



(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

⊕ www.ijirset.com 🖂 ijirset@gmail.com 🖄 +91-9940572462 🕓 +91 63819 07438





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405175|

AI Assistive Live Session Transcription for Hearing and Speech Impaired

Anshad PE¹, Abdul Farhan N², Suhail A³, Vipin Krishna K⁴, Dr.Nathan B⁵

UG Scholar, Department of Artificial Intelligence and Data Science, Dhaanish Ahmed Institute of Technology,

Coimbatore, India¹

UG Scholar, Department of Artificial Intelligence and Data Science, Dhaanish Ahmed Institute of Technology,

Coimbatore, India²

UG Scholar, Department of Artificial Intelligence and Data Science, Dhaanish Ahmed Institute of Technology,

Coimbatore, India³

UG Scholar, Department of Artificial Intelligence and Data Science, Dhaanish Ahmed Institute of Technology,

Coimbatore, India⁴

Associate Professor, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology,

Coimbatore, India⁵

ABSTRACT: Effective communication is vital for sharing information, ideas, and emotions. However, virtual meeting platforms like Zoom, Microsoft Teams, and Google Meet often fail to accommodate individuals with hearing and speech impairments, creating significant barriers to inclusivity. While sign language offers a means of communication for deaf individuals, its interpretation remains challenging for non-signers. Existing sign language recognition technologies are limited in accuracy and accessibility, making them unsuitable for seamless integration into virtual platforms. This project introduces an AI-driven system designed to bridge the communication gap between deaf and hearing participants in virtual meetings. The system employs advancements in deep learning, particularly Temporal Convolutional Networks (TCNs), to enable two-way communication in real-time. It includes three core modules: a Sign Recognition Module (SRM) that interprets signs using TCN, a Speech Recognition and Synthesis Module (SRSM) powered by Hidden Markov Models, which converts spoken language into text, and an Avatar Module (AM) that visually translates speech into corresponding signs. The Avatar Module is essential for visually representing spoken language in sign language format, ensuring non-signers can effectively communicate with sign language users in an intuitive and engaging way. Trained on Indian Sign Language, the system facilitates communication across diverse groups, including deaf, mute, hard-of-hearing, visually impaired, and non-signers. Its integration into popular virtual meeting platforms through a user-friendly web-based interface enhances accessibility and participation. This solution represents a significant advancement in fostering inclusivity and accessibility in virtual meeting environments.

KEYWORDS: Sign Language Recognition, Temporal Convolution Network (TCNs), Avatar Module, Virtual Platform, Real Time Communication, AI-Driven Systems.

I. INTRODUCTION

The Mind IT Solution is one of the few IT system integration, professional service and software development companies in Macedonia that works with Enterprise systems and companies which has sister concern in Trichy. As a privately owned company, The Mind IT Solution provides IT Consultancy, software design and development as well as professional services and hardware deployment and maintenance to the following verticals: Government (Local and Central), Financial Services (insurance, banking and clearing house), Telecommunications, Energy and Utilities, Health Care and Education





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405175|

II. RELATED WORK

Existing virtual communication platforms such as Zoom, Microsoft Teams, and Google Meet have significantly advanced digital collaboration. However, they lack integrated real-time sign language interpretation, leaving individuals with hearing and speech impairments at a disadvantage. Several approaches and technologies have attempted to bridge this accessibility gap with varying degrees of success.

1. Assistive Technologies: Traditional devices like hearing aids and cochlear implants enhance auditory perception but fall short in addressing the needs of sign language users. Manual methods, such as pen-and-paper communication, are impractical for dynamic and fast-paced virtual interactions.

2. Sign Language Interpreters:Human interpreters have long been used to facilitate communication between deaf and hearing individuals. However, their availability, cost, and the requirement for prior arrangement limit their use in spontaneous and scalable digital interactions.

3. Computational Approaches:Support Vector Machines (SVMs): Used for static sign recognition based on image features, but they struggle with large datasets and continuous sign sequences, limiting real-time application.

Ensemble Learning Models: Techniques like Random Forests and Gradient Boosting improve classification accuracy by aggregating multiple learners. However, they demand high computational resources, reducing practicality for real-time deployments.

Gesture Recognition by Learning from Poses (GRBP): This approach extracts skeletal pose features for gesture recognition. While accurate, it relies on high-quality skeletal data, which is not always feasible in webcam-based systems.

4. Deep Learning Innovations:Recent studies have explored CNNs, LSTM, and TCN architectures to improve recognition accuracy and sequence modeling. For example, Temporal Convolutional Networks (TCNs) have shown superior performance over RNNs and LSTMs in handling long-range dependencies with lower memory consumption and higher parallelization.

