PERFORMANCE-OPTIMIZEDREAL-TIMESIGN LANGUAGE RECOGNITION

[1]AFSHA SULTANA, [2]P. PRASHANTHI, [3]M. NITHYA, [4]M.RUCHITHA

[1]Assistant Professor, CSE (AI&ML), Bhoj Reddy Engineering College for Women.

[234]B.TECH, SCHOLAR, CSE (AI&ML), Bhoj Reddy Engineering College for Women.

ABSTRACT

The Sign Language Recognition (SLR) system is developed to bridge the communication gap between individuals with hearing and speech impairments and the general public by converting hand gestures into readable text using computer vision and deep learning techniques. Sign language is one of the most important communication methods for people who cannot hear or speak; however, many individuals are not familiar with it, which creates communication barriers. This project aims to provide a real-time, efficient, and user-friendly solution that automatically recognizes hand gestures and translates them into meaningful text output. The proposed system uses a webcam to capture live video input of hand gestures. The captured frames are processed using computer vision techniques to detect and track hand movements. A hand-tracking algorithm identifies key hand landmarks such as finger joints and palm positions. These landmarks are extracted as numerical keypoints that represent the structure and movement of the hand. Since gestures involve motion over time, sequential keypoint data from multiple frames is collected to accurately represent each gesture. This sequential data is then used to train a deep learning model. A Long Short-Term Memory (LSTM) neural network is employed for gesture recognition. LSTM is a type of Recurrent Neural Network (RNN) that is well-suited for handling sequential data and learning temporal patterns. The model is trained using different gesture sequences so that it can learn and distinguish between various hand movements. During real-time execution, the system continuously captures hand gestures, extracts keypoints, and feeds them into the trained LSTM model. The model predicts the performed gesture and displays the corresponding text on the screen, enabling users to communicate effectively without requiring any specialized hardware.

The system offers several advantages over traditional sign language recognition methods. It eliminates the need for expensive sensor gloves or wearable devices and relies only on a webcam and a computer. The system is cost-effective, easy to use, and capable of recognizing gestures in real time. It also provides good accuracy using deep learning techniques and can be extended to include additional gestures. Furthermore, the system is designed to function under varying lighting conditions and backgrounds, making it practical for real-world applications. Overall, the Sign Language Recognition system demonstrates the effective use of artificial intelligence, computer vision, and deep learning to develop an assistive communication tool. The project highlights the potential of technology in improving accessibility and inclusivity for people with disabilities. With further enhancements such as speech output, mobile integration, and multi-language support, the system can be expanded into a more advanced communication platform for real-time interaction.

1.INTRODUCTION

1.1 INTRODUCTION OF THE PