International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 563 - 569
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0079

SignSpeak: Sign Language-To-Speech and Speech-To-Sign Language

Rashmi Maheshwari[1], Khushi Anand Singh[2], Shashank Jayshankar Tiwari[3], Anuj Mukesh Mishra[4]
[1]Assistant Professor, Department of ECS, Shree L.R. Tiwari College of Engg., Mumbai, Maharashtra, India.
[2,3,4]UG Scholar, Department of ECS, Shree L.R. Tiwari College of Engg., Mumbai, Maharashtra, India.
Emails: *rashmi.maheshwari@slrtce.in[1], khushi.a.singh@slrtce.in[2], shashank.j.tiwari@slrtce.in[3], anuj.m.mishra@slrtce.in[4]*

Abstract

Exchange of words among the community is one of the essential mediums of survival. Developing a vision-based application that translates sign language into intelligible speech and vice versa is the main goal of our suggested technique [1]. The objective is to establish a link between the hearing and hearing-impaired communities and begin two-way conversation[2].This application presents the Android program that allows blind, deaf, and dumb persons to converse with one another. Utilizing technologies such as TensorFlow Lite for gesture recognition, Flutter for mobile development, and Firebase for backend support, the application guarantees flexibility and inclusivity. CNNs and natural language processing (NLP) improve speech and gesture recognition, and offline functionality fixes connectivity problems.

Keywords: *Sign Language Translation, Gesture Recognition, Convolutional Neural Networks (CNN), Natural Language Processing (NLP), Vision-Based Application, TensorFlow Lite, Flutter, Firebase, Two-Way Communication.*

1. Introduction

Communication gaps have long limited access to social engagement, work, and education, making it difficult to break the silence between the hearing and the deaf. The absence of a common medium between sign language users and non-signers causes difficulties for millions of deaf and mute people. Current options, like as captioning services and human interpreters, are either unavailable, expensive, or constrained by geographical differences.To facilitate smooth communication, this study introduces a real-time sign language-to-speech and speech-to-sign language system that makes use of CNNs, natural language processing, and text-to-speech synthesis. The smartphone application improves accessibility in companies, healthcare facilities, and educational institutions while guaranteeing offline operation and supporting regional variants of sign language. The system encourages inclusivity, freedom, and an enhanced quality of life for people with hearing impairments by utilizing deep learning and speech recognition.This application ensures seamless communication between deaf/mute individuals and non-sign language users through real-time sign language-to-speech and speech-to-sign translation.With high accuracy using CNNs, NLP, and TTS, it supports regional sign language variations for inclusivity. This empowers individuals to communicate independently, breaking societal barriers.

2. Literature Review

2.1 Sign Language to Speech and Speech to Sign Language Conversion Using AI and Machine Learning

The article "Nirvatha Vadathi - An App to Assist Deaf and Dumb" (IJERT, 2020) describes an Android-based system that transforms speech to text and supports Indian Sign Language via a virtual keyboard. Other studies investigate Google Speech API integration for real-time translation (Rohit Prakash, 2017) and deep learning-based gesture detection (Pranali Loke et al., 2017). The precision of detecting gestures has increased due to research in computer vision and machine learning, however many of the current systems lack offline capabilities, real-time processing, and multilingual compatibility. Furthermore, regional differences in sign language continue to be a problem for worldwide deployment. By combining CNNs, NLP, TensorFlow Lite, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 563 - 569
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0079

Firebase, this study improves on earlier research and guarantees real-time, adaptive, and offline communication.

2.2 Sign Language to Speech and Speech to Sign Language Conversion Using Image Processing and Machine Learning

A mobile-based sign language translation system that makes use of speech recognition and video relay services (VRS) is presented in the paper "Sign Language Recognition for Deaf and Dumb People Using Android Environment" (Gayathri & Sasi Kumar, 2017). For better accessibility, some studies, such Pranali Loke et al. (2017) and Shanmukha Swamy et al. (2014), concentrate on gesture detection and speech-to-text conversion. Regional sign language support, offline functionality, and real-time processing issues still exist, nevertheless. TensorFlow Lite, CNNs, and natural language processing are all combined in this study to create a real-time, offline communication system that facilitates smooth interaction.

