International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 04 April-2026 | Impact Factor: 3.5

Sign Language Detection Using Deep Learning Techniques

Mrs. Sarla Devi[*1], L. Vivek[*2], Y. Srinidhi[*3], K. Shashank[*4], A. Deekshith[*5]

[*1]Assistant Professor, Department of CSE (Data Science), ACE Engineering College, Hyderabad, Telangana, India

[*2,*3,*4,*5]III B.Tech Students, Department of CSE (Data Science), ACE Engineering College, Hyderabad, Telangana, India.

How to Cite this Article:

Vivek, L., Srinidhi, Y., Shashank, K. & Deekshith, A. (2026). Sign Language Detection Using Deep Learning Techniques. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04). https://doi.org/10.55041/ijcope.v2i4.245

https://doi.org/10.55041/ijcope.v2i4.245

ABSTRACT

Communication between hearing-impaired individuals and others remains a major challenge due to the lack of understanding of sign language. The proposed Sign Language Detection System addresses this issue by using deep learning techniques to recognize hand gestures in real time and convert them into both text and speech output. The system uses Convolutional Neural Networks (CNN) implemented with TensorFlow and Keras for accurate gesture classification, while OpenCV is used for image capture and preprocessing through a standard webcam. Unlike traditional methods that rely on handcrafted features or special hardware, this system automatically learns visual patterns from data and provides a complete communication solution by integrating text-to-speech functionality. The system is efficient, cost-effective, and user-friendly, making it suitable for real-world applications and assistive communication.

Keywords: Sign Language Detection, Deep Learning, Convolutional Neural Networks (CNN), Computer Vision, Gesture Recognition, Text-to-Speech, TensorFlow, Keras, OpenCV, Assistive Technology

INTRODUCTION

Sign language is a vital form of communication used by hearing-impaired and mute individuals, relying on hand gestures, facial expressions, and body movements to convey meaning. However, most people are not familiar with sign language, which creates a significant communication gap between sign language users and the rest of society. This gap makes everyday interactions such as education, healthcare access, and public communication difficult, highlighting the need for an effective and accessible solution.

To address this problem, sign language recognition systems use computer vision and deep learning techniques to interpret hand gestures captured through a camera. Traditional methods relied on handcrafted features like skin color detection, contour extraction, or sensor-based devices such as data gloves, which are often costly, less accurate, and sensitive to environmental conditions. In contrast, deep learning approaches, particularly Convolutional Neural Networks (CNN), automatically learn complex visual patterns from image data, enabling more accurate and robust gesture recognition in real-time scenarios.

The proposed system uses CNN models implemented with TensorFlow and Keras, along with OpenCV for real-time image capture and preprocessing. It recognizes hand gestures through a standard webcam and converts them into both text and speech output using text-to-speech techniques. This approach provides a low-cost, efficient, and user-friendly solution that not only improves communication accessibility but also demonstrates the practical application of deep learning in assistive technologies.

Authors	Title / Reference	Technology Used	Limitations	Year
C M Naveen Kumar & Vanitha A.	Deep Learning Based Recognition of Sign Language	MediaPipe, CNN	Struggles with continuous real-time video processing; highly dependent on consistent backgrounds and lighting	2024
Natarajan et al.	Hybrid Deep Neural Architecture for Sign Language Recognition	Computer Vision, CNN, Bi-LSTM	High computational cost makes mobile deployment difficult; latency issues during live translation.	2023
Muthu Mariappan H. et al.	Real-Time Sign Language Recognition System	Fuzzy C-Means, OpenCV	Achieved only 75% accuracy; struggled with complex gestures.	2021

Anitha et al.	Gesture Based Sign Language Recognition	Deep Neural Networks	Primarily focused on recognition; lacks seamless translation.	2025
Hira Hameed	Privacy-Preserving British Sign Language Recognition	Radar, VGG16	Privacy-focused but limited to a small set of emotional signs.	2022
Jeet Debnath	Real-Time Gesture Based Sign Language Recognition System	CNN, Computer Vision	Lacks seamless translation to audio; focus is primarily on recognition accuracy.	2024
D. Shofia Priyadharshani	Integrated Voice Output Functionality for Sign Learners	Deep Learning, Voice Output	Identifies a lack of seamless transitions between different sign languages for non-signers.	2024
M. Soundarya	Sign Language Recognition Using Machine Learning	LSTM Algorithms	Traditional camera-based systems suffer from poor lighting and privacy concerns.	2024
Piyusha V. & Sanket D. (et al.)	Detection and Interpretation of Indian Sign Language Using LSTM Networks	LSTM Algorithms	Translates isolated words, not continuous sentences. Accuracy is highly sensitive to lighting, camera angles, and hand orientation.	2023

II. METHODOLOGY

The development of the Sign Language Detection System follows a structured approach covering all stages from data preparation to real-time deployment.

