IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Impact Factor: 5.731
Volume 5, Issue 12, December 2025

Real-Time Sign Language Interpretation Using CNN and LSTM Networks

Divya U H, Dr. Heena Kousar, Rashmi T V, Harshitha R, Chandu V C, Divya S
Assistant Professor, Department of Computer Science & Engineering,
Professor, Department of Computer Science & Engineering
Assistant Professor, Department of Computer Science & Engineering
Students, Department of Computer Science and Engineering
East Point College of Engineering and Technology, Bengaluru, India

Abstract: *For those with speech and hearing problems, sign language is an essential communication tool. Nonetheless, there is a substantial barrier separating sign language users from non-signers, which restricts accessibility in everyday interactions, healthcare, and education. The DL (deep learning)-based sign language interpretation system described in the present paper utilises LSTM (Long Short-Term Memory) networks for dynamic sequence classification, CNNs (Convolutional Neural Networks) for static gesture identification. The system has a user-friendly interface utilizing Streamlit for smooth interaction and uses Mediapipe for real-time hand landmark detection. Conventional rule-based recognition systems are outperformed by experimental evaluations, which show an accuracy of 88% for dynamic sequences and 92% for static gestures. The suggested method improves accessibility and inclusivity by offering real-time gesture translation into text, facilitating efficient communication for the community of people who are hard-of-hearing and deaf.*

Keywords*:* Sign Language Recognition, Deep Learning, CNN, LSTM, Gesture Interpretation, Accessibility

I. INTRODUCTION

The majority of people who use sign language, an organized visual language, are deaf or hard of hearing. Its conveyed using body movement, hand gestures, facial expressions. In spite of its significance, there is still communication gap among non-signers and sign language users, which causes problems in everyday interactions, healthcare, and education. Traditional methods, such as human interpreters and text-based systems, offer solutions but are often expensive, inaccessible in real-time, unable to capture full expressiveness of sign language. Need for an automated and efficient sign language interpretation system has become increasingly important to foster inclusivity or accessibility [1]. Real-time gesture recognition systems that can correctly understand sign language have been made possible by recent developments in artificial intelligence (AI) and DL [2][3]. In this paper, a DL-based system that combines LSTM networks for dynamic gesture classification and CNNs for static gesture identification is presented. System provides user-friendly interface with Streamlit, allowing for smooth interaction, and uses Mediapipe for real-time hand landmark detection. Suggested system effectively bridges communication gap among non- signers and sign language users by utilizing these strategies to obtain high accuracy in identifying both individual letters and complex phrases.

II. BACKGROUND STUDY

Using facial expressions, hand gestures, sign language body movements is vital communication tool for people with speech and hearing problems. Accessibility in professional, clinical, and educational environments is difficult, nevertheless, because communication hurdles occur when communicating with non-signers. Development of automatic sign language recognition systems is essential because traditional options, such text- based systems and human translators, are sometimes expensive and unavailable in real-time. The accuracy of gesture detection has greatly increased because to developments in DL and computer vision. Using computer vision methods and a DL model, the authors of [5]

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 5, Issue 12, December 2025
Impact Factor: 5.731

created system which can translate sign language videos in both speech and text. Their approach emphasized real-time translation, aiming to enhance communication accessibility for hearing-impaired individuals by integrating gesture recognition with audio output for broader usability.

A thorough examination of AI technologies used in sign language was given by Papastratis et al. [6], who concentrated on computer vision, DL, and multimodal techniques that encompass the analysis of hand, face, and body movements. The paper highlights current challenges in real-time recognition and emphasizes the need for robust, scalable AI systems to enable inclusive human- computer interaction for deaf community.

Using CNN and LSTM architectures, Kothadiya et al.

[7] presented DeepSign, a DL-based system for sign language detection and identification. Their model shows promise for real-time applications in assistive communication technology and successfully recognizes both static and dynamic gestures, with high accuracy on benchmark datasets. The main drawback of the DeepSign system is its limited generalization to real-world environments and diverse signing styles, as it was primarily tested on controlled datasets.

While LSTM networks are better at capturing sequential relationships in dynamic gestures, CNNs are better at identifying static hand gestures. Additionally, real-time hand tracking technologies like Mediapipe and OpenCV enhance precision in gesture detection. By combining these methods, the suggested system aims to improve accessibility for hard-of-hearing or deaf community by offering a precise, real-time sign language interpretation tool.

III. PROPOSED METHODOLOGY

Suggested system, seen in figure 1, integrates CNNs for static gesture detection with LSTM networks for dynamic gesture classification, utilizing DL techniques for real-time sign language interpretation. With help of Mediapipe, system can parse camera video input, identify hand landmarks, or translate motions into text outputs. The method guarantees great accuracy in identifying both discrete hand signs and continuous gesture sequences by utilizing combination of sequence modeling and computer vision. Methodology consists of multiple stages: pre-processing, feature extraction, sequence modelling, classification, and output generation.