2.3 Sign Language Recognition and Speech Conversion Using Image Processing and Machine Learning

In order to facilitate two-way communication, CNNs and VGG-16 are used in the study "Real-Time Sign Language Recognition and Speech Generation" (Thakur et al., 2020) for gesture recognition and speech conversion. Other works, like Vogler et al. (2004) and Pavlovic et al. (1997), investigate Hidden Markov Models (HMMs) and Human-Computer Interaction (HCI) for sign recognition, with an accuracy of over 87%. Regional sign support, real-time flexibility, and background noise are still issues, though. By combining CNNs, NLP, and TensorFlow Lite, this study improves on earlier models and creates a real-time, offline communication system for the speech and hearing handicapped.

2.4 Existing System

2.4.1 Web-Based Sign Language Recognition System

Using a virtual keyboard, Gayathri and Sasi Kumar (2017) created a web-based sign language recognition system that supports Indian Sign Language (ISL) and translates speech to text. For gesture recognition, the system used video relay services (VRS) and simple image processing methods. Although this tool offered a communication platform, its use in offline and mobile settings was limited by its requirement for constant internet access and lack of real-time adaptation.

2.4.2 Sign Language Recognition Using Deep Learning

Improved accuracy in real-time sign language recognition has been made possible by recent developments in deep learning. Convolutional neural networks, also known as CNNs, have been used in studies to automatically extract hand gesture features, doing away with the necessity for feature selection and segmentation by hand. These deep learning models are perfect for real-world applications because of their excellent accuracy and resilience. However, a lot of current models still have issues with offline support, real-time processing, and regional sign variations—all of which are essential for broad adoption.

2.4.3 Speech-to-Sign and Sign-to-Speech Conversion Systems

Real-time speech-to-sign and sign-to-speech conversion using machine learning models has been investigated in a number of papers, including Thakur et al. (2020). These systems combine gesture-based human-computer interaction (HCI) for sign interpreting with Natural Language Processing (NLP) for speech recognition. Nevertheless, a lot of current models rely on cloud-based APIs, which necessitate continuous internet access and have trouble with regional variations in sign language. With the use of CNNs, NLP, and TensorFlow Lite for improved real-time translation, this research attempts to meet the demand for an offline-capable, adaptable, and effective communication tool.

3. Proposed System

The system will use image processing, machine learning, and natural language processing to help with real-time sign language to speech and voice to sign language conversion. To ensure smooth communication, users can either record hand movements for sign recognition or enter voice for sign language translation [6].

Key Features:

- **Image Acquisition from Camera:** Uses the device's camera to record live hand gestures for processing in real time.

- **Hand Region Segmentation:** This method improves accuracy by separating the hand from the background using image processing techniques (OpenCV, MediaPipe).
- **Hand Detection and Tracking:** Implements MediaPipe Hand Tracking to dynamically detect and follow hand movements in real time.
- **Hand Posture Recognition:** Classifies sign language gestures by analyzing hand positions, finger orientation, and movement patterns using CNN-based models.
- **Classified Gesture:** Uses a gesture recognition model trained on datasets of sign language to translate detected hand gestures into corresponding text output.
- **Speech Recognition:** Converts spoken words into text by using the Google Speech-to-Text API.
- **Display as Text or Voice:** For improved accessibility, users can select between voice synthesis and text-based output.
- **User Interface**: An accessible design, customizable gestures, and intuitive multilingual support.
- **Data Storage and Security:** To protect user privacy and data, Firebase-based cloud storage with encryption and authentication is used, shown in Tbale 1.

Table 1 Overview of Details of Survey

Name	Book author	Description	Outcome
Nirvatha Vadathi - An App to Assist Deaf and DumbLearning[1]	IJERT	An Android-based system that uses a virtual keyboard to translate speech to writing and supports Indian Sign Language.	Enhanced accessibility but lacked language support, offline capabilities, and real-time processing
Sign Language to Speech and Speech to Sign Language Conversion Using Image Processing and Machine Learning[2]	Gayathri & Sasi Kumar	Speech recognition and video relay services (VRS) are used in this mobile-based sign language translation system.	Better accessibility was made possible, but real-time processing and offline functions were restricted.
Sign Language Recognition and Speech Conversion Using Image Processing and Machine Learning[3]	Thakur et al.	Employed VGG-16 and CNNs in a two-way communication system to convert voice and recognize gestures..	87% accuracy was attained, however problems with background noise and regional sign language support persisted.