Requirement Analysis:
The project begins by identifying the key functionalities such as real-time gesture detection, image processing, gesture classification, and conversion of recognized gestures into text and audio. The system is designed to work using a standard webcam without requiring specialized hardware, ensuring accessibility and low cost.

Data Collection and Preprocessing:
A dataset of hand gesture images representing different sign language symbols is collected. The images are preprocessed using OpenCV techniques such as resizing, normalization, and background noise reduction to ensure consistency and improve model performance. Data augmentation techniques may also be applied to increase dataset diversity and prevent overfitting.

Model Design and Training:
A Convolutional Neural Network (CNN) is designed using TensorFlow and Keras to automatically learn features from gesture images. The model is trained on the prepared dataset to classify different hand gestures accurately. Training involves multiple epochs, optimization, and validation to achieve high accuracy and generalization.

System Implementation:
The system integrates the trained CNN model with OpenCV for real-time video capture. Frames from the webcam are processed and passed to the model for prediction. The predicted gesture is then converted into corresponding text output.

Text-to-Speech Conversion:
The generated text is further processed using text-to-speech modules to produce audio output. This allows the system to provide both visual and auditory feedback, making communication more effective.

Testing and Deployment:
The system is tested for accuracy, response time, and robustness under different lighting and background conditions. Integration testing ensures smooth interaction between modules. The final system is deployed as a real-time application that can run on standard computing devices.

III. MODEL EVALUATION

This section presents the evaluation of the Sign Language Detection System across multiple dimensions including functionality, usability, accuracy, and performance. The system was tested using real-time webcam input and predefined gesture datasets to validate the correctness of gesture recognition and output generation.

Evaluation Aspect	Performance
Functional Evaluation	All major features such as real-time gesture capture, image preprocessing, gesture prediction, and text-to-speech conversion work correctly in test scenarios.
Usability Evaluation	The system provides a simple and user-friendly interface with real-time feedback, making it easy for users to interact without prior technical knowledge.
Recognition Accuracy	The CNN model accurately classifies hand gestures under controlled conditions and shows reliable performance across different test samples.
Processing Speed	Gesture recognition and output generation are performed in near real-time with minimal delay, ensuring smooth communication flow.
Output Quality	The system generates clear and correct text output along with understandable audio output using text-to-speech conversion.
Error Handling	The system handles unclear gestures, invalid inputs, and detection failures by providing appropriate feedback or ignoring incorrect predictions.
Robustness	The system performs well under moderate variations in lighting and background, though extreme conditions may affect accuracy.

3.1 Comparison and Considerations

The Sign Language Detection System provides a significant improvement over traditional communication methods that rely on human interpreters or manual learning of sign language. Compared to earlier approaches that used handcrafted features or sensor-based devices, the proposed system uses deep learning techniques to automatically learn gesture patterns, resulting in better accuracy and flexibility. The use of a standard webcam eliminates the need for expensive hardware, making the system more accessible and cost-effective.

Unlike traditional image processing methods that struggle with complex backgrounds and variations, the CNN-based approach improves recognition capability by learning features directly from data. The integration of both text and audio output makes the system more practical compared to solutions that provide only text-based communication. However, the system may face limitations in handling dynamic gestures, complex sign sequences, and extreme environmental conditions. Future improvements can include the use of advanced models such as LSTM for sequence recognition, improved datasets for better accuracy, and mobile-based deployment for wider accessibility.

IV. RESULT

The result of the Sign Language Detection System is obtained through real-time gesture recognition using a webcam. The system captures live video input and processes each frame using OpenCV for hand detection and preprocessing. The detected hand region is then analyzed using a trained Convolutional Neural Network (CNN) model, which identifies the corresponding sign language gesture.

During execution, the system successfully detects the hand and maps key landmark points using a hand landmark detection model (e.g., MediaPipe). These landmarks represent the structure and position of the fingers and joints, which are then used as input features for gesture classification. As shown in **Figure 1**, the system accurately identifies the gesture and displays the predicted result on the screen (e.g., “Left: C”), indicating that the hand sign corresponds to the letter ‘C’. The bounding box around the hand ensures precise region detection, while the skeletal structure overlay enhances visualization of gesture tracking.

Figure 1: Real-time hand detection, landmark tracking, and gesture classification output (Letter ‘C’)

