

**BRIDGING COMMUNICATION GAPS USING AI-BASED GESTURE RECOGNITION:
SIGNCONNECT.APP**

*** Sakshi Ravindra Singh & ** Asst. Prof. Hemangi J. Talele**

* Dept. of CS & IT, B. K. Birla College (Empowered Autonomous), Kalyan.

** Dept. of CS & IT, B. K. Birla College (Empowered Autonomous), Kalyan.

Abstract:

SignConnect.App presents a practical assistive communication system aimed at supporting interaction between hearing- and speech-impaired individuals and non-sign language users. The application interprets static hand gestures and converts them into readable text and synthesized speech using computer vision techniques. Hand landmark extraction is performed through MediaPipe, while OpenCV, Tkinter, and pyttsx3 are employed for gesture processing, interface design, and offline speech output, respectively. The system is designed to recognize English alphabet gestures along with a limited set of commonly used words relevant to daily communication. In contrast to many existing solutions that rely on specialized hardware or internet-dependent processing, the proposed application operates entirely offline using a standard webcam. Experimental testing under typical indoor conditions indicates reliable recognition performance with minimal response delay, suggesting that the system offers a cost-effective and accessible solution for real-world assistive communication scenarios.

Keywords: Sign Language, Hand Gesture Recognition, Assistive Technology, Artificial Intelligence, Accessibility

Copyright © 2026 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

Introduction:

Communication plays a central role in everyday social interaction; however, individuals with hearing and speech impairments often experience difficulty when communicating with those unfamiliar with sign language. While sign language remains an effective and expressive medium within the hearing-impaired community, its limited use among the broader population continues to create barriers in education, healthcare, and public environments.[1]

Assistive systems capable of translating hand gestures into text or speech can help reduce these barriers and promote inclusive interaction [2]. Hand gesture recognition plays a key role in human-computer interaction but faces challenges such as background complexity, motion induced rotation errors, and

computational delay.

Earlier approaches to gesture recognition frequently depended on sensor-based gloves or complex learning models, which increased system cost and limited portability. Such constraints reduce practicality in real-world use, particularly in resource-constrained settings. In response, SignConnect.App is developed as a lightweight, vision-based solution that functions in real time using a conventional webcam. By incorporating offline text-to-speech capabilities, the system provides immediate feedback without requiring network connectivity [3]. The work focuses on achieving a balance between affordability, usability, and recognition reliability, making the system suitable for everyday assistive communication.

Literature Review:

This section explores a background work on sign language. The purpose of this section is to present a review of the context and progress of the research theme. Various researcher studied this domain, proposing different methods for text-to-speech capabilities using sign language [4]. It is an open field, offering opportunities for future advancements to enhance prediction accuracy.

Kadous et al. proposed a model proposed glove-based sign language recognition systems that provided accurate hand measurements; however, these systems were expensive, non-portable, and unsuitable for everyday use[5]. Fels and Hinton (1993) developed hardware-intensive gesture recognition systems that required multiple sensors, significantly limiting their usability and accessibility[6]. Freeman and Roth et al. introduced contour-based hand tracking techniques, which were effective under controlled conditions but highly sensitive to lighting variations and background clutter[7]. Starner et al. utilized Hidden Markov Models (HMMs) for American Sign Language recognition, achieving promising results but only in constrained and controlled environments[8].

Mitra and Acharya et al. reviewed rule-based and machine learning-based gesture recognition systems, noting that rule-based approaches were lightweight but less robust, while ML-based systems required large datasets and higher computational resources[9]. Jain et al. proposed CNN-based hand gesture recognition models that achieved high accuracy but demanded significant computational power, making them unsuitable for low-end systems[10].

Cui et al. (2025) provided a systematic review of deep vision-based real-time hand gesture recognition techniques, highlighting the role of deep learning models including CNN, LSTM and attention mechanism in improving recognition accuracy, response time, and robustness under complex

environmental conditions. This review compares these models, analyzes evaluation metrics and datasets, and highlights existing research gaps related to accuracy, efficiency, and real-world adaptability.[11]

Porwal et al. presented a mathematical and computational framework for real-time American Sign Language (ASL) alphabet recognition using fingertip landmark geometry extracted via MediaPipe. In this research approach employed geometric normalization, statistical feature analysis (PCA and LDA), and lightweight classification techniques, achieving high accuracy with low computational latency, making it suitable for real-time assistive applications [12].

