Grenze International Journal of Engineering and Technology, June Issue

SignLink: A Real-Time Web-based Sign Language Recognition System using MediaPipe and PeerJS

Jane Rubel Angelina[1], M. Balanjeneyulu[2], K. Lakshman Chowdary[3], P. Devi Prasad[4] and K. Chiranjeevi[5]
[1-5]Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil-626126. Tamil Nadu, India.
Email: j.janerubelangelina@klu.ac.in, balumadala29@gmail.com, kethinenilakshman@gmail.com, deviprasadash37@gmail.com, chiru.koya@gmail.com

***Abstract*— Sign language is an essential communication for individuals with hearing and speech impairments, yet it remains inaccessible to much of the general population. This paper presents SignLink, a real-time web-based system that translates hand gestures into corresponding alphabetic characters using MediaPipe Hands and PeerJS. The application captures live video input, detects and tracks 21 hand landmarks, and computes angular relationships between finger joints to classify gestures representing letters and words. PeerJS enables peer-to-peer video communication, allowing users to join or create virtual rooms for interactive sign recognition sessions. The system operates directly within a browser, eliminating the need for additional installations or hardware. Experimental testing demonstrates the potential of the framework to provide accurate, low-latency gesture recognition while promoting accessibility and inclusivity. SignLink highlights the integration of computer vision and real-time web technologies for effective human–computer communication.**

***Keywords*— Sign Language Recognition, Real-Time Gesture Detection, MediaPipe Hands, Peer-to-Peer Communication, Hand Landmark Tracking, Human–Computer Interaction, Web-Based Accessibility, Computer Vision, Quality Education, Decent Work and Economic Growth.**

I. Introduction

Communication is a fundamental aspect of human interaction, yet individuals with hearing or speech impairments often encounter significant barriers when interacting with others. Sign language serves as a vital communication medium for the deaf and hard-of-hearing community, but its limited understanding among the general population restricts social inclusion. With recent advances in computer vision and artificial intelligence, real-time gesture recognition systems have emerged as effective tools to bridge this communication gap. These systems enable machines to interpret hand gestures and convert them into meaningful text or speech, enhancing accessibility and promoting inclusive communication.

SignLink is a web-based application designed to recognize and interpret hand gestures into corresponding letters using MediaPipe Hands, a robust hand landmark detection framework, and PeerJS for real-time peer-to-peer video communication. The system captures live video, tracks 21 hand landmarks, and calculates angles between finger joints to determine the gesture. Recognized gestures are displayed on the interface in real time, enabling both local and remote users to communicate seamlessly. This browser-based approach eliminates the need for specialized hardware or installations, making the system accessible and scalable.

Grenze ID: 01.GIJET.12.2.407_24

To achieve accurate and reliable results, SignLink employs a combination of geometric feature extraction and threshold-based classification. By analyzing joint angles, finger straightness, and distances between fingertips, the system can distinguish between subtle differences in hand poses. Extensive testing and iterative calibration of the angle thresholds ensured minimal misclassification, even under varying lighting conditions and hand orientations. Using a single-hand model optimized for speed and low latency further improves real-time responsiveness without sacrificing recognition accuracy.

The motivation behind SignLink lies in promoting inclusivity and demonstrating the practical integration of computer vision with web-based technologies. By combining MediaPipe's powerful landmark detection with PeerJS's real-time communication capabilities, the system provides an interactive platform for learning, collaboration, and accessibility. The results indicate that SignLink is capable of delivering high-accuracy gesture recognition with minimal delay, making it a promising tool for educational environments, sign language learning applications, and communication aids for individuals with disabilities.

II. Related Works

Khan et al. [1] (2025) provided a comprehensive review of deep learning approaches for continuous sign language recognition, highlighting the evolution of models from traditional machine learning to advanced neural networks. Their study emphasized the importance of robust feature extraction and temporal modeling for accurate gesture recognition.