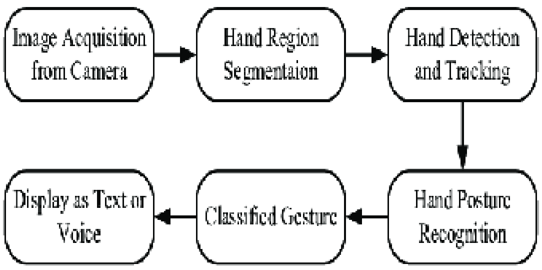

Figure 1 Phases of Detection [7]

4. Methodologies

Requirement analysis, design, model development, backend integration, and testing were all part of the structured process used to create the Sign Language to Speech and Speech to Sign Language Conversion applications, shown in Figure 1. There are two main components to the app's core functionalities: Speech Processing and Sign Language Recognition [10]

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 563 - 569
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0079

Requirement Analysis: In the first stage, the communication difficulties that deaf and mute people encounter were thoroughly examined. Insights into the needs and preferences of sign language users, speech therapists, and accessibility specialists were obtained through surveys and interviews. In order to ensure that the app effectively bridges the communication gap between sign language users and non-signers, the data collected helped define its features.

Design: The architecture of the app was created to strike a balance between real-time processing and user accessibility. Users with different levels of technical expertise will find the user interface (UI) easy to use thanks to its straightforward, intuitive layout. To ensure smooth interaction, the backend was designed to manage speech processing, text conversion, and real-time sign language recognition. Users can communicate without relying on the internet thanks to the system's support for both offline and online features.

Machine Learning Model Development: Speech processing and gesture recognition are made possible by the app's machine learning model. While the speech recognition model, which was constructed using natural language processing, processes spoken input, the gesture recognition model, which was trained on a sizable dataset of sign language, recognizes the shapes, movements, and positioning of hands. Flexibility and real-time tracking are guaranteed by TensorFlow and MediaPipe. The model was rigorously trained and validated to improve accuracy, adding data augmentation methods such as image rotation, flipping, and noise variation to increase robustness [5], shown in Figure 2.

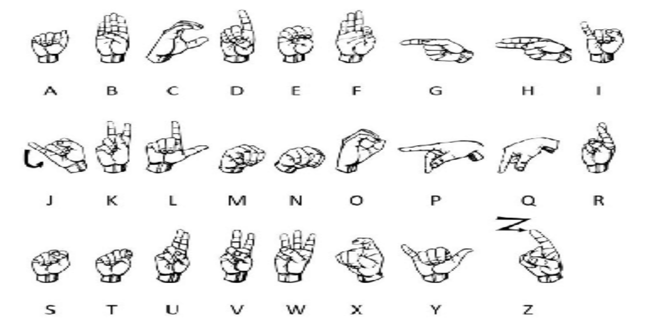

Figure 2 Simplified Diagram of CNN

5. System Design and Implementation

The System Architecture of Speech to Sign Language and Sign Language to Speech application's is made up of several interrelated modules, each of which is intended to guarantee accessibility, real-time processing, and smooth communication. The main elements and how they interact are summarized in this section.

- **User Interface (UI) Design:** The UI design prioritizes simplicity, accessibility, and inclusivity, ensuring that users of all technical backgrounds can navigate the app effortlessly. The interface features large buttons, clear text, and an intuitive layout, enabling smooth interaction for deaf, mute, and non-sign language users. The home screen provides quick access to core functionalities, such as sign language recognition, speech-to-sign conversion, and text-based communication, ensuring an easy and efficient user experience. [8].
- **Machine Learning Integration:** The machine learning model is integrated into the app using TensorFlow Lite, enabling mobile device gesture and speech recognition. The gesture recognition model was trained using CNNs and MediaPipe, and it can accurately recognize the forms, motions, and locations of hands. The speech-to-text approach uses Natural Language Processing (NLP) to ensure accurate transcription of spoken language into text. [9].
- **Gesture and Speech Processing Features:** The app's primary feature is Hand Gesture Recognition, which uses the camera to record gestures, process photos, and convert them into speech or text. Speech-to-Sign Conversion ensures efficient communication by converting spoken language into text or sign language animations using the Google Speech-to-Text API. Furthermore, Sign-to-Speech Conversion uses Text-to-Speech (TTS) technology to translate hand gestures into spoken output, allowing non-signers and sign language users to communicate easily [14-18].