Initial research in sign language recognition largely emphasized hardware-assisted methods. Early glove-based systems enabled precise tracking of finger movements but were often costly and inconvenient for continuous use. Similar hardware-heavy approaches demonstrated acceptable accuracy while introducing challenges related to comfort, maintenance, and scalability.

To overcome these limitations, researchers explored camera-based gesture recognition techniques. Contour and appearance-based hand tracking methods reduced reliance on physical sensors but proved sensitive to environmental factors such as lighting variation and background complexity. Statistical modelling approaches, including Hidden Markov Models, improved recognition consistency under controlled conditions; however, their performance often declined in unconstrained real-world settings [13]. With advances in machine learning, data driven techniques gained attention. Comparative studies highlighted the trade-off between rule based systems, which are computationally efficient but less adaptive, and learning-based models, which require extensive training data and processing resources. Although convolutional neural network-based approaches achieved high recognition accuracy, their dependency

on powerful hardware and increased latency limited their suitability for lightweight real-time applications. The introduction of MediaPipe Hands provided an efficient framework for real-time hand landmark detection with relatively low computational overhead [14]. Subsequent studies integrated this framework with classification techniques to improve gesture recognition performance. Recent research trends emphasize the importance of offline, low-cost assistive systems capable of functioning reliably in everyday environments.

Research Gap:

Despite continued progress in gesture recognition technologies, many existing systems remain constrained by hardware dependency, high computational demands, or reliance on internet connectivity. Additionally, several approaches focus solely on visual output without providing integrated speech feedback. These limitations reduce usability in practical assistive communication scenarios. There is a clear need for a compact, offline system that combines gesture recognition with both textual and auditory output while maintaining low cost and real-time responsiveness. SignConnect.App is designed to address these requirements through a rule-based approach built upon MediaPipe hand landmark detection and offline speech synthesis [15].

Research Methodology :

The proposed system follows a modular design that enables efficient real-time gesture recognition. By dividing the application into independent functional units, the design simplifies processing, improves execution speed, and allows seamless interaction between gesture detection, classification, and output generation components.

This structured design ensures that each module operates independently while contributing to the

overall performance of the system, making maintenance and future enhancements easier. The modular approach also enhances system stability, as changes in one component do not affect the entire workflow, thereby supporting accurate and consistent gesture recognition in real-time environments.

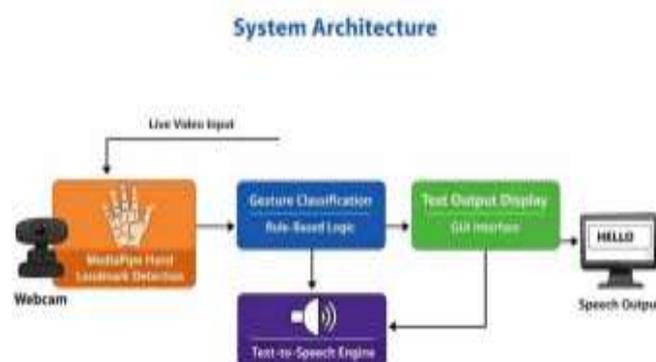


Figure 1: System Architecture

The application captures live video input through a webcam, which is processed frame-by-frame using MediaPipe to extract twenty-one hand landmarks. These landmarks represent key finger joints and palm positions, forming the basis for gesture interpretation. A rule-based gesture classification module evaluates the relative positions of these landmarks to identify corresponding alphabets or functional words. The detected output is displayed as text within the graphical interface and can be converted into speech using an offline text-to-speech engine. A custom dataset was manually created by the authors, consisting of static hand gestures representing English alphabets (A–Z) and commonly used words such as “Hello,” “Yes,” “No,” and “Thank You.” Data samples were collected under varying lighting conditions, hand sizes, orientations, and backgrounds to enhance robustness. During dataset creation, challenges such as similar finger configurations and inconsistent illumination was encountered and addressed through refined gesture rules.