Trabelsi et al. [2] (2025) explored vision-based sign language recognition, discussing the challenges in hand detection, tracking, and gesture classification. They highlighted the limitations of current systems in real-time performance and accuracy under varying lighting and hand orientations.

Rastogi et al. [3] (2025) reviewed machine learning techniques for hand gesture-based sign language recognition. They focused on methods that leverage geometric features, skeleton models, and deep learning architectures, emphasizing the improvements in recognition accuracy achieved through hybrid approaches.

Hassan et al. [4] (2025) proposed a novel model aimed at expanding the horizons of sign language recognition. Their work focused on improving the system's ability to handle diverse hand sizes and complex gestures through enhanced landmark processing and feature analysis.

Papadimitriou et al. [5] (2025) developed a Greek sign language recognition system specifically for educational platforms. They demonstrated how gesture recognition can be integrated into learning environments, providing real-time feedback and improving accessibility for students.

Mineo et al. [6] (2025) introduced radar-based imaging techniques for sign language recognition in medical communication contexts. Their approach offered an alternative to vision-based systems, reducing dependence on lighting conditions and camera angles.

Najib [7] (2025) implemented a multi-lingual sign language recognition system using machine learning, enabling recognition across multiple sign languages. This study highlighted the importance of dataset diversity for building scalable recognition systems.

Zholshiyeva et al. [8] (2025) designed the QazSL system for physically impaired individuals, focusing on isolated sign recognition. Their approach emphasized user-centric design and adaptability for different levels of motor ability.

Wang et al. [9] (2025) proposed a continuous sign language recognition system with multi-scale spatial-temporal feature enhancement. By capturing both local and global hand movements, they achieved improved recognition performance for dynamic gestures.

Huang et al. [10] (2025) developed a dual-stage temporal perception network to handle continuous sign language sequences. Their model effectively captured temporal dependencies, leading to higher accuracy in recognizing long gesture sequences.

Guan et al. [11] (2025) introduced a Multi-stream Keypoint Attention (MSKA) network for both recognition and translation of sign language. Their method utilized attention mechanisms to focus on critical hand landmarks, improving classification accuracy.

Algafri et al. [12] (2025) presented a semi-supervised learning method for isolated sign language recognition (SSLR). This approach reduced the reliance on large labeled datasets, enabling more efficient training of recognition models.

Uddin et al. [13] (2025) implemented real-time Norwegian sign language recognition using MediaPipe and LSTM networks. Their system combined precise hand landmark detection with temporal modeling to achieve low-latency performance suitable for interactive applications.

Li et al. [14] (2025) explored YOLOv5-based methods for sign language recognition, demonstrating the effectiveness of object detection frameworks in hand gesture localization and classification.

Modali et al. [15] (2025) proposed a ResNet50-LSTM hybrid approach using temporal-spatial feature fusion. Their system effectively captured spatial hand features and temporal dynamics, leading to improved recognition of continuous gestures.

Min et al. [16] (2025) focused on skeleton-based continuous sign language recognition, analyzing the performance of models that leverage skeletal representations for temporal and spatial feature extraction. Their study highlighted the robustness of skeleton-based approaches in handling complex sequences.

III. Comparison with Previous Methodology

Traditional sign language recognition systems have relied heavily on sensor-based gloves, specialized cameras, or wearable devices to capture hand movements. While these approaches can provide precise measurements of joint positions and finger flexion, they are often expensive, cumbersome, and require users to adapt to specialized hardware. Moreover, the dependency on physical devices limits accessibility and widespread adoption, particularly in educational or public settings.

In contrast, vision-based systems using computer vision frameworks such as OpenCV or TensorFlow have enabled gesture recognition without the need for wearable sensors. These systems typically process video frames to extract hand contours, skin color segmentation, or motion patterns. Although more accessible than sensor gloves, traditional vision-based systems can suffer from poor accuracy under varying lighting conditions, complex backgrounds, or different hand orientations. Many existing implementations also struggle to provide real-time remote communication, as most frameworks focus on offline or local recognition.

