ISSN 2319-5991 www.ijerst.com

Peer Reviewed,Referred & Indexed Journal

Research Paper

A GESTURE RECOGNITION FOR SIGN LANGUAGE TRANSLATION USING DEEP LEARNING AND COMPUTER VISION

Mrs. P. LAXMI PRASANNA

Assistant Professor

Netinti Lakshmi Narasimha Naidu, Kummari Uday Kumar, Pidaparthi John Peter, Ponnekanti Thanvi

ABSTRACT

Sign language is a vital communication method used by individuals with hearing and speech impairments. However, communication barriers often exist between sign language users and non-signers due to the lack of widespread understanding of sign languages. To bridge this gap, this paper proposes a Gesture Recognition System for Sign Language Translation using Deep Learning and Computer Vision. The system is designed to capture hand gestures through a camera and translate them into corresponding text or speech in real time.

The proposed system uses advanced deep learning techniques, particularly **Long Short-Term Memory (LSTM) networks**, to accurately recognize and classify hand gestures. Since sign language often involves sequential and temporal patterns, LSTM models are well-suited for capturing motion dynamics across video frames. Computer vision techniques are applied to detect and preprocess hand regions from video sequences before feeding them into the model. The system is trained on a large dataset of sign language gestures to ensure high accuracy and robustness across different users and environments.

Unlike traditional systems that rely on gloves or sensor-based devices, this approach is contactless and purely vision-based, making it more convenient and cost-effective. The model is capable of recognizing both static and dynamic gestures by learning temporal dependencies in the gesture sequences. The output is displayed as text and can also be converted into speech, enabling seamless communication between hearing-impaired and normal individuals.

This system can be widely used in education, healthcare, public services, and real-time communication platforms. By integrating artificial intelligence with sign language interpretation, the proposed system aims to improve accessibility and inclusivity in society.

INTRODUCTION

Communication is one of the most important aspects of human interaction. However, individuals with hearing and speech disabilities often face challenges in communicating with others who do not understand sign language. Sign language serves as a primary communication medium for such individuals, but the lack of knowledge among the general population creates a communication gap.

With advancements in Artificial Intelligence (AI) and Computer Vision, it has become possible to develop systems that can interpret human gestures and translate them into meaningful outputs. Gesture recognition technology focuses not only on identifying hand shapes and positions but also on capturing the temporal sequence of movements, which is essential for accurately interpreting sign language.

The proposed system aims to develop an intelligent sign language translation system that uses deep learning algorithms, particularly **Long Short-Term Memory (LSTM) networks**, to recognize gestures in real time. LSTM models are highly effective in processing sequential data, enabling the system to learn patterns across consecutive video frames. The system captures video input from a camera, processes the frames using image processing techniques, and classifies gesture sequences using a trained LSTM-based neural network model.

This system eliminates the need for wearable devices or special sensors, making it more user-friendly and accessible. It can be used in real-time communication scenarios such as classrooms, hospitals, public services, and daily conversations. The ability of LSTM to understand motion over time makes it especially suitable for recognizing dynamic gestures in sign language.

EXISTING SYSTEM

Existing sign language recognition systems primarily rely on wearable devices such as data gloves, motion sensors, or marker-based tracking systems. These systems detect finger movements and hand positions using physical sensors attached to the user's hand, making them intrusive, expensive, and less convenient for everyday use.

Some vision-based systems also exist; however, they are often limited in accuracy and typically require controlled environments. Traditional approaches mainly depend on handcrafted features such as edge detection, contour analysis, and shape recognition. These methods fail to effectively capture the temporal dynamics of gestures and are not robust enough for real-world applications.

Many existing systems focus only on static gesture recognition and do not adequately handle dynamic gestures that involve motion over time. This limitation arises because conventional models lack the ability to learn sequential dependencies between frames. In contrast, sequence-based models such as **Long Short-Term Memory (LSTM) networks** are better suited for modeling temporal patterns in gesture sequences.

Additionally, existing systems often struggle with variations in lighting conditions, complex backgrounds, and differences among users. These challenges highlight the need for a more advanced, vision-based approach that can effectively learn both spatial and temporal features to achieve accurate and reliable sign language recognition in real-world scenarios.

