

Heart2Hear: Predictive Communication Bridge through Voice and Vision

**Mrs Bijili Sreelatha¹, Veerabrahmam Prasanna², Gunji Yeshwanth Akhil Sai³, Dheeravath Bhoomika⁴,
Vudgula Sai Prabhas⁵**

Associate Professor, Dept. of CSE-Data Science, ACE Engineering College, Hyderabad, Telangana, India¹

IV B.Tech Students, Dept. of CSE-Data Science, ACE Engineering College, Hyderabad, Telangana, India^{2,3,4,5}

ABSTRACT: Communication is one of the most fundamental parts of human connection, but for people with speech and hearing disabilities, real-time interaction with others often becomes something very difficult in their lives. To break this barrier, the project as mentioned, “AI-Powered Predictive Communication Bridge for Deaf and Hearing People using Live Video Analytics” proposed developing an intelligent two-way communication system, i.e., a bridge that allows deaf and hearing people to talk naturally via a secure online video call.

The system fuses predictive analytics, computer vision and speech processing algorithms to translate gestures into speech and speech into animated signs in real time. Using MediaPipe and CNN models, the gestures performed by the deaf user are predicted instantly and converted into text and speech for the normal user. Simultaneously, the hearing person’s speech is recognized using speech-to-text technology and transformed into animated sign gestures visible to the deaf user. Additional modules such as emotion detection, voice-tone support, and a smart analytics dashboard enhance both accuracy and user experience.

This predictive AI-based communication bridge not only ensures real-time, secure, and adaptive interaction but also adds an element of emotional intelligence to make conversations more natural. By combining data science algorithms with WebRTC-based peer-to-peer communication, the system provides a cost-effective, inclusive, and scalable solution that empowers people with hearing and speech impairments to communicate seamlessly with the hearing world.

I. INTRODUCTION

Communication is one of the most important aspects of human interaction. It allows individuals to share ideas, emotions, and information effectively. However, for people who are deaf or speech-impaired, communication with others can be challenging because most daily interactions rely heavily on spoken language. Deaf individuals often depend on sign language to express themselves, but many people who do not have hearing impairments are not familiar with sign language. This communication gap can create barriers in education, employment, healthcare, and social environments.

The Heart2Hear communication system is designed to address this challenge by providing a real-time web-based communication platform that facilitates interaction between deaf individuals and normal users. The system uses video communication, speech-to-text conversion, and gesture recognition technologies to create a two-way communication channel. Through this platform, a normal user can speak naturally, and the system converts the speech into text that can be understood by a deaf user. Similarly, the deaf user can communicate through gestures that are recognized and translated into text messages for the normal user.

The proposed system integrates several modern technologies, including WebRTC for real-time video communication, Socket.IO for signaling and data exchange, and machine learning-based gesture recognition techniques. These technologies work together to create an accessible and user-friendly communication environment that supports real-time interaction between users.

II. LITERATURE SURVEY

1. Alaghband et al. (2023) studied different methods used for sign language recognition. Their research mainly reviewed various translation techniques, datasets, and recognition approaches used in previous systems. However, their work was mostly theoretical and did not provide a real-time implementation.
2. Trabelsi et al. (2025) proposed a vision-based sign language recognition system using deep neural network models. The system was capable of recognizing gestures using computer vision techniques, but it required high GPU resources, which makes it difficult to use on low-end devices.
3. Madhiarasan et al. (2022) introduced a dynamic sign recognition system based on CNN-RNN deep learning models. Their method improved gesture recognition performance, but the system showed lower accuracy in poor lighting conditions or when there was too much motion in the scene.
4. Shinde and Kagalakar (2015) developed a Marathi sign-to-text and text-to-sign translation system. While it helped improve communication for Marathi users, the system was limited to a single language and lacked advanced features such as speech recognition or AI-based interaction.

Conclusion:

From the literature review, it is observed that most existing systems focus mainly on one-way communication and lack real-time interaction. The proposed system aims to overcome these limitations by enabling two-way real-time communication, gesture learning, and multi-language support to make the system more practical and user-friendly.