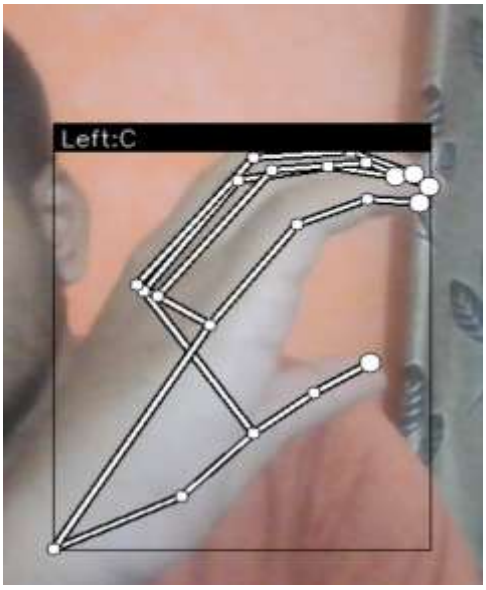
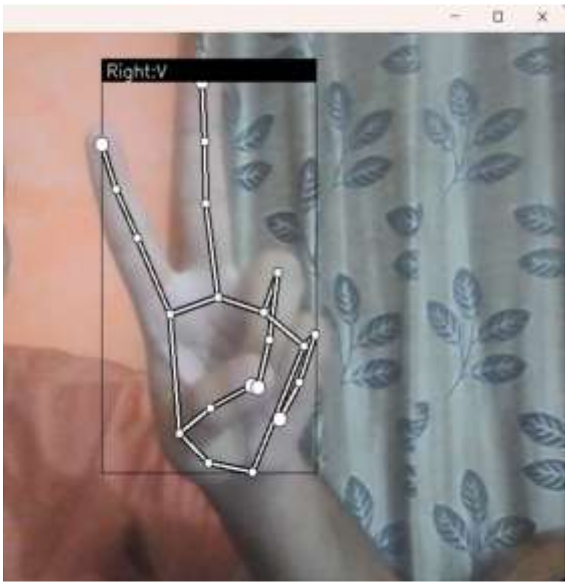


Figure 2: Real-time hand detection, landmark tracking, and gesture classification output (Letter 'V')



Figures 1 and 2 demonstrate the system's ability to accurately recognize and classify different hand gestures in real time.Once the gesture is recognized, it is converted into text output in real time. This text is further processed by a text-to-speech module to generate audio output, enabling both visual and auditory communication. Although the audio output cannot be displayed in the result images, it is generated successfully during system execution.

The results demonstrate that the system can accurately detect and classify hand gestures under real-time conditions with clear visual feedback. The integration of gesture recognition with both text and speech output provides a complete communication solution. However, the system performs best under controlled lighting and simple backgrounds, and accuracy may decrease in complex environments or with rapid hand movements.

V. CONCLUSION

The Sign Language Detection System successfully demonstrates the application of deep learning and computer vision in improving communication for hearing-impaired individuals. By using Convolutional Neural Networks (CNN) for gesture recognition and integrating OpenCV for real-time image processing, the system accurately identifies hand gestures and converts them into both text and speech output. This provides a complete and practical communication solution that bridges the gap between sign language users and others.

The system is designed to be cost-effective and accessible, as it uses a standard webcam instead of specialized hardware. Its real-time performance, combined with text-to-speech functionality, makes it suitable for everyday use in various environments such as education, public services, and personal communication. Although the system performs well under controlled conditions, it may face challenges with complex gestures, dynamic sign sequences, and varying environmental conditions.

Overall, the project highlights the potential of deep learning in assistive technologies and provides a foundation for further improvements. Future enhancements can include support for continuous sign language recognition, improved model accuracy with larger datasets, and deployment on mobile platforms to increase usability and reach.

ACKNOWLEDGEMENT

We express our sincere gratitude to our project guide and respected faculty members for their continuous support, guidance, and valuable suggestions throughout the development of this project. Their insights and encouragement played a crucial role in the successful completion of this work.

We also thank our institution for providing the necessary resources and environment to carry out this project effectively. Finally, we appreciate the efforts and cooperation of all team members, whose dedication and teamwork contributed significantly to the successful implementation of the system.

VI. REFERENCES:

• Anitha et al. (2022). A Static Hand Gesture Based Sign Language Recognition System using Convolutional Neural Networks. IEEE International Conference on Business Analytics for Technology and Security (ICBATS).

• Hira Hameed. (2022). Privacy-Preserving British Sign Language Recognition Using Deep Learning. IEEE Engineering in Medicine & Biology Society (EMBC).

• Piyusha V., Sanket D., et al. (2023). Detection and Interpretation of Indian Sign Language Using LSTM Networks. Journal of Intelligent Systems and Control (JISC).

• Natarajan et al. (2023). Hybrid Deep Neural Architecture for Sign Language Recognition. IEEE Xplore.

• Jeet Debnath. (2024). Real-Time Gesture Based Sign Language Recognition System. IEEE International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS).

• D. Shofia Priyadharshani. (2024). A Comprehensive Application for Sign Language Alphabet and World Recognition, Text-to-Action Conversion for Learners, Multi-Language Support and Integrated Voice Output Functionality. International Journal of Advanced Research in Science, Communication and Technology (IJARSCT).

• M. Soundarya. (2024). Sign Language Recognition Using Machine Learning. IEEE Xplore.

• C M Naveen Kumar & Vanitha A. (2024). Deep Learning Based Recognition of Sign Language. IEEE Conference on Data Science and Information System (ICDSIS).