International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

AI Based Sign Language to Text Converter

Gangadhar Hongal[1], Iranna Agni[2], Mahantesh Halingali[3], Rohit More[4], Dr. Nagaraj Bhat[5]
* KLS Vishwanathrao Deshpande Institute of Technology Haliyal, India Department of Electronics and Communication

---***---

Abstract- Communication between hearing-impaired individuals and the general public is often limited due to the lack of widespread knowledge of sign language. Recent advancements in artificial intelligence have enabled the development of gesture-recognition technologies that can help overcome this gap by translating hand signs into readable text. In this work, an AI-based, real-time sign language conversion system is proposed to recognize hand gestures and map them to their corresponding textual meanings. The design of the system follows a structured A systematic process that involves collecting data, segmentation, data cleaning and preparation, identification of key features, and training of the model gesture recognition. Each stage contributes to refining the captured gesture data, ensuring that the model learns clear and meaningful patterns for accurate prediction.

The goal of the proposed system is to offer a fast, dependable, and accessible communication tool that supports hearing- and speech-impaired individuals in public, educational, and assistive environments. By efficiently interpreting commonly used signs, the system can significantly enhance interaction and reduce communication barriers. Experimental results validate that the model performs reliably in real-time scenarios, demonstrating its capability to deliver consistent outcomes. Overall, this AI-driven promoting inclusive communication and contributing to the development of assistive technologies that enhance social participation for individuals with communication challenges.
.

1. INTRODUCTION

Facial expressions are among the most natural ways through Communication is among the most fundamental needs of human interaction. People find all sorts of ways to get their messages across. They use spoken words and body movements. They rely on facial expressions and plenty of subtle cues too. For folks who cannot hear or speak, sign language turns into their primary way to express themselves. It works as a full language on its own. It has its own grammar and structured hand shapes. It includes various visual parts as well.

Still, sign language often does not receive much recognition beyond the community of people with hearing issues. That unfamiliarity creates real problems in everyday dealings. Those issues show up in schools and hospitals. They appear in workplaces and public spots. They even happen right at home sometimes.

Things are shifting now with advances in Artificial Intelligence and computer vision tech. These newer setups can analyze hand gestures pretty accurately. They make it feasible to turn sign language into written text or spoken words. The tech helps out individuals with hearing disabilities a lot. It also pushes for a society that feels more open and easier for everyone to navigate. The motivation for this project arises from the real-world Need for user-friendly tools that enable seamless interaction between sign language users and non- signers.

The present work focuses on designing a practical, real-time AI-based sign language to text converter that can be used on everyday devices such as laptops or mobile cameras. The system captures hand gestures, preprocesses them to enhance clarity, extracts meaningful features, and then classifies each gesture using a trained model. The output is displayed as readable text, enabling a smooth conversation flow between users. This approach takes inspiration from established image- processing methodologies—as seen in other vision-based systems such as facial expression detection while extending them into the domain of sign language interpretation.

Overall, the introduction of AI into gesture recognition has the potential to make communication more inclusive and accessible. The proposed system intends to contribute toward a society where hearing-impaired individuals can interact freely without depending on interpreters or specialized services. By developing a lightweight, accurate, and easy-to- use sign-language recognition model, this research shows how technology can meaningfully support human connection.

2. LITERTATURE REVIEW

The Research in sign language recognition has grown quickly because of growing demand for assistive technologies that help communication between hearing- impaired individuals and the general population. Over the years, researchers have explored a wide range of techniques from traditional image-processing-based approaches to advanced deep-learning architectures—to improve the accuracy, flexibility, and real-time performance of gesture recognition systems.

Abdullah Baihan and team (2024) proposed a modified deep-learning rapidly due to the rising need for assistive technologies that aid Memory (LSTM) units to recognize continuous sign sequences. Their work emphasizes the challenges of signer variability, background complexity, and movement speed, and introduces a hybrid optimization method to improve model convergence and recognition accuracy. This study is significant because it highlights how temporal modelling in LSTM networks can capture the dynamic nature of sign gestures, which often involve sequential hand motions.