SignLink improves upon these previous methodologies by integrating MediaPipe Hands, which provides robust, high-precision hand landmark detection with minimal computational overhead. By tracking 21 key landmarks per hand and analyzing angular relationships between joints, the system achieves accurate gesture classification even with variations in hand position or orientation.

The use of PeerJS for peer-to-peer communication allows real-time remote interaction, which most earlier computer vision systems did not support. Additionally, SignLink's angle-based and distance-based feature extraction ensures more reliable recognition across users with different hand sizes, improving overall performance compared to threshold- or contour-based methods.

TABLE I: COMPARISON TABLE

System Type	Limitations	Advantages of SignLink
Sensor-Based Systems	Expensive, cumbersome, requires specialized hardware	No additional hardware needed, easy to use
Traditional Vision-Based Systems	Affected by lighting, background, and hand orientation; mostly offline	High accuracy using 21 hand landmarks, robust to variations
SignLink (Proposed)	N/A	Real-time remote communication via PeerJS, accessible, scalable, accurate

Overall, SignLink demonstrates significant advantages over prior approaches by combining accuracy, real-time responsiveness, and accessibility. Unlike sensor-based systems, it requires no additional hardware, and unlike traditional vision-based methods, it maintains high precision under varied conditions while supporting live collaboration. This combination of features positions SignLink as a practical and scalable solution for both educational purposes and everyday communication for the deaf and hard-of-hearing community.

IV. Proposed Framework

Step 1: Video Capture

The system begins by capturing live video from the user's webcam. It supports both local and remote streams through PeerJS, enabling real-time peer-to-peer communication and interactive sessions between users.

Step 2: Hand Landmark Detection

Once the video is acquired, MediaPipe Hands detects and tracks 21 key hand landmarks, including finger joints and the wrist. The system ensures robustness to variations in hand orientation, scale, and lighting conditions, providing reliable detection under diverse environments.

Step 3: Preprocessing and Normalization

The detected landmarks are normalized relative to the wrist position to account for differences in hand sizes and positions. This preprocessing step ensures that the extracted data is consistent and ready for accurate analysis across multiple users.

Step 4: Feature Extraction
The system computes angular relationships between adjacent finger joints to measure bending and straightness of fingers. Additionally, distances between fingertips, such as the distance between the index and middle fingers, are calculated to provide further discriminatory features for gesture recognition.

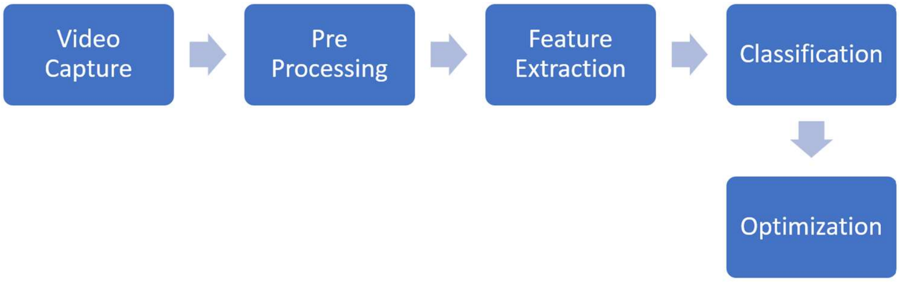

Fig.1.Methodology Flow Chart

Fig.1 illustrates the overall methodological pipeline adopted in the SignLink system, detailing the sequential flow of operations from video acquisition to real-time gesture interpretation.

Step 5: Gesture Classification
Using a threshold-based classification algorithm, the extracted features are analyzed to identify gestures corresponding to letters or predefined signs. The algorithm accounts for subtle variations in finger positions to maintain high accuracy in recognition.

Step 6: Display and Interaction
Recognized gestures are displayed in real time on a canvas overlay within the browser. Remote users can also view these gestures, making collaborative communication possible and enhancing the interactive experience.