Data Flow Diagram of Gesture Recognition Process

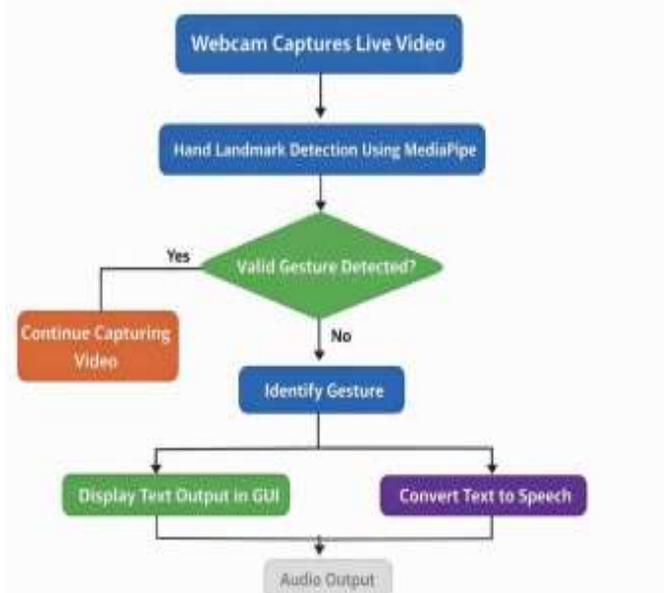


Figure 2: Data Flow Diagram of Gesture Recognition Process

Here Gesture detection relies on geometric relationships between landmarks rather than deep learning models. This approach was chosen to minimize computational overhead, enable offline execution, and allow easy personalization. This design choice allows the system to remain lightweight while still delivering dependable recognition performance without requiring large training datasets or high-end hardware. By focusing on relative landmark positions and spatial relationships, the method supports consistent operation across different users and environments, making the application easier to deploy, adapt, and maintain for real world assistive communication scenarios.

Results and Analysis:

The system was evaluated under standard indoor conditions using a conventional webcam. The evaluation involved multiple test sessions conducted with different users to assess stability and repeatability of the results. Gesture recognition performance remained consistent across variations in hand

orientation and moderate changes in lighting conditions. The system maintained smooth real-time operation without noticeable delays or crashes during continuous use. These observations indicate that the application is reliable for practical day-to-day communication in indoor environments. Experimental results show that the system successfully recognizes static alphabet gestures with an average accuracy of approximately 95 percent. Functional word recognition achieved accuracy values ranging between 92 and 96 percent. The response latency for gesture detection and output generation remained below one second, making the system suitable for real-time communication.

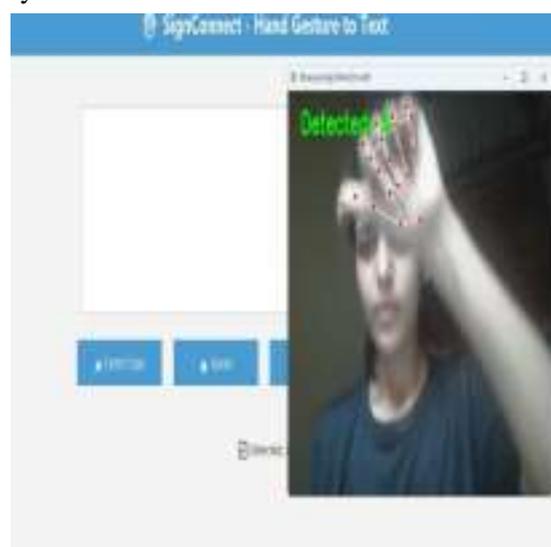




Figure 3: Interface Hand Gesture to text

These performance outcomes demonstrate that the proposed system maintains reliable recognition across varying hand orientations and lighting conditions commonly found in indoor environments. The consistently low response time enhances conversational flow by providing near-instant feedback, while the high accuracy confirms the effectiveness of combining MediaPipe-based landmark extraction with rule-based gesture interpretation for practical, real world assistive communication.



