International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Hand Gesture Sign Language Recognition System

Suchitra M*

MCA Student, Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India

ABSTRACT :

This project presents a real-time hand gesture recognition system designed to interpret American Sign Language (ASL) alphabets, numbers, and specific phrases such as "Thank you," "How are you?" and "What are you doing?" Utilizing Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, OpenCV, and Media Pipe, the system captures live hand gestures using a laptop camera and processes them for accurate interpretation. The system enhances communication for individuals with hearing impairments by translating gestures into text or speech, offering a practical, accessible tool for real-world applications. The project aims to improve accessibility and inclusivity, bridging the communication gap for individuals with hearing impairments and making it easier to communicate in various environments.

Keywords: Convolutional Neural Network (CNNs), ASL recognition, real-time hand gesture tracking, OpenCV.

1. Introduction :

The Hand Gesture Sign Language Recognition System (HGSLRS) utilizing Machine Learning has emerged as a ground-breaking tool for bridging communication gaps for individuals with hearing impairments. Traditional methods, such as relying on human interpreters, often limit spontaneity and autonomy, making real-time communication a challenge [1]. In contrast, HGSLRS offers an automated, accurate solution by interpreting complex hand gestures using advanced machine learning techniques, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. This system provides real-time translation of sign language gestures into text or speech, ensuring effective and seamless communication without the delays associated with manual interpretation [2]. In practical terms, HGSLRS offers increased independence for users in a variety of environments. For example, the system can be integrated into mobile devices, laptops, or wearable technology, ensuring accessibility wherever users go. This flexibility makes it especially valuable in settings such as educational institutions, workplaces, or public service environments, where immediate and reliable communication is essential. By eliminating the need for human intermediaries, the system allows individuals to engage in social interactions freely, promoting inclusivity and confidence in everyday conversations [3]. Another significant benefit of HGSLRS is its capacity to capture subtle hand movements and expressions, ensuring high accuracy in gesture recognition. The use of CNNs enables the system to recognize intricate hand shapes and positions, while LSTM networks capture the temporal dynamics of gestures, allowing the system to interpret continuous sign language fluently. This level of precision reduces the likelihood of miscommunication, enhancing the reliability of the system in real-time conversations [4]. The scalability of HGSLRS also makes it suitable for widespread adoption. Whether implemented in personal devices or public service infrastructures, the system can be easily adapted to different user needs without a significant increase in complexity or cost. Additionally, the data collected through the system can provide insights into communication patterns, further refining the system's accuracy over time and adapting to a wide range of hand gestures and linguistic nuances [5].

Overall, the HGSLRS represents a significant leap forward in enhancing communication accessibility for individuals with hearing impairments. By streamlining communication processes, reducing dependency on human interpreters, and promoting real-time interaction, the system empowers users to engage more fully in social and professional environments. As technology continues to evolve, HGSLRS holds the promise of fostering a more inclusive and accessible society for all.

2. Literature Survey

Sharma et al. [6] pioneered the use of deep learning techniques, specifically Convolutional Neural Networks (CNNs), for recognizing hand gestures in sign language. This approach marked a significant shift from traditional image processing techniques, as CNNs leverage hierarchical feature extraction, greatly improving the recognition accuracy of static hand gestures. This foundational work demonstrated the potential of deep learning models to handle complex visual data, laying the groundwork for subsequent advancements in the field.

Similarly, Bousbai et al. [7] introduced an ensemble of convolutional neural networks and capsule networks, improving American Sign Language (ASL) recognition by increasing model diversity and expanding datasets. This ensemble-based approach contributed to higher accuracy and robustness in sign language recognition, reflecting how multiple models working together can capture a wider range of gesture variations.

Building on these advancements, Aparna et al. [8] proposed a deep learning model using CNN and Long Short-Term Memory (LSTM) networks for Indian sign language recognition. This combination allows for effective capture of spatio-temporal information, crucial for recognizing the fluid movements of sign language gestures. Their method achieved high accuracy, showcasing the power of integrating temporal data for dynamic gesture recognition.

Azad et al. [9] further advanced gesture recognition with a multilevel temporal sampling method and Weighted Depth Motion Map (WDMM) for 3D hand gestures. Their approach reduced intra-class dissimilarities and increased inter-class similarity in depth maps, contributing to improved 3D hand gesture recognition accuracy.