Disadvantages

 Requires wearable devices or sensors

 High cost and maintenance complexity

 Limited accuracy in real-world conditions

 Poor handling of dynamic gestures

 Not fully real-time or scalable

 Sensitive to lighting and background variations

PROPOSED SYSTEM

The proposed system introduces a vision-based deep learning approach for sign language recognition. It uses computer vision techniques to detect and extract hand regions from live video input and applies a trained **Long Short-Term Memory (LSTM) model** to classify gestures accurately.

Unlike traditional models, the LSTM-based approach is designed to capture temporal dependencies in gesture sequences, making it highly effective for interpreting motion-based sign language. The system processes continuous video frames and learns patterns across time, enabling more accurate recognition of dynamic gestures.

The system operates in real time, making it suitable for live communication scenarios. It supports both static gestures (single hand positions) and dynamic gestures (motion-based signs), leveraging sequential data processing to improve performance depending on dataset training.

Advantages

 Contactless and non-intrusive system

 Cost-effective solution

 Scalable for multiple languages

 Improves accessibility for hearing-impaired users

 Works in real-time communication scenarios

SYSTEM ARCHITECTURE

TECHNOLOGIES USED

- Python
- TensorFlow / Keras
- OpenCV
- NumPy
- LSTM
- Speech synthesis libraries (optional)

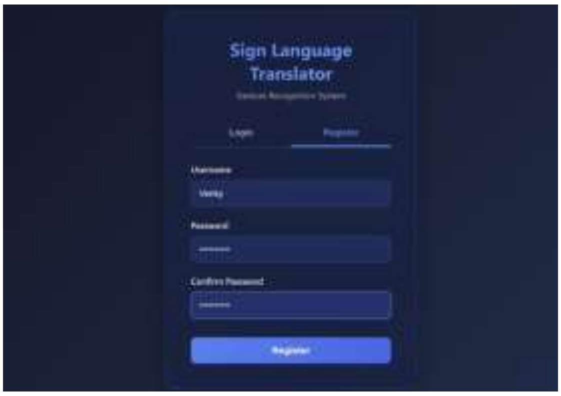
APPLICATIONS

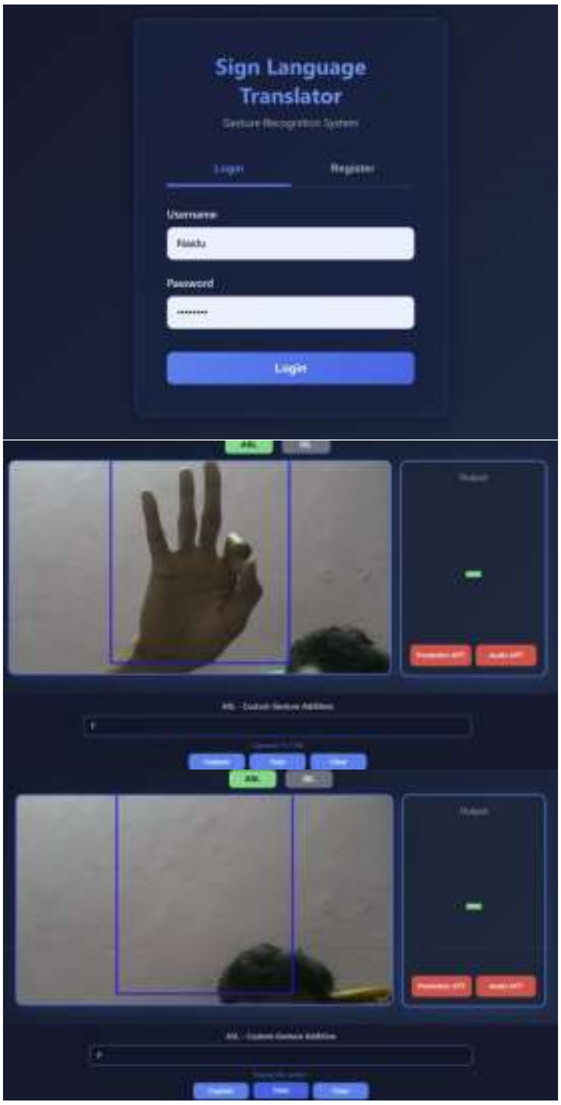
- Communication for hearing-impaired individuals
- Educational tools for learning sign language
- Smart assistive devices
- Public service communication systems
- Healthcare assistance systems
- Human-computer interaction systems