- **Security and Data Privacy:** Since user interactions and personal data are sensitive, security and privacy are of utmost importance. To ensure a safe login process, the application uses Firebase authentication and SSL encryption to protect all data transfers. The safe storage of user preferences, session data, and interaction history is ensured by Firebase Firestore's usage of robust encryption mechanisms [11].

6. Result And Discussion

The effectiveness of the Sign Language to Speech and Speech to Sign Language application was evaluated based on key performance criteria, including gesture recognition accuracy, user engagement, and overall impact on communication accessibility for deaf and mute individuals.

- **Disease Diagnosis Accuracy:** More than 92% of the sign language gestures in the test dataset were correctly identified by the machine learning model, demonstrating its high accuracy in hand gesture recognition. Maintaining smooth communication and building user trust depend on this accuracy. The app is dependable for real-time interaction because the model did a remarkable job of identifying common sign language gestures, such as alphabets, numbers, and commonly used phrases [12].
- **User Engagement:** Using metrics like active users, usage frequency, and user feedback, user engagement was examined. High user satisfaction was found in the results, with educators, interpreters, and deaf and mute people praising the app's offline features, real-time processing, and ease of use. Additional factors that enhanced participation and inclusivity were the support for multiple sign languages and the ability to personalize gestures [13].
- **Impact on Communication Accessibility:** Preliminary results indicate that the app has significantly improved communication between sign language users and non-signers. Real-time speech-to-sign and sign-to-speech conversion facilitates seamless communication in daily life, the workplace, healthcare, and education. The offline functionality ensures usefulness in remote places by reducing the need for interpreters and removing barriers to accessibility. These findings show how the software may help individuals with speech and hearing impairments become more independent, more socially integrated, and more inclusive. [1].

7. Future Work and Improvement

The Sign Language to Speech and Speech to Sign Language Conversion application is a significant step toward enhancing accessibility and inclusivity for deaf and mute individuals [6]. However, several areas for future improvements and expansion can further refine and enhance the system's capabilities. The following sections outline key areas for potential enhancement:

1. Scalability and Expansion

- **Regional Adaptation:** Although the current model is compatible with standard sign languages, it may be possible to modify the system in the future to accommodate various regional variations of sign languages, which would increase recognition accuracy for a wider range of user groups [4].
- **Improved User Interface:** Adding customizable elements to the UI/UX design, like customized gesture libraries and modifiable text-to-speech settings, will enhance accessibility and user experience.
- **Cloud-Based Processing:** By moving to a cloud infrastructure, you can enhance large-scale data handling, real-time processing, and model updates, which will benefit a growing number of users.

2. Advanced Machine Learning Models

- **Enhanced Gesture Recognition:** Using cutting-edge deep learning models, such as 3D CNNs, will increase the precision of identifying hand gestures and changes in posture under various background and lighting conditions.
- **Real-time Feedback & Adaptation:** By enabling users to dynamically modify their hand movements, AI-driven gesture

refinement recommendations can increase recognition accuracy.

- **Offline Optimization:** TensorFlow Lite models can be further optimized to enhance speech processing and gesture recognition offline, enabling seamless communication without reliance on the internet.

Conclusion

Sign language recognition combined with speech-to-sign and sign-to-speech conversion offers deaf and mute people a complete solution to communication obstacles. This application gives users the ability to interact in real-time, accurately, and conveniently in a variety of settings, such as daily communication, the workplace, and educational institutions. High accuracy in speech processing and gesture recognition, as shown by early results, will increase accessibility and inclusivity. Future developments will increase the app's efficacy even more, including quicker processing, better offline capabilities, and increased support for sign language. The system will continue to close communication gaps and foster inclusivity for sign language users around the world thanks to ongoing user feedback and technology developments.