Figure 6: Bar Chart – Recognition Accuracy for Gestures and Pie Chart – Overall Success vs. Failure Rate.

The bar chart illustrates consistent accuracy across different gestures, while the pie chart highlights a high overall success rate, indicating the effectiveness of the rule-based approach combined with MediaPipe. The absence of specialized hardware further reinforces the system's practicality for everyday use.

Discussion:

The evaluation results suggest that SignConnect.App offers a reliable means of translating static hand gestures into both text and speech in real-time settings. When compared with sensor-dependent or deep learning-centric systems, the proposed approach reduces implementation complexity and resource requirements without a substantial loss in recognition accuracy. The ability to operate offline enhances both privacy and reliability, particularly in environments with limited or inconsistent internet access. Furthermore, the simplicity of the graphical interface supports ease of use for individuals with varying levels of technical familiarity, reinforcing the system's applicability for everyday assistive communication.

References:

1. Kumar, A., K. Thankachan, and M.M. Dominic. Sign language recognition. in 2016 3rd international conference on recent advances in information technology (RAIT). 2016. IEEE.
2. Balamanikandan, A., et al. Gesture to Speech: A Wearable Solution for Sign Language Translation. in 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS). 2024. IEEE.
3. Bhushan, S., et al., Advancing Text-to-Speech Systems for Low-Resource Languages: Challenges, Innovations, and Future Directions. IEEE Access, 2025.
4. Trivedi, A., et al., Speech to text and text to speech recognition systems-Areview. IOSR J. Comput. Eng, 2018. 20(2): p. 36-43.
5. Kadous, M.W. Machine recognition of Auslan signs using PowerGloves: Towards,large-lexicon recognition of sign language. in Proceedings of the Workshop on the Integration of Gesture in Language and Speech. 1996. DE Wilmington.
6. Fels, S.S. and G.E. Hinton, Glove-talk: A neural network interface between a data glove and a



- speech synthesizer. *IEEE transactions on Neural Networks*, 1993. **4**(1): p. 2-8.
7. Freeman, W.T. and M. Roth. Orientation histograms for hand gesture recognition. in *International workshop on automatic face and gesture recognition*. 1995. Zurich, Switzerland.
8. Starner, T., B. Schiele, and A. Pentland. Visual contextual awareness in wearable computing. in *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No. 98EX215)*. 1998. IEEE.
9. Mitra, S. and T. Acharya, *Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2007. **37**(3): p. 311-324. 10. Jain, R., R.K. Karsh, and A.A. Barbhuiya, *Encoded motion image-based dynamic hand gesture recognition. The visual computer*, 2022. **38**(6): p. 1957-1974.
11. Cui, C., M.S. Sunar, and G.E. Su, *Deep vision-based real-time hand gesture recognition: a review. PeerJ Computer Science*, 2025. **11**: p. e2921.
12. Porwal, A., *GEOMETRIC AND STATISTICAL ANALYSIS OF FINGERTIP LANDMARK FEATURES FOR REAL-TIME AMERICAN SIGN LANGUAGE RECOGNITION. International Journal of Applied Mathematics*, 2025. **38**(5s): p. 972- 982.
13. Mark, G. and Y. Steve, *The application of hidden Markov models in speech recognition. Foundations and Trends® in Signal Processing*, 2024. **1**(3): p. 195-304. 14. Veluri, R.K., et al. *Hand gesture mapping using MediaPipe algorithm. in Proceedings of Third International Conference on Communication, Computing and Electronics Systems: ICCCES 2021*. 2022. Springer.
15. Harris, M. and A.S. Agoes. *Applying hand gesture recognition for user guide application using MediaPipe. in 2nd International Seminar of Science and Applied Technology (ISSAT 2021)*. 2021. Atlantis Press.

Cite This Article:

Singh S.R. & Asst. Prof. Talele H.J. (2026). *Bridging Communication Gaps Using AI-Based Gesture Recognition: SignConnect.App.* In **Aarhat Multidisciplinary International Education Research Journal**: Vol. XV (Number I, pp. 86–91) **Doi:** <https://doi.org/10.5281/zenodo.18608739>