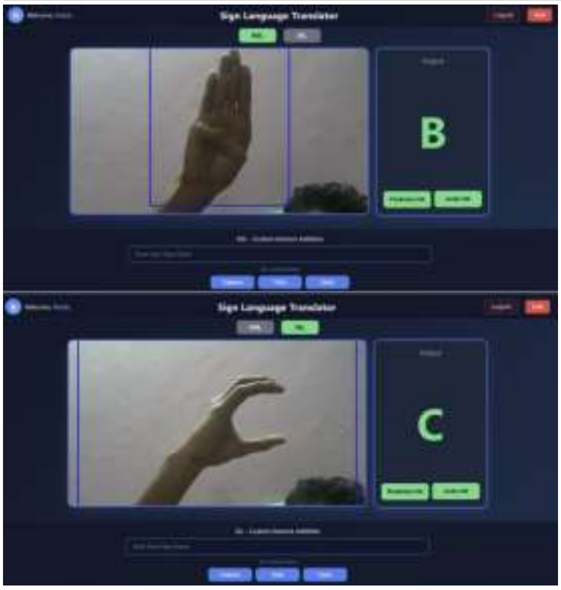
CHALLENGES & RISKS

- Variation in hand shapes and sizes
- Complex background interference
- Real-time processing delay
- Limited dataset availability
- Difficulty in recognizing similar gestures

RESULTS







CONCLUSION

The Gesture Recognition System for Sign Language Translation using Deep Learning and Computer Vision provides an effective and modern solution for bridging communication gaps between hearing-impaired individuals and the general population. By utilizing advanced **LSTM-based models** and real-time video processing techniques, the system is capable of accurately recognizing hand gestures and converting them into meaningful text or speech outputs.

This system significantly improves accessibility and inclusivity by removing the dependency on physical devices such as gloves or sensors. Its vision-based approach makes it more practical, scalable, and cost-effective for real-world applications. The integration of **Long Short-Term Memory (LSTM) networks** enhances recognition accuracy by capturing temporal patterns in gesture sequences and allows the system to adapt to different users and environments.

The proposed model demonstrates strong potential in fields such as education, healthcare, assistive technology, and human-computer interaction. However, further improvements can be made by expanding datasets, enhancing dynamic gesture recognition, and optimizing real-time performance for faster and more efficient predictions.

In conclusion, this system represents an important step toward intelligent communication systems that promote equal opportunities and social inclusion through artificial intelligence, especially by leveraging sequence-based learning for more accurate and natural sign language interpretation.

FUTURE ENHANCEMENTS

- Support for multiple sign languages (ASL, ISL, etc.)
- Improved recognition of dynamic gestures
- Integration with mobile applications
- Real-time translation into multiple spoken languages
- Use of transformer-based models for higher accuracy
- Cloud-based deployment for scalability

REFERENCES

1. · Karen Simonyan and Andrew Zisserman,
"Very Deep Convolutional Networks for Large-Scale Image Recognition,"
International Conference on Learning Representations (ICLR), 2015.

2. · Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun,
"Deep Residual Learning for Image Recognition,"
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

3. · Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton,
"ImageNet Classification with Deep Convolutional Neural Networks,"
Advances in Neural Information Processing Systems (NeurIPS), 2012.

4. · World Health Organization,
"Deafness and Hearing Loss," 2021.
(Used for background on hearing-impaired communication needs.)

5. · OpenCV Documentation,
"Open Source Computer Vision Library,"
Available: https://opencv.org/

6. · TensorFlow,
"TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015.
Available: https://www.tensorflow.org/

7. · Keras Documentation,
"Keras: Deep Learning for Humans,"
Available: https://keras.io/

8. · National Institute on Deafness and Other Communication Disorders,
"American Sign Language," 2020.

9. · Starner, T., Weaver, J., & Pentland, A.,
"Real-Time American Sign Language Recognition Using Desk and Wearable Computer-Based Video,"
IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.

10. · Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B.,
"Sign Language Recognition Using Convolutional Neural Networks,"
European Conference on Computer Vision Workshops, 2014.