Another important study by Shahad Thamear Abd Al-Latief et al (2024) provides a broad comparative analysis of various deep-learning models applied to sign language recognition. Their evaluation includes CNN-based classifiers, 3D-CNN architectures, and recurrent models, showcasing how each technique performs under different lighting conditions, gesture complexity, and dataset sizes. The authors conclude that deep learning models clearly outperform conventional handcrafted feature extractors especially in tasks involving fine finger movements and multi-frame gesture sequences. This work forms a strong foundation for understanding model selection and optimization in AI-based sign recognition systems.

F. M. Najib (2024) investigated the use of machine learning algorithms combined with AI techniques to interpret sign language gestures. The study examines how hand-shape descriptors, geometric features, and motion-based cues can be integrated with ML classifiers to achieve reliable recognition. This research is particularly relevant because it demonstrates how hybrid systems—those that merge classic computer vision with modern AI—can provide strong performance even with limited computational resources.

Furthermore, advances in deep learning have allowed models to grasp complex hand shapes without explicit feature engineering. Many recent studies confirm that CNN-based models deliver strong performance across different backgrounds, lighting conditions, and user differences making them highly suitable for assistive communication tools.

Collectively, the reviewed literature shows a clear shift toward AI-driven, end-to-end learning systems that can operate in real time with high accuracy. These studies highlight essential challenges such as background noise, dynamic gestures, signer diversity, and low-light conditions.

At the same time, they demonstrate significant progress in building reliable recognition models suitable for practical deploymentThe findings from these studies directly contribute to the development of this project, which seeks to build a practical, real-time sign language– to–text converter using AI techniques.

3. DESIGN METHODOLOGY

The design methodology of the proposed AI-Based Sign Language to Text Converter follows a structured, step-by- step pipeline that ensures reliable gesture recognition and accurate text conversion. The system is divided into several functional blocks, beginning from gesture acquisition and ending with real-time text output. Each stage plays a specific role in transforming raw hand images into meaningful information.

3.1 System Flow Diagram

The complete working model of the system is represented in the flow diagram below.

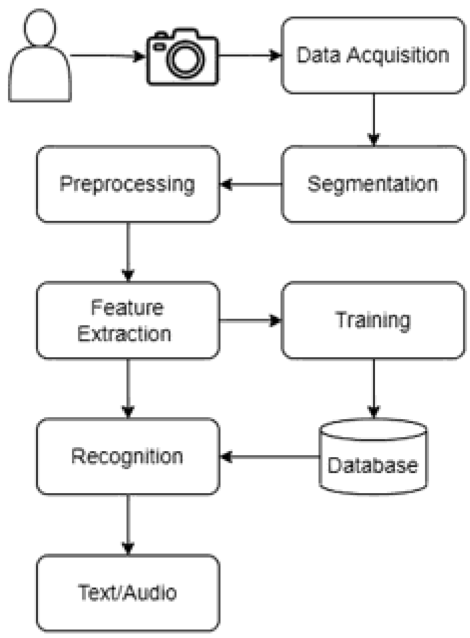

Fig 3.1.1: Flow Diagram

This diagram outlines the major steps used in the project, including data collection, segmentation, preprocessing, feature extraction, model training, recognition, and final output.

3.2 Data Acquisition

The system starts by capturing hand gesture images through a live webcam or mobile device camera. The user places their hand in front of the camera, and continuous video frames are recorded. Proper lighting and a clear background help the system detect gestures more accurately. The captured frames are then forwarded for further processing.

3.3 Segmentation

After acquiring the image, the next task is to isolate the hand region from the background. Segmentation helps focus only on the relevant portion of the frame. Techniques such as skin-color filtering, background subtraction, or contour detection are used to extract the hand area. his step makes sure that irrelevant objects or background elements do not interfere with recognition.

3.4 Preprocessing

The segmented hand image undergoes several preprocessing steps to improve clarity and reduce noise:

- Conversion to grayscale
- Noise reduction using Gaussian or median filtering
- Normalization for brightness and contrast
- Resizing the image to a fixed dimension

These preprocessing operations make the input more consistent and help the recognition model perform better.