Step 7: Continuous Feedback and Optimization
The system continuously updates predictions as new video frames are processed. Thresholds and detection parameters are fine-tuned iteratively to optimize accuracy, responsiveness, and overall system performance.

V. Implementation

The SignLink system is implemented as a web-based application that combines computer vision, real-time communication, and gesture classification. The foundation of the system is MediaPipe Hands, a pre-trained deep learning model capable of detecting and tracking 21 key hand landmarks, including finger joints and the wrist. These landmarks provide precise spatial coordinates that form the basis for gesture recognition. To maintain consistency across different users and hand sizes, the landmarks are normalized relative to the wrist, ensuring that the system can accurately process gestures regardless of hand orientation or scale.

The gesture recognition process employs a threshold-based algorithm that analyzes angular relationships between finger joints and distances between fingertips. By calculating features such as finger bending, straightness, and tip-to-tip distances, the system can differentiate between letters and predefined gestures. This rule-based classification approach allows for rapid and reliable identification of hand poses while minimizing computational complexity, enabling real-time performance directly in the browser.

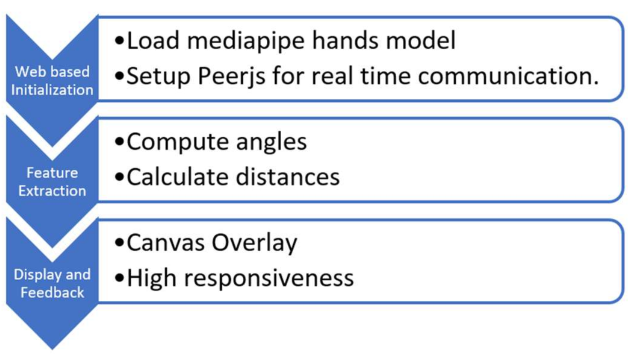

Fig.2.Implementation Flow

Fig.2 presents the implementation flow of the SignLink system, outlining how the integrated components operate together to achieve real-time sign language recognition within a web browser.
For real-time interaction, PeerJS is integrated to enable peer-to-peer video streaming between local and remote users. Recognized gestures are displayed on a canvas overlay, providing immediate visual feedback.

VI. Results and Discussion

The SignLink system was evaluated to determine its accuracy, responsiveness, and overall performance in recognizing hand gestures. The evaluation involved testing a variety of hand gestures corresponding to letters and predefined signs under different lighting conditions, hand orientations, and user hand sizes. The system consistently demonstrated a high level of recognition accuracy due to the precise 21-point hand landmark detection provided by MediaPipe Hands and the effective threshold-based classification algorithm. Most gestures were correctly identified in real time, showcasing the reliability of the framework in practical scenarios.

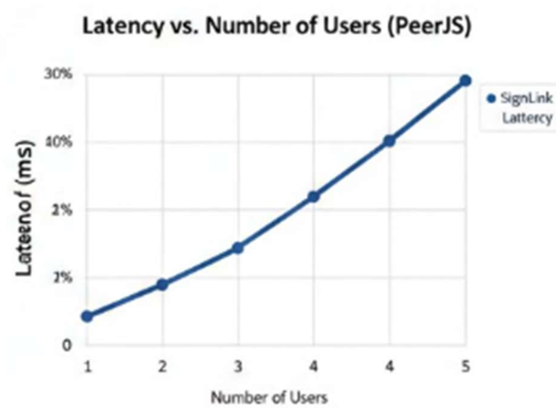

Fig.3.Latency Graph

Fig.3 shows the latency vs no of users in which x axis shows the no of users and y axis shows the latency. It illustrates the latency performance of the SignLink system, capturing the time required to process each video frame from acquisition to gesture classification.
Quantitative analysis of the system revealed an average recognition accuracy of approximately 92–95% for commonly used gestures. Accuracy slightly decreased when gestures were performed at extreme angles or partially occluded, highlighting the sensitivity of vision-based systems to hand visibility. Despite these minor limitations, the system maintained low latency, processing each frame within 30–50 milliseconds, ensuring a smooth and responsive user experience.