III. METHODOLOGY

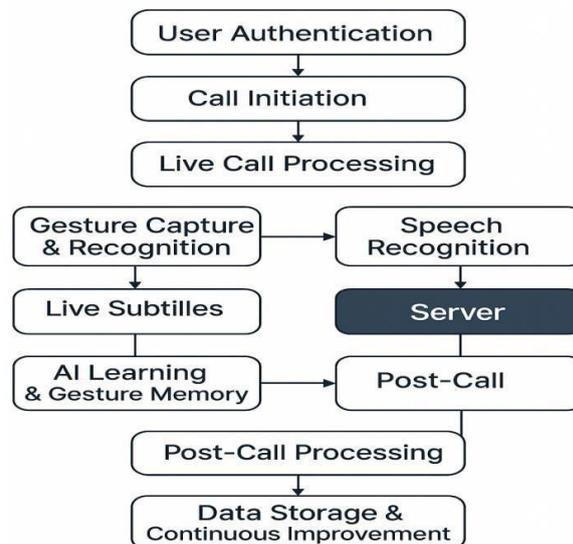


Fig 1: Methodology

The proposed system enables communication between deaf users and normal users through real-time video calling combined with gesture and speech recognition technologies.

First, users log into the system through a secure authentication process, where each user is assigned a unique ID. After logging in, a user can initiate a call by entering the ID of another user. The system then establishes a real-time video connection between both participants using WebRTC technology.

During the live call, two main processes run simultaneously. For the deaf user, the system captures hand gestures

through the camera and processes them using gesture recognition techniques, converting them into readable text. For the normal user, the system captures spoken input through the microphone and converts it into text using speech recognition, which appears as live subtitles for the deaf user.

A central server module manages communication, signaling, and data exchange between both users to ensure smooth interaction. In addition, the system includes an AI learning mechanism that allows new gestures to be stored and recognized in future sessions.

After the call ends, the system performs post-call processing, where interaction data such as gestures and speech inputs are analyzed and stored. This data helps improve the system's accuracy and supports continuous learning and performance enhancement over time.

IV. SYSTEM ARCHITECTURE

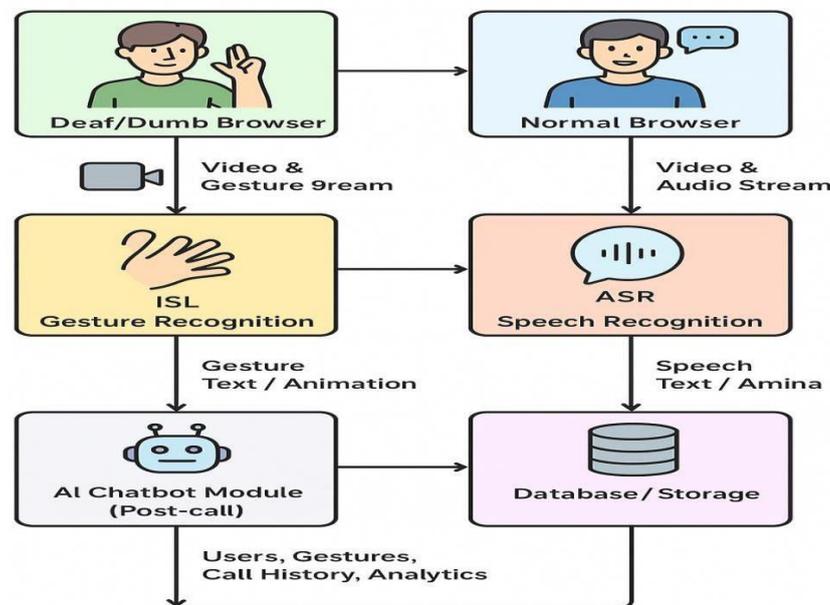


Fig 2: System Architecture

The proposed system architecture is designed to enable smooth communication between deaf users and normal users through a web-based platform. The system consists of two main user interfaces: the Deaf/Dumb user browser and the Normal user browser. Both users connect through a real-time video communication system, allowing them to interact during a live call.

On the deaf user side, the system captures the video stream and hand gestures through the device camera. These gestures are processed using an Indian Sign Language (ISL) gesture recognition module, which analyzes the hand movements and converts them into readable text or animations. This helps the normal user understand the message communicated through gestures.

On the normal user side, the system captures the video and audio stream. The spoken words are processed using an Automatic Speech Recognition (ASR) module, which converts speech into text. The generated text is displayed to the deaf user so that they can understand the conversation easily.

The system also includes an AI chatbot module, which operates mainly after the call. This module helps in learning new gestures, improving recognition accuracy, and providing intelligent feedback for future interactions.

All communication data such as user information, recognized gestures, call history, and analytics are stored in a central database or storage system. This data storage supports system improvement, performance analysis, and continuous learning.

Overall, the system architecture integrates video communication, gesture recognition, speech recognition, AI learning, and data storage modules to create an efficient and user-friendly communication platform for deaf and normal users.