3.5 Feature Extraction

Feature extraction focuses on identifying important characteristics of the hand gesture. The system analyzes:

- Hand contours and shape
- Number and position of fingers
- Orientation and movement patterns
- Key geometric features

extracted features act as distinguishing markers that help classify different signs accurately.

3.6 Training and Classification

The extracted features are used to train an AI model, usually a Convolutional Neural Network (CNN). During training, the model learns the structure of each gesture and develops the ability to differentiate one sign from another.

In the classification stage, live gesture frames are compared with the learned patterns, and the model predicts the most suitable gesture category.

3.7 Text Output Generation

After a gesture is identified by the model, it is immediately translated into text and shown on the screen, enabling real- time interaction. This immediate feedback aids communication between sign language users and non-users. If needed, the output can also be converted to speech, providing an extra level of accessibility.
The overall workflow of the system is designed so that each stage supports the next, beginning with gesture capture and ending with the generation of the textual output. By following a clear and organized sequence—data collection, segmentation of the hand region, preprocessing, feature extraction, model training, and final classification—the system delivers consistent performance and reliable accuracy. This organized approach enhances recognition reliability and also makes the system simpler to implement and practical for daily use, particularly for helping individuals with hearing or speech difficulties.

4. RESULT AND DISCUSSION

The system was tested using a live webcam so its real-time gesture recognition abilities could be closely examined. Throughout the evaluation, the camera captured a continuous stream of frames, and the model automatically isolated the hand, identified important landmark points, and processed them to classify the gesture being shown. This live setup made it possible to observe how well the system handled natural and unscripted hand movements.

During testing, the model consistently identified commonly used signs—including "Thank You," "I Love You," "No," and "Hello"—with high confidence. The display of landmark points on the detected hand also helped illustrate how accurately the system followed finger positions and joint movements, demonstrating the reliability of the gesture-tracking mechanism.

Beyond accuracy, the responsiveness and stability of the system were also evaluated. Even when lighting or background conditions shifted slightly, the model continued to track gestures smoothly. Small changes in hand angle, distance from the camera, or movement speed had little impact on recognition quality, showing that the system could handle minor variations without losing performance.

These observations suggest that the system is practical for everyday use, where gestures are rarely perfect or repeated in the same wayOverall, the results show that the proposed solution is reliable, user-friendly, and effective for supporting communication between sign language users and the general public.

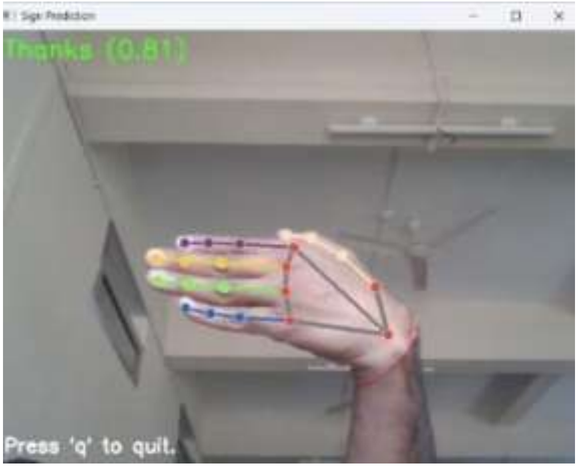

Fig 4.1.1: Recognized Gesture – "Thanks" (Confidence: 0.81)

The displayed result illustrates the system recognizing the "Thanks" gesture. The hand's landmark points were accurately detected, allowing the model to classify the gesture with a confidence score of 0.81. The clear mapping of the fingers made it easier for the classifier to differentiate this sign, even under normal indoor lighting conditions.

Fig 4.1.2: Recognized Gesture – "I Love You" (Confidence: 0.86)

The "I Love You" gesture, which has a distinctive combination of extended and folded fingers, was also correctly identified by the system. The landmark points were detected consistently, and the model produced a confidence score of 0.86, reflecting strong recognition performance. Its ability to accurately distinguish between raised and bent fingers played a key role in correctly interpreting this uniquely shaped gesture.