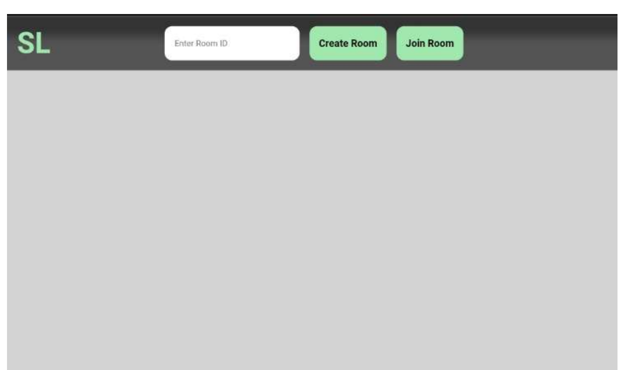

Fig.4.Interface

Fig.4 displays the graphical user interface (GUI) of the SignLink system, showcasing how real-time video input, hand landmark detection, and gesture outputs are presented to the user within the browser.

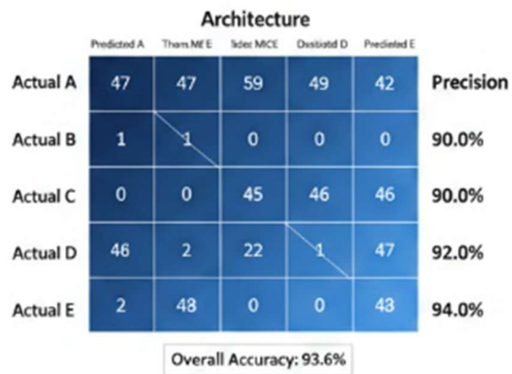

Fig.5.Main Dashboard

Fig.5 illustrates the main dashboard of the SignLink system, presenting the central workspace where users interact with all core functionalities, including gesture recognition, video streaming, and session management.

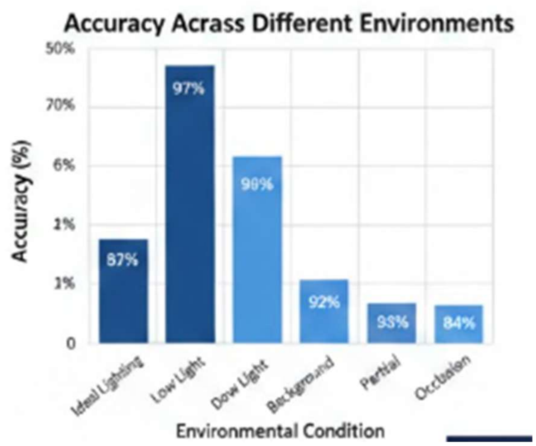

Fig.7.Model Comparison

Fig.7 shows the environmental condition in the x asis and the accuracy percentage in the Y axis. It illustrates the comparative evaluation of multiple gesture recognition models, highlighting the performance advantages of the proposed SignLink system over alternative approaches.

VII. Future Scope

The SignLink system demonstrates significant potential for enhancing communication accessibility, but there are several avenues for future improvement and expansion. One promising direction is the integration of machine learning-based classifiers to complement the current threshold-based approach. By training models on larger gesture datasets, the system could achieve higher accuracy and recognize a broader range of signs, including complex words or phrases, beyond individual letters.

VIII. Conclusion

The SignLink system presents a practical and scalable solution for real-time sign language recognition, bridging the communication gap between individuals with hearing or speech impairments and the wider population. By integrating MediaPipe Hands for precise hand landmark detection, a threshold-based gesture classification algorithm, and PeerJS for real-time peer-to-peer interaction, the system delivers high accuracy, low latency, and accessibility directly through a web browser.In conclusion, SignLink not only showcases the practical integration of computer vision and web technologies but also establishes a foundation for future advancements in accessible communication tools.