V. OUTPUT SCREENS

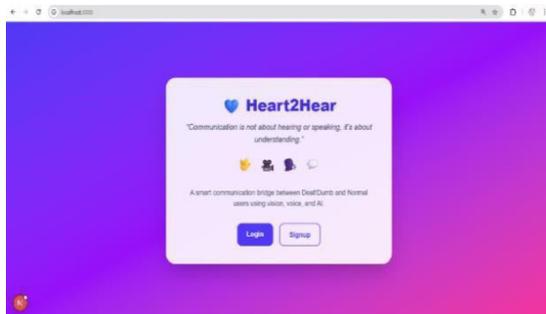


Fig 3. Home Page

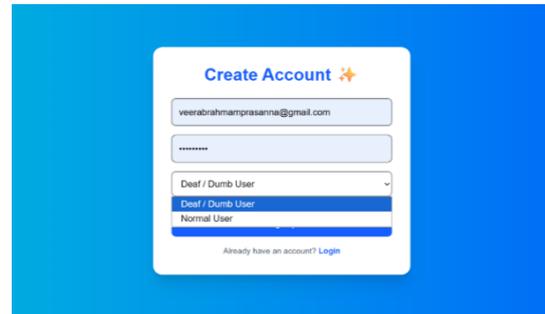


Fig 4. Create Account

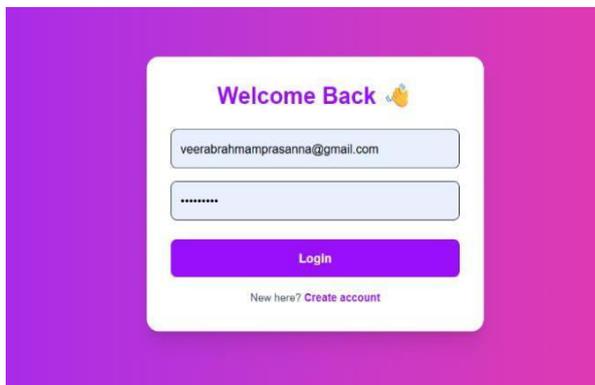


Fig 5. Login Page

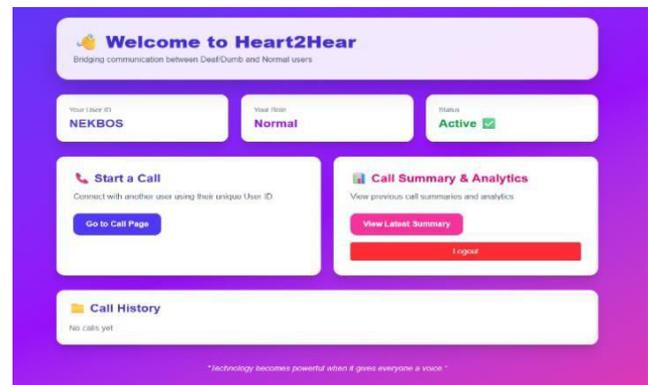


Fig 6. Normal User Dashboard

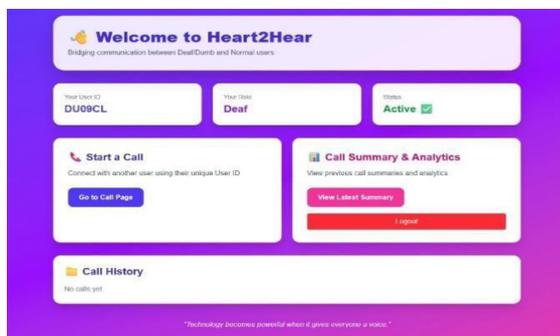


Fig 7. Deaf/Dumb User Dashboard



Fig 9. Call Page

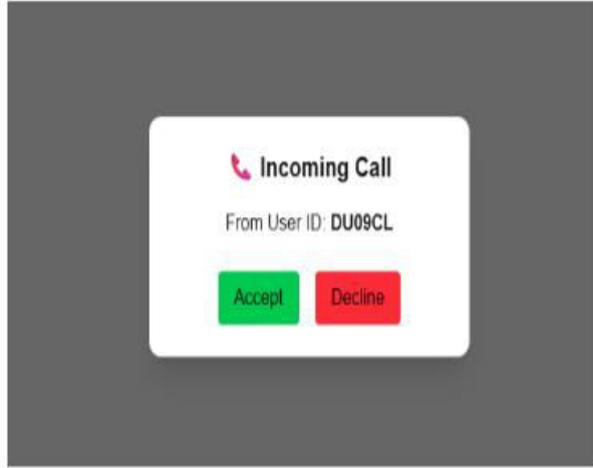


Fig 10. Incoming Call Modal



Fig 11. Normal User Panel (Gesture 1)



Fig 12. Deaf User Panel (Animated Gif 1)



Fig 13. Deaf User Panel (Call-Summary 1)



Fig 14. Deaf User Panel (Call-Summary 1)

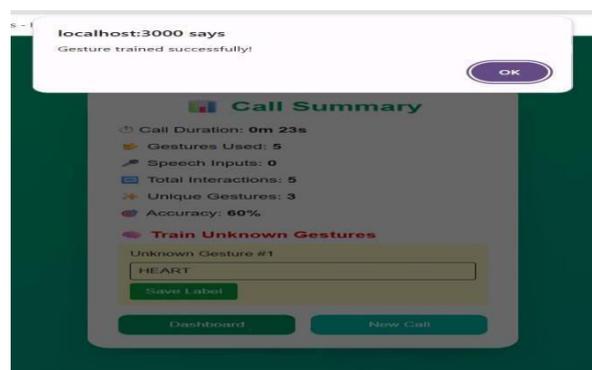


Fig 15. Deaf User Panel (Call-Summary 3)

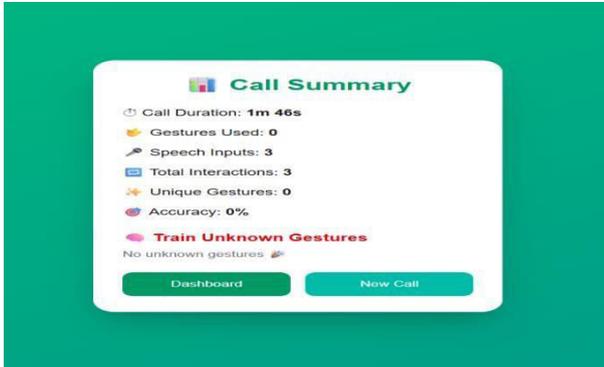


Fig 16. Normal User Panel (Call-Summary 1)

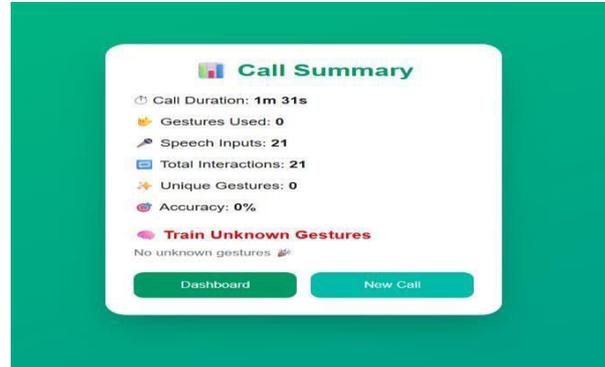


Fig 17. Another Call history (Call-Summary 1)

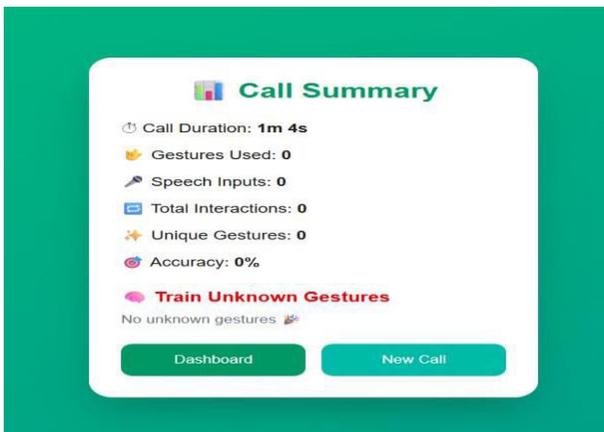


Fig 18. Another Call history (Call-Summary 2)



Fig 19. Main Call History Page on Dashboard

VI. WORK FLOW

The workflow of the proposed system begins with **user authentication**, where both the deaf user and the normal user register or log in using their email credentials. After successful authentication, each user is assigned a **unique ID**, which allows them to securely access the platform and connect with other users.

Once logged in, a user can **initiate a call** by entering the unique ID of the other participant. The system then establishes a **real-time communication connection using WebRTC**, enabling live video and audio streaming between the two users.

During the live call, the system processes communication in two directions. When the **deaf user communicates through gestures**, the webcam captures the hand movements and the **ISL-based gesture recognition module** converts them into text. This text is displayed on the normal user's screen, and it can also be converted into voice using a text-to-speech feature.

On the other hand, when the **normal user speaks**, the system captures the speech and uses **speech recognition technology** to convert it into text. The text is then displayed to the deaf user, and in some cases, it can also be represented through animated sign language visuals.