Fig 4.1.3: Recognized Gesture – "No" (Confidence: 0.96)

In this example, the system identified the "No" gesture with a confidence score of 0.96, indicating very high accuracy. The model captured the curved hand shape and the specific finger positioning with precision. This gesture consistently produced one of the highest confidence values during testing, demonstrating the model's strong ability to interpret more compact hand formations.

Fig 4.1.4: Recognized Gesture – "Hello" (Confidence: 0.82)

For the "Hello" gesture, the system accurately identified the open-palm posture and correctly mapped all the finger landmarks. Even when the hand shifted slightly in angle or distance from the camera, the model continued to recognize the gesture reliably, producing a confidence score of 0.82. This demonstrates the system's ability to handle broad, open-hand movements effectively.

4.2 Overall Performance

Across all tests, the system maintained stable detection and produced predictions with confidence values mostly above 0.80. The landmark tracking worked smoothly in real time, and the model responded without noticeable delay. Gestures with clear and distinct finger patterns gave higher accuracy, while gestures with subtle differences sometimes produced slightly lower confidence values.

4.3 Discussion

The results indicate that the AI-based system can accurately recognize basic sign language gestures. Good lighting and proper hand placement improved prediction performance. The system performed best for gestures with distinct finger positions. Minor performance drops were observed when the hand was partially outside the camera frame or when lighting was very low. Overall, the system worked reliably and demonstrated that AI can be effectively used for gesture-to-text conversion.

4.4 Real-Time System Response

During testing, the system showed smooth and continuous real-time performance. The webcam captured frames at a stable rate, and the AI model processed each gesture without noticeable delay. The landmark detection remained consistent even when the hand moved slightly, allowing the system to update predictions instantly. This quick response makes the prototype suitable for real- time communication scenarios where immediate output is required. The system also displayed each gesture's confidence score, helping evaluate how accurately the model understood the sign.

4.5 Discussion on System Limitations

Although the system performed well overall, some limitations were noted during testing. Accuracy slightly decreased in low-light conditions, as the hand's landmark points were harder for the model to detect clearly. Recognition also became less stable when gestures were performed either very close to or far from the camera. In a few instances, fast movements or fingers overlapping each other led to brief misclassifications. These issues can be minimized by expanding the training dataset to include more varied examples and by enhancing the model's ability to adapt to different lighting conditions.

5. CONCLUSION

The AI-based Sign Language to Text Converter created in this project is able to recognize commonly used hand gestures and translate them into readable text with good accuracy. It works in real time and performs consistently under standard indoor lighting. Using AI models and gesture tracking, smooth communication between sign language users and non-signers is achievable.

The outcomes achieved match the goals set at the beginning of the project accurate gesture detection, real-time response, and a user-friendly working prototype. This project demonstrates that AI can significantly enhance communication accessibility for individuals with hearing or speech impairments. With further improvements, the system can be expanded to recognize more gestures and be implemented on mobile or practical platforms as mentioned in the project objectives.

4. REFERENCES

[1] A. Alkhoraif, S. Almalki, B. Alnujaidi. "Transformer and Ensemble-Based Models for Word-Level Sign Language Recognition" 2025.

[2] Shahad Thamear Abd Al-Latief et al., "Deep Learning for Sign Language Recognition," 2024.

[3] Abdullah Baihan, Alutaibi, Mohammed Alshehri, Sunil Kumar Sharma, "Sign Language Recognition Using Modified Deep Learning Network and Hybrid Optimization," 2024.

[4] S. Kumar, R. Patil, A. Jain. "Real-Time Vision-Based Indian Sign Language Recognition System" 2024.

[5] D. Kumari, R. S. Anand. "Isolated Video-Based Sign Language Recognition using Hybrid CNN-LSTM with Attention Mechanism" 2024.

[6] Y. S. N. Rao, K. Srinivas, A. Reddy. "Dynamic Sign Language Recognition and Translation – A Systematic Review" 2024.

[7] S. Tan, H. Wijaya, K. Ang. "A Review of Deep-Learning Approaches for Sign Language Recognition and Translation" 2024.

[8] R. Kumar, S. Mehta, P. Reddy. "A Real-Time American Sign Language Gesture Recognition System using MediaPipe and Convolutional Neural Networks" 2023.