In 2023, Zhong et al. [10] introduced a hybrid approach combining 3D-CNNs and Transformers for real-time hand gesture recognition. This method outperformed existing techniques in both speed and accuracy, highlighting the potential of combining different neural network architectures for more efficient gesture recognition systems.

Liao et al. [11] presented the B3D ResNet method, which significantly enhanced dynamic sign language recognition by improving the ability to recognize complex gestures in larger video sequences. Their approach demonstrated that deeper network architectures can better capture intricate hand movements.

Chong Tan et al. [12] focused on gesture segmentation using a fusion of color and depth information, alongside an SVM classifier. Their method achieved higher recognition rates and greater robustness than traditional methods, indicating the value of combining different types of data for more reliable gesture recognition.

Finally, Sahoo et al. [13] demonstrated the effectiveness of fine-tuning pre-trained CNN models using a score-level fusion technique. This approach improved real-time gesture recognition even in low-resource datasets, making it ideal for user-independent interfaces where training data may be limited.

3. Proposed Methodology

The proposed methodology for the hand gesture sign language recognition system involves using a camera to capture live images of hand gestures, followed by image enhancement and hand detection to ensure clarity. Key features are extracted using CNNs, and a machine learning model processes these features to recognize specific gestures, which are then mapped to corresponding text outputs. This structured approach enables accurate real-time translation of sign language into a user-friendly format.

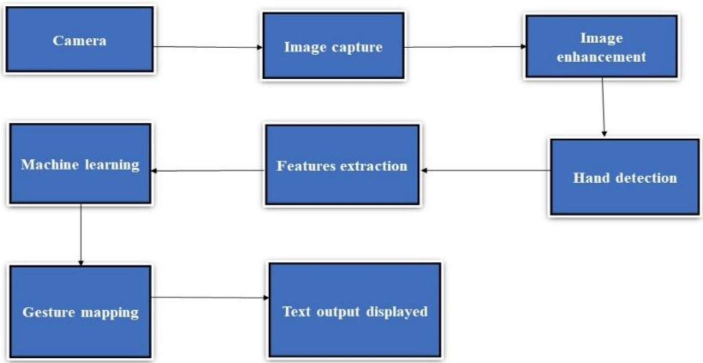


Fig 1. General architecture of the proposed work

- Camera and Image capture: A live camera (like the one on a laptop) captures images of the hand gestures in real-time, serving as the Input to the system.
- Hand *Detection:* The system detects the hand in the captured images using software tools like OpenCV and Media Pipe, ensuring only the hand is focused on for gesture recognition.
- Feature *Extraction:* Important details of the hand (like its shape and movement) are identified using machine learning models, especially Convolutional Neural Networks (CNNs), which help in understanding the gestures.
- Gesture *Capturing and Recognition:* The system then analyses these hand features to recognize specific gestures. For dynamic gestures involving movement, Long Short-Term Memory (LSTM) networks are used to interpret them correctly.
- Machine *Learning:* The system uses machine learning models to classify gestures. These models have been trained to recognize gestures.
- Output: Once the gesture is recognized, it is converted into text or speech for communication, making the system easy to use in real-time scenarios.

3.1 CNN Models

A Convolutional Neural Network (CNN) is a deep learning algorithm designed to process visual data by automatically extracting features from input images. In your Hand Gesture Sign Language Recognition System, CNN plays a crucial role in detecting and recognizing hand gestures from live camera feeds. The CNN architecture consists of convolutional layers that apply filters to detect features like edges and textures, followed by pooling layers that down sample the data, reducing its complexity while retaining essential information. After several convolutional and pooling layers, the feature maps are flattened and passed through fully connected layers for classification. The system learns through backpropagation, adjusting weights to improve accuracy over time. CNN enables your project to efficiently capture spatial hierarchies of hand gestures, making real-time recognition possible by identifying patterns and features unique to different sign language gestures, contributing to seamless and spontaneous communication for the hearing-impaired community.