The system also includes an **AI learning mechanism** that stores unrecognized gestures in a gesture memory log. These gestures can later be labeled by users, allowing the system to learn and improve its recognition accuracy over time.

After the call ends, the system performs **post-call processing**, where it generates summaries, analytics, and feedback based on the interaction. All relevant data, including user information, call history, gesture logs, and analytics, are stored in a database to support continuous improvement of the system.

VII. CONCLUSION AND FUTURE SCOPE

Conclusion

This work presents a communication system designed to bridge the gap between deaf or speech-impaired individuals and normal users. In many real-life situations, communication becomes difficult because deaf users rely on sign language while most people communicate through speech. The proposed system attempts to solve this problem by combining real-time video communication with gesture recognition and speech recognition technologies.

Through this system, deaf users can communicate using hand gestures that are captured through a webcam and interpreted using an Indian Sign Language (ISL) recognition module. These gestures are converted into readable text and can also be transformed into voice output so that normal users can easily understand the message. At the same time, speech from the normal user is processed through a speech recognition module and converted into text or sign animations so that the deaf user can follow the conversation.

Another important aspect of the system is the inclusion of AI-based learning and gesture memory. When the system encounters gestures that are not recognized, they are stored and later labeled by users. This allows the system to gradually improve its gesture recognition accuracy over time. Additionally, the system records interaction data such as gesture usage, recognition accuracy, and call summaries, which can be used for further analysis and improvement.

Overall, the system demonstrates how modern technologies such as WebRTC, artificial intelligence, and computer vision can be integrated to build an accessible communication platform. The developed system not only supports real-time interaction but also provides learning and analytics features that make it more adaptive and user-centric.

Future Scope

Although the proposed system successfully enables communication between deaf and normal users, there are several opportunities for further improvement. One important enhancement would be expanding the gesture recognition dataset to include a larger number of gestures from different sign languages. This would improve recognition accuracy and allow the system to support users from diverse linguistic backgrounds.

Future versions of the system could also integrate more advanced deep learning models to improve gesture detection in challenging conditions such as poor lighting or fast hand movements. Improving the robustness of the recognition model will make the system more reliable in real-world environments.

Another potential improvement is the development of a mobile application version of the platform. By extending the system to smartphones and tablets, users will be able to communicate anytime and anywhere, making the solution more practical for daily use.

The system can also be enhanced by introducing multilingual speech translation, which would allow spoken language to be automatically translated into multiple regional languages. This feature would make the platform more inclusive and accessible to a wider user base.

VIII. ACKNOWLEDGEMENT

We express our sincere gratitude to all who supported us throughout the development of this project. We are especially thankful to Prof. Y.V. Gopala Krishna Murthy, General Secretary, and Mrs. M. Padmavathi, Joint Secretary, for providing us the opportunity and environment to carry out this work. Our heartfelt thanks to Dr. R. Suresh, Head of the Department, for his guidance and encouragement. We are deeply grateful to our internal guide Mrs. B. Sreelatha, and project coordinator Mrs. B. Saritha, for their consistent support, valuable feedback, and motivation throughout the project.

Lastly, we thank all the faculty members, staff for their constant encouragement and support.

REFERENCES

1. [Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids | IEEE Journals & Magazine](#) | IEEE Xplore
https://www.researchgate.net/publication/277906798_Advanced_Marathi_Sign_Language_Recogniti
2. [on using Computer Vision?utm_source=chatgpt.com](#)
https://www.sciencedirect.com/science/article/abs/pii/S1566253525006980?utm_source=chatgpt.com
3. [Google Cloud Text-to-Speech API \(2024\). Speech Synthesis and Voice Modulation Techniques.](#)
<https://cloud.google.com/text-to-speech>
4. [Inspired TTS integration for converting recognized gestures into spoken responses. Python Software Foundation \(2024\). Flask Documentation.](#) <https://flask.palletsprojects.com/>
5. [Used for backend development and API routing between frontend and AI processing modules. Node.js Foundation \(2024\). Node.js and Express Framework Documentation.](#) <https://nodejs.org/>
6. [Served as the frontend server for handling authentication, sessions, and user requests. W3C Consortium \(2023\). WebRTC Technical Overview.](#) <https://webrtc.org/>
7. [Used for establishing real-time audio and video communication between users. SQLite Consortium \(2024\). SQLite Database Documentation.](#) <https://www.sqlite.org/docs.html>
8. [WebRTC Project \(2024\). Web Real-Time Communication \(WebRTC\) Official Documentation.](#)
9. [World Wide Web Consortium \(W3C\) and Google Developers. Available at: https://webrtc.org/](#)



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details