III. METHODOLOGY

The proposed system, **Sign Meet**, utilizes a multi-modal AI framework designed to enable inclusive communication between deaf and non-deaf users within virtual meeting environments. It integrates computer vision, deep learning, natural language processing, and real-time video conferencing tools to facilitate sign-to-speech and speech-to-sign translation. The methodology is divided into the following components:

- Web-Based Virtual Communication Platform: The Sign Meet system is a web application built using Python (Flask), MySQL, and Jitsi for video conferencing. It supports multiple user roles including administrators, deaf users, and non-deaf users. The platform enables real-time video communication with integrated AI-powered modules for sign and speech translation.
- Sign Recognition Module (SRM): This module uses Temporal Convolutional Networks (TCNs) to recognize sign language gestures from video input. Preprocessing steps include grayscale conversion, Gabor filtering for noise reduction, and Region Proposal Network (RPN)-based segmentation to isolate gestures. The module extracts features using Fully Connected Layers (FCLs) and classifies signs using Convolutional Neural Networks (CNNs).
- Speech Recognition and Synthesis Module (SRSM):Utilizing Hidden Markov Models (HMMs), this module converts spoken language to text and synthesizes speech from textual inputs. Speech preprocessing involves Wavelet Transform and feature extraction via Mel-Frequency Cepstral Coefficients (MFCC). This enables accurate interpretation of non-deaf users' speech for conversion into sign language.





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405175|

- Avatar Module (AM): An AI-generated avatar animates recognized signs for visual interpretation. This component ensures intuitive understanding by deaf users by dynamically converting text or speech into corresponding sign gestures using a 3D animated model.
- **Multilingual Interpretation**: The system includes multilingual translation of recognized signs and speech using NLP techniques. Users can set language preferences, allowing translated outputs (text and sign) to be presented in multiple languages, enhancing accessibility.
- Integration and Deployment: The system is deployed with Jitsi integration for live meetings. The trained SignNet model is embedded in the backend, allowing seamless real-time processing of gestures and speech. Notifications and meeting links are managed via SMS and email alerts.

IV. EXPERIMENTAL RESULTS

The implementation of the Sign Meet system was evaluated through comprehensive testing across multiple user roles and interaction scenarios in virtual environments. The system successfully enabled seamless communication between deaf and non-deaf participants through real-time sign language recognition, speech synthesis, and avatar-based interpretation.







www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

<section-header>

V. CONCLUSION

The Sign Meet platform demonstrates how artificial intelligence can be effectively harnessed to create inclusive digital communication environments for individuals with hearing and speech impairments. By integrating advanced modules such as Temporal Convolutional Networks (TCNs) for sign recognition, Hidden Markov Models (HMMs) for speech processing, and real-time avatar animation, the system enables seamless, two-way interaction between deaf and hearing participants in virtual meetings.

The platform successfully addresses the limitations of traditional assistive methods and human interpreters by offering real-time translation, visual feedback, and multi-language support. Through integration with Jitsi and a user-friendly web interface, Sign Meet ensures accessibility, scalability, and ease of adoption across diverse settings, including education, remote work, and public communication.

In conclusion, Sign Meet represents a significant advancement toward inclusive virtual communication, reducing digital barriers and empowering the hearing and speech-impaired community to engage confidently and independently in the digital age.

REFERENCES

1. Gadekallu, T. R., Srivastava, G., & Liyanage, M. (2022). Hand gesture recognition based on a Harris hawks optimized convolution neural network. Computers & Electrical Engineering, 100, Article ID 107836.

2. Chowdhary, G. T. R. C. L., & Parameshachari, B. D. (2022). Computer Vision and Recognition Systems: Research Innovations and Trends. CRC Press.

3. Riaz, M. M., & Zhang, Z. (2021). Surface EMG Real-Time Chinese Language Recognition Using Artificial Neural Networks. In Intelligent Life System Modelling Image Processing and Analysis (Vol. 1467). Springer.

4. Halvardsson, G., Peterson, J., Soto-Valero, C., & Baudry, B. (2021). Interpretation of Swedish sign language using convolutional neural networks and transfer learning. SN Computer Science, 2(3), 1–15.

5. Likhar, P., Bhagat, N. K., & R. G. N. (2020). Deep Learning Methods for Indian Sign Language Recognition. 2020 IEEE 10th International Conference on Consumer Electronics (ICCE-Berlin), 1–6.

6. Li, F., Shirahama, K., Nisar, M. A., Huang, X., & Grzegorzek, M. (2020). Deep Transfer Learning for Time Series Data Based on Sensor Modality Classification. Sensors, 31(20), 4271.

7. Bird, J. J., Ekárt, A., & Faria, D. R. (2020). British sign language recognition via late fusion of computer vision and leap motion with transfer learning to American sign language. Sensors, 20(18), 5151.









INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com