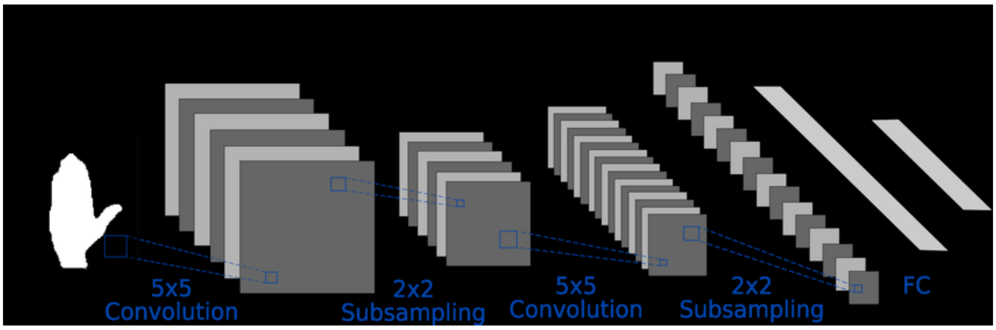


Fig 2. CNN Architecture Diagram

4. Experimental results and discussion :

In this section, the Experimental results were discussed and the screenshots of the results were provided below.

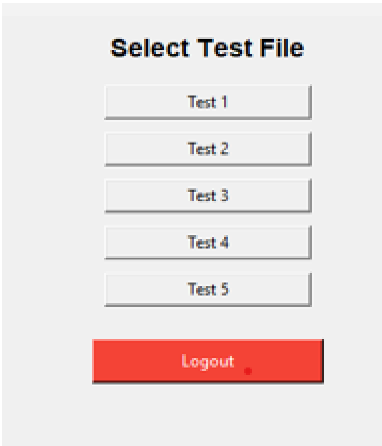


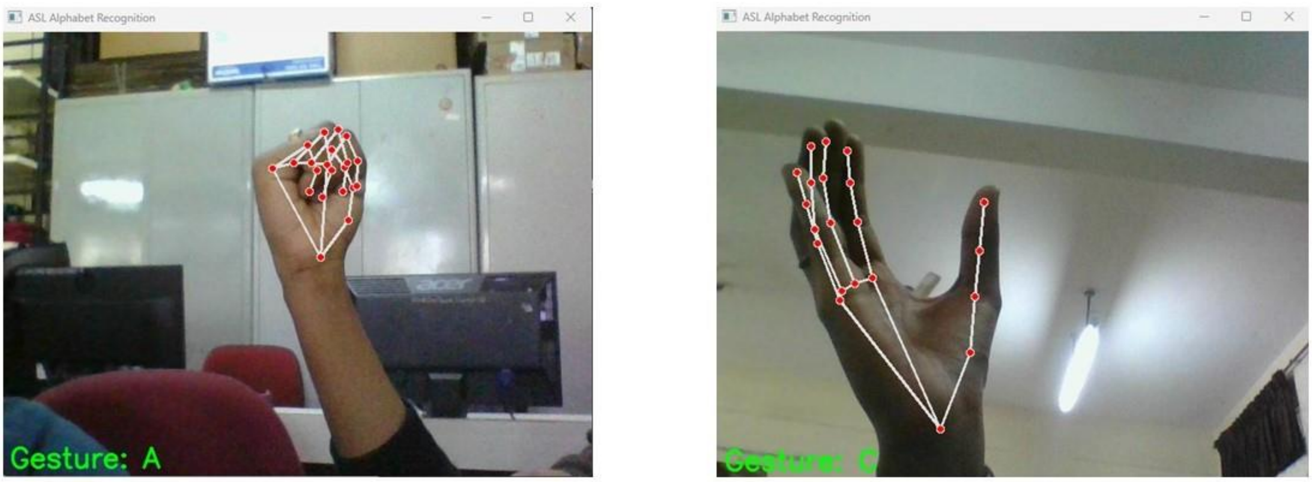
Fig 3. Select test files

Test 1: Recognizes alphabet gestures (A-Z).
Test 2: Recognizes numbers.
Test 3: Identifies thumbs up and thumbs down gestures.
Test 4: Detects simple sentences such as "What are you doing?" and "How are you?", super.
Test 5: Recognizes the "Thank you" gesture.

Test 1

The Alphabet Recognition Output Page is where users receive the results of their hand gesture recognition for alphabetical inputs. After performing a gesture representing any letter from A to Z in front of the camera, the system processes the input and displays the corresponding letter on this page.

Fig 4. Test 1 result



Test 2

The Number Recognition Output Page is where users receive the results of their hand gesture recognition for numerical inputs. After performing a number gesture representing any digit from 0 to 10 in front of the camera, the system analyzes the input and displays the corresponding number on this page.

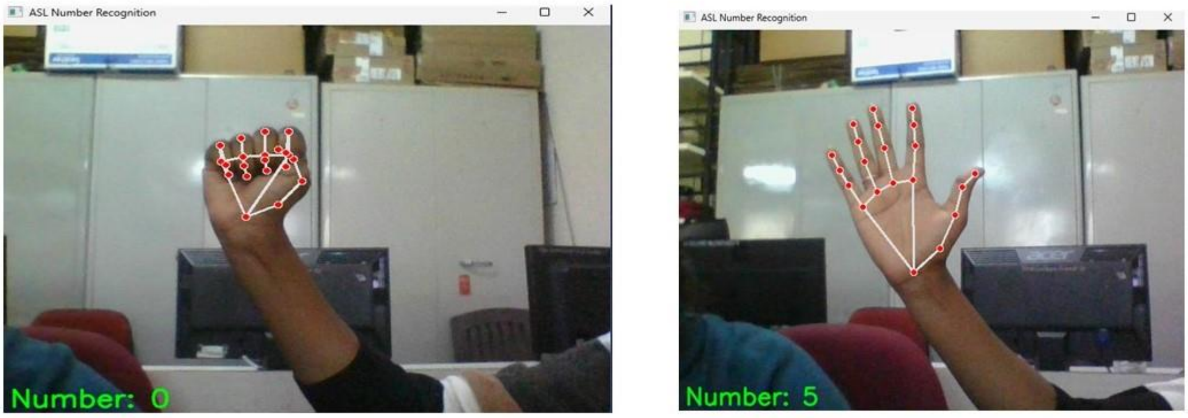


Fig 5. Test 2 result

Test 3

The Thumbs Up and Thumbs Down Recognition Output Page is where users receive the results of their hand gesture recognition for specific gestures. After performing either the thumbs up or thumbs down gesture in front of the camera, the system analyzes the input and displays the corresponding gesture result on this page.

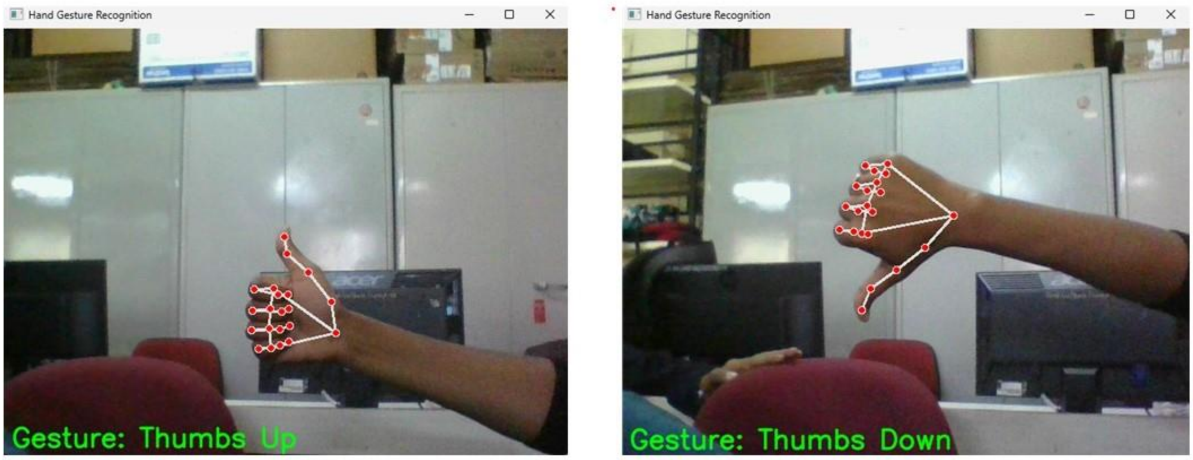
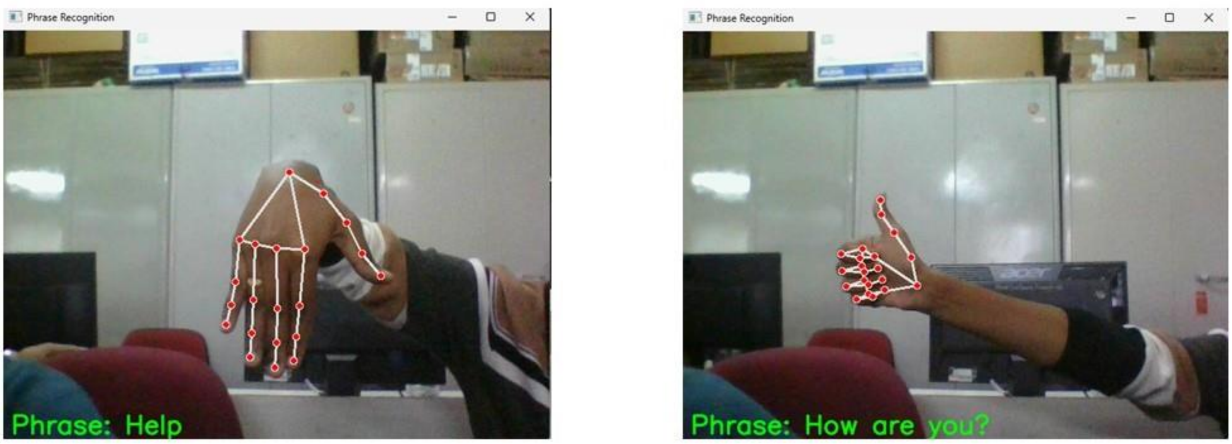


Fig 6. Test 3 result

Test 4

The Simple Sentences Recognition Output Page is where users receive the results of their hand gesture recognition for predefined simple sentences. After performing a gesture that represents a specific sentence, such as "What are you doing?" or "How are you?" or help in front of the camera, the system analyzes the input and displays the corresponding sentence on this page.

Fig 7. Test 4 result



Test 5

Output Page is where users receive the results of their hand gesture recognition for the specific "Thank You" gesture. After performing the "Thank You" gesture in front of the camera, the system processes the input and displays the corresponding phrase, "Thank You," on this page.

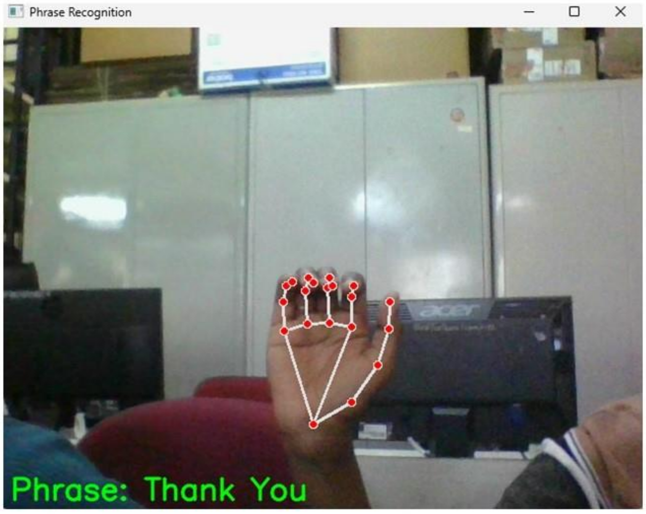


Fig 8. Test 5 result

Conclusion :

In conclusion, our hand gesture sign language recognition system marks a significant advancement in bridging communication barriers for individuals with hearing impairments by accurately translating sign language gestures into text through the use of machine learning and computer vision. Moving forward, we aim to enhance the system’s accuracy, real-time performance, and ability to adapt to multiple sign languages. Future work will include expanding the dataset with diverse gestures from different dialects and lighting conditions, improving real-time translation for live conversations, and integrating NLP to facilitate contextual understanding of phrases and sentences. Additionally, the system will incorporate wearable technology like smart glasses and gloves to further enhance accessibility. Ethical considerations, particularly user privacy, remain central to our approach, and collaboration with the deaf community, experts, and relevant organizations will ensure that the system is inclusive, culturally sensitive, and impactful globally.

Acknowledgement

I am grateful to my project guide Prof. Shivanand Gornale, Professor, Department of Computer Science, Rani Channamma University, Belagavi for his valuable guidance for completion of this work.

Conflict of Interest

There is no conflict of interest declared by authors.

REFERENCES :

[1] Siby, T., Pal, S., Arlina, J., & Nagaraju, S. (2022). Gesture based Real-Time Sign Language Recognition System. *2022 International Conference on Connected Systems & Intelligence (CSI)*, 1-6. https://doi.org/10.1109/CSI54720.2022.9924024.

[2] Saeed, Z., Zainol, Z., Zaidan, B., & Alamoodi, A. (2022). A Systematic Review on Systems-Based Sensory Gloves for Sign Language Pattern Recognition: An Update From 2017 to 2022. *IEEE Access*, 10, 123358-123377. https://doi.org/10.1109/ACCESS.2022.3219430.

[3] Zhou, Z., Tam, V., & Lam, E. (2022). A Portable Sign Language Collection and Translation Platform with Smart Watches Using a BLSTM-Based Multi-Feature Framework. *Micromachines*, 13. https://doi.org/10.3390/mi13020333.

[4] Young, J., Langlotz, T., Cook, M., Mills, S., & Regenbrecht, H. (2019). Immersive Telepresence and Remote Collaboration using Mobile and Wearable Devices. *IEEE Transactions on Visualization and Computer Graphics*, 25, 1908-1918. https://doi.org/10.1109/TVCG.2019.2898737.

[5] Ellis, R., Rönnberg, J., & Plejert, C. (2023). Combined effects of age and hearing impairment on utterances and requests for clarification in spontaneous conversation and a referential communication task. *International journal of language & communication disorders*. https://doi.org/10.1111/1460-6984.12940.

[6] Sharma, S., & Singh, S. (2021). Vision-based hand gesture recognition using deep learning for the interpretation of sign language. *Expert Syst. Appl.*, 182, 115657. https://doi.org/10.1016/J.ESWA.2021.115657.

[7] Bousbai, K., Morales-Sánchez, J., Merah, M., & Sancho-Gómez, J. (2022). Improving hand gestures recognition capabilities by ensembling convolutional networks. *Expert Systems*, 39. https://doi.org/10.1111/exsy.12937.

[8] Aparna, C., & Geetha, M. (2020). CNN and Stacked LSTM Model for Indian Sign Language Recognition. , 126-134. https://doi.org/10.1007/978-981-15-4301-2_10.

[9] Azad, R., Asadi-Aghbolaghi, M., Kasaei, S., & Escalera, S. (2019). Dynamic 3D Hand Gesture Recognition by Learning Weighted Depth Motion Maps. *IEEE Transactions on Circuits and Systems for Video Technology*, 29, 1729-1740. https://doi.org/10.1109/TCSVT.2018.2855416.

[10] Zhong, E., del-Blanco, C., Berjón, D., Jaureguizar, F., & García, N. (2023). Real-Time Monocular Skeleton-Based Hand Gesture Recognition Using 3D-Jointsformer. *Sensors (Basel, Switzerland)*, 23. https://doi.org/10.3390/s23167066.

[11] Liao, Y., Xiong, P., Min, W., Min, W., & Lu, J. (2019). Dynamic Sign Language Recognition Based on Video Sequence With BLSTM-3D Residual Networks. *IEEE Access*, 7, 38044-38054. https://doi.org/10.1109/ACCESS.2019.2904749.

[12] Tan, C., Sun, Y., Li, G., Jiang, G., Chen, D., & Liu, H. (2020). Research on gesture recognition of smart data fusion features in the IoT. *Neural Computing and Applications*, 1-13. https://doi.org/10.1007/s00521-019-04023-0.

[13] Sahoo, J., Prakash, A., Plawiak, P., & Samantray, S. (2022). Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. *Sensors (Basel, Switzerland)*, 22. https://doi.org/10.3390/s22030706.