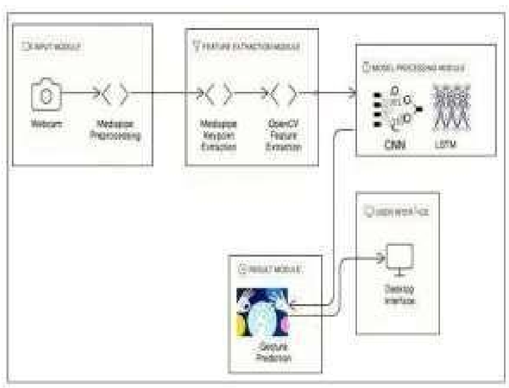

Figure 1. The proposed methodology

Pre-processing stage: This stage involves capturing video frames from a webcam and converting them into grayscale images to reduce computational complexity. Mediapipe is then applied to detect hand landmarks, including finger positions and palm orientation, ensuring precise feature extraction. This step improves system's ability to track gestures effectively, even in varying lighting conditions.

Feature extraction stage: By detecting patterns in hand forms and orientations, the system analyzes static gestures using CNN models. This enables the algorithm to accurately identify single letters and simple words. Extracted features are then structured into a format suitable for further processing in the next stage.

Sequence modelling stage: The technology analyzes static motions by spotting patterns in hand forms and orientations using CNN models. This makes it possible for the algorithm to accurately identify single letters and basic words. This improves the system's ability to interpret complete sentences rather than just isolated signs.

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 5, Issue 12, December 2025
Impact Factor: 5.731

Classification stage: It assigns recognized gestures to their corresponding text representations. Using a trained deep learning model, the system matches extracted gesture features with predefined sign language vocabulary, ensuring correct identification of hand signs. The model continuously learns and improves recognition accuracy with additional training data.

Output generation: This stage displays the recognized gestures as text output on a user-friendly interface developed using Streamlit. The system provides real- time feedback, ensuring a seamless interaction experience. By using this method, the suggested solution can close communication gap among non-signers and sign language users, increasing accessibility and inclusivity.

IV. EXPERIMENTAL ANALYSIS

Accuracy, real-time performance, as well as capacity to identify both static and dynamic motions were the main criteria used to assess the suggested sign language interpreting system. Dataset of different sign language motions, such as letters, sentences, and phrases, was used to test the system. The below given figures give insight into how the signs are recognized, and they generate the text.

The figure 2 demonstrates recognition of the ASL (American Sign Language) alphabet for letter "P" using deep learning techniques. The system processes hand gesture images with CNN and LSTM models to classify the sign accurately, enhancing accessibility for sign language users.

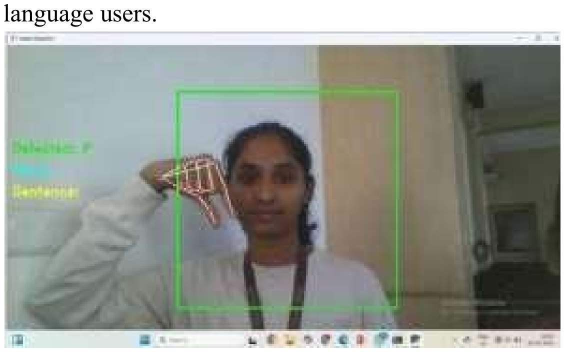

Figure 2. ASL Alphabet Recognition for the Letter "P"

Figure 3. ASL-Based Sentence Formation Using Gesture Recognition

The figure3 demonstrates the conversion of ASL alphabet gestures into complete sentences using DL- based recognition system. Model assembles words into structured sentences, with the enter key facilitating sentence formation, improving real-time sign language communication.

IJETIR

ISSN (Online) 2583-0554

IJETIR
Impact Factor: 5.731

International Journal of Emerging Technologies and Innovative Research (IJETIR)

Volume 5, Issue 12, December 2025

Figure 4. Recognition for the word “All the best”

Figure 5. Recognition for the word “I Love You”

The figure 4 and figure 5 collectively demonstrate the model’s ability to recognize as well as translate ASL gestures into meaningful text outputs, enhancing accessibility for non-verbal communication.

