the project endeavors to enhance communication accessibility for individuals facing language and communication challenges.

The integration of AI-driven gesture recognition into educational environments holds the promise of revolutionizing how students and educators interact, particularly for those with communication difficulties. The project's commitment to inclusivity and user-friendliness underscores its potential to create a more equitable learning experience.

The Gesture Speak system, with its innovative approach, allows users to express themselves through intuitive gestures, providing an alternative means of communication beyond traditional spoken or written language. The emphasis on user-friendly interfaces ensures that educators and students can seamlessly incorporate this technology into their daily interactions, promoting a more inclusive and supportive educational environment.

As we anticipate the successful implementation of Gesture Speak, it is poised to bring about a transformative shift in communication dynamics within educational settings. The future scope of this project includes further refinement of gesture recognition algorithms, exploration of multi-modal communication options, and potential integration with other assistive technologies. The continuous incorporation of user feedback will be instrumental in refining and expanding the system's capabilities.

Moreover, the project acknowledges the importance of compliance with data privacy regulations, ensuring ethical and legal standing in the development and deployment of the Gesture Speak system. Ongoing efforts to address security concerns and scalability for various educational institutions will contribute to the project's long-term success and widespread adoption.

In conclusion, Gesture Speak has the potential to become an indispensable tool in educational technology, transcending communication barriers and fostering a more inclusive learning environment. By bridging the communication gap through AI, this project stands as a testament to the transformative power of innovative technology in promoting accessibility, equity, and engagement within educational spaces.

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 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
6. 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
7. 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