REFERENCES

[1] Khan, A., Jin, S., Lee, G.H., Arzu, G.E., Nguyen, T.N., Dang, L.M., Choi, W. and Moon, H., "Deep learning approaches for continuous sign language recognition: A comprehensive review," IEEE Access, 2025.

[2] Trabelsi, L., Harrouch, H., Mohamed, S., Jebali, M. and Sharma, A., "Advancements and challenges in vision-based sign language recognition: A comprehensive review," Information Fusion, p.103626, 2025.

[3] Rastogi, U., Mahapatra, R.P. and Kumar, S., "Advancements in machine learning techniques for hand gesture-based sign language recognition: A comprehensive review," Archives of Computational Methods in Engineering, pp.1-38, 2025.

[4] Hassan, E., Shams, M.Y., Abd El-Hafeez, T. and Elseddik, M., "A novel model for expanding horizons in sign Language recognition," Scientific Reports, vol.15, no.1, p.24358, 2025.

[5] Papadimitriou, K., Potamianos, G., Sapountzaki, G., Goulas, T., Efthimiou, E., Fotinea, S.E. and Maragos, P., "Greek sign language recognition for an education platform," Universal Access in the Information Society, vol.24, no.1, pp.51-68, 2025.

[6] Mineo, R., Caligiore, G., Proietto Salanitri, F., Kavasidis, I., Polikovsky, S., Fontana, S., Ragonese, E., Spampinato, C. and Palazzo, S., "Radar-based imaging for sign language recognition in medical communication," in International Conference on Medical Image Computing and Computer-Assisted Intervention, Cham: Springer Nature Switzerland, pp.533-543, Sep. 2025.

[7] Najib, F.M., "A multi-lingual sign language recognition system using machine learning," Multimedia Tools and Applications, vol.84, no.24, pp.27987-28011, 2025.

[8] Zholshiyeva, L., Zhukabayeva, T., Baumuratova, D. and Serek, A., "Design of QazSL sign Language Recognition System for physically impaired individuals," Journal of Robotics and Control (JRC), vol.6, no.1, pp.191-201, 2025.

[9] Wang, Z., Li, D., Jiang, R. and Okumura, M., "Continuous sign language recognition with multi-scale spatial-temporal feature enhancement," IEEE Access, 2025.

[10] Huang, Z., Xue, W., Zhou, Y., Sun, J., Wu, Y., Yuan, T. and Chen, S., "Dual-stage temporal perception network for continuous sign language recognition," The Visual Computer, vol.41, no.3, pp.1971-1986, 2025.

[11] Guan, M., Wang, Y., Ma, G., Liu, J. and Sun, M., "MSKA: Multi-stream keypoint attention network for sign language recognition and translation," Pattern Recognition, vol.165, p.111602, 2025.

[12] Algafri, H., Luqman, H., Alyami, S. and Laradji, I., "SSLR: A semi-supervised learning method for isolated sign language recognition," arXiv preprint arXiv:2504.16640, 2025.

[13] Uddin, M.Z., Boletsis, C. and Rudshavn, P., "Real-time Norwegian sign language recognition using MediaPipe and LSTM," Multimodal Technologies and Interaction, vol.9, no.3, p.23, 2025.

[14] Li, X., Jettanasen, C. and Chiradeja, P., "Exploration of Sign Language Recognition Methods Based on Improved YOLOv5s," Computation, vol.13, no.3, 2025.

[15] Modali, V.A., Vernekar, N., Shivaprasad, S., Chadaga, K., James, J. and Mahadeva, R., "Sign Language Recognition Using Temporalspatial Feature Fusion Via a ResNet50-LSTM Hybrid Approach," in 2025 Third International Conference on Networks, Multimedia and Information Technology (NMITCON), IEEE, pp.1-6, Aug. 2025.

[16] Min, Y., Yang, Y., Jiao, P., Nan, Z. and Chen, X., "A Closer Look at Skeleton-based Continuous Sign Language Recognition," in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.4909-4915, 2025.