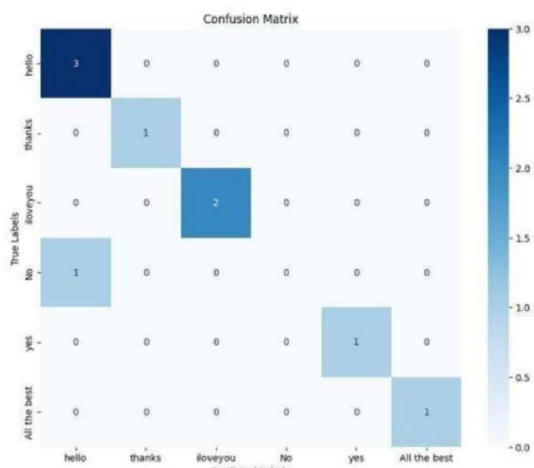

Figure 6. Confusion Matrix

The model shows higher degree of accuracy in recognizing different sign language movements, with a test accuracy of 88.89%. Figure 6 illustrates model's performance by displaying number of indicators that have been successfully and wrongly identified. While off-diagonal elements show misclassifications, diagonal elements show accurate predictions. The matrix showed that model correctly identified multiple instances of "hello," "I love you," and other gestures, with only a few misclassifications. The loss value of 0.6584 suggests that the model has room for further optimization to reduce errors.

V. CONCLUSION

With its efficient, scalable, and real-time gesture recognition approach, the proposed DL-based sign language interpretation system significantly improves accessibility for those with speech and hearing impairments. The method guarantees good recognition and text-to-sign language accuracy by using CNNs for static gestures and LSTM networks for dynamic gestures. Its usability in practical applications like public services, healthcare, and education is further improved by the incorporation of Mediapipe for hand landmark detection and an intuitive user interface.

Future enhancements will focus on expanding the vocabulary set to involve wider range of sign languages, incorporating regional or linguistic variations to improve adaptability. Additionally, optimizing real-time processing speed through model compression techniques, involving quantization and pruning, will enable smoother execution, especially on low-power devices. A major goal is to deploy the system on mobile platforms, making it more accessible to broader user base. However, numerous challenges need to be addressed, involving ensuring accurate recognition across diverse signing styles, handling variations in lighting conditions and background noise, and reducing computational complexity for efficient performance on mobile devices. Moreover, improving real-time latency is crucial to ensure seamless communication, making further research into edge computing and lightweight deep learning models essential. Addressing these challenges will enhance the overall efficiency and practicality of the system, making it a more inclusive and widely applicable tool for sign language interpretation.

REFERENCES

[1] H. Liang and J. Yuan, "Gesture Recognition Using Skeleton-Based Features and Deep Learning," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 11, pp. 3245-3259, Nov. 2018.

[2] A. Sharma and M. Bansal, "Review on sign language recognition in context of vision-based and deep learning," Measurement: Sensors, vol. 24, 2022, Art. no. 100507. https://doi.org/10.1016/j.measen.2022.100507

[3] M. A. Asad, S. A. Khan, M. A. Khan, S. A. R. Zaidi and H. Hwang, "Sign Language Recognition: A Deep Survey," Expert Systems with Applications, vol. 164, 2021, Art. no. 113794.htps://doi.org/10.1016/j.eswa.2020.113794

[4] Supriya Agre1, Sohan Wasker, Aman Vashishtha, Harshavardhan Latkar, Atharva Kanse, "Real-time Conversion of Sign Language to Text and Speech, and vice-versa", Volume 10, Issue 11, JETIR, (2023).

[5] G. Sekhar Reddy, A. Sahithi, P. Harsha Vardhan, P. Ushasri, "Conversion of Sign Language Video to Text and Speech" International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 10, Issue V, May (2022).

[6] Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K., & Daras, P., Artificial intelligence technologies for sign language. Sensors, 21(17), 5843 (2021).

[7] Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil- González, A. B., & Corchado, "Deepsign: Sign language detection and recognition using deep learning", Electronics, 11(11), 1780, (2022).

[8] Mahesh Kumar N B1, "Conversion of Sign Language into Text ", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9 (2018).

[9] Al-Qurishi, M., Khalid, T., & Souissi, R., Deep learning for sign language recognition: Current techniques, benchmarks, and open issues. IEEE Access, 9, 126917- 126951. (2021).

[10] Adeyanju, I. A., Bello, O. O., & Adegboye, M. A., Machine learning methods for sign language recognition: A critical review and analysis. Intelligent Systems with Applications, 12, 200056, (2021).

[11] Rastgoo, R., Kiani, K., & Escalera, Sign language recognition: A deep survey. Expert Systems with Applications, 164, 113794, (2021).

[12] Sharma, A., Sharma, N., Saxena, Y., Singh, A., & Sadhya, D, Benchmarking deep neural network approaches for Indian Sign Language recognition. Neural Computing and Applications, 33(12), 6685-6696, (2021).