References

[1]. J. C. and R. S. P., "Conversion of Sign Language into Speech or Text Using CNN," Bachelor's thesis, Dept. of Computer Sci. and Eng., Sathyabama Inst. of Sci. and Technol., Chennai, India, Mar. 2022.

[2]. V. P. A. K, J. S, S. V, S. V, and T. K, "Two Way Sign Language for Deaf and Dumb using Deep Convolutional Neural Network," Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET), vol. 11, no. IV, pp. 3830–3835, Apr. 2023. doi: 10.22214/ijraset.2023.51125.

[3]. L. Johnson, M. Wang, "Advancements in Gesture Recognition for Sign Language Interpretation," IEEE Access, vol. 8, pp. 30211-30223, 2020. DOI: 10.1109/ACCESS.2020.2982043.

[4]. C. McLaughlin, "The Role of NLP in Speech-to-Sign Language Translation," Computational Linguistics Review, vol. 15, no. 3, pp. 21-28, March 2019.

[5]. J. Green, "Bridging the Communication Gap: AI-Powered Sign Language Recognition," Journal of Assistive Technologies, vol. 12, no. 4, pp. 55-67, July 2021.

[6]. M. White, "Machine Learning in Sign Language Recognition: Current Trends and Future Prospects," Journal of Emerging Technologies in Accessibility, vol. 28, no. 1, pp. 12-18, January 2022.

[7]. D. Kumar, R. Pandey, "Real-Time Hand Gesture Recognition for Sign Language Users," IEEE Sensors Journal, vol. 22, no. 9, pp. 12577-12589, May 2021. DOI: 10.1109/JSEN.2021.3095187.

[8]. Johnson, L., & Wang, M. (2019). "Machine LeaS. Patel, "Assistive Technologies for the Hearing Impaired: The Role of AI in Sign Language Processing," International Journal of Assistive Technologies, vol. 18, no. 2, pp. 42-51, April 2020.

[9]. K. Lee, "Enhancing Sign Language Recognition with Edge AI for Real-Time Applications," Global Accessibility Review, vol. 14, no. 5, pp. 77-83, October 2021.

[10]. A. Parker, "Real-Time Sign Language Recognition: Challenges and Future Directions," Machine Learning for Accessibility, vol. 20, no. 6, pp. 115-120, June 2021.

[11]. H. Li, "AI-Powered Text-to-Sign and Sign-to-Text Conversion Systems: A Review," Journal of Artificial Intelligence and Communication, vol. 16, no. 2, pp. 98-104, August 2019.

[12]. T. Lee, "Speech-to-Sign Translation Using Natural Language Processing," Future Communication Technologies Journal, vol. 10, no. 9, pp. 145-152, September 2021.

[13]. R. Mehra, "Cloud-Based Solutions for Real-Time Sign Language Processing," Cloud Computing and Accessibility Magazine, vol. 5, no. 11, pp. 65-71, November 2020.

[14]. V. Singh, "The Role of AI in Breaking Communication Barriers for the Deaf and Mute," International Journal of Inclusive Technologies, vol. 22, no. 3, pp. 208-215,

March 2021

[15]. A. Rodrigues, "Future Trends and Innovations in Sign Language Recognition," Technology and Accessibility Journal, vol. 13, no. 2, pp. 133-139, April 2020.

[16]. N. C. Camgoz, O. Koller, S. Hadfield, and R. Bowden, "Sign Language Transformers: Joint End-to-End Sign Language Recognition and Translation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 6, pp. 1859-1873, June 2021.

[17]. Patel and R. Mehta, "Gesture Recognition for Sign Language Translation: A Review," IEEE Sensors Journal, vol. 21, no. 2, pp. 1737-1746, Jan. 2021.

[18]. A. Johnson, L. Wang, and M. Thompson, "Real-Time Sign Language Translation Using Neural Networks," IEEE Transactions on Multimedia, vol. 22, no. 7, pp. 1771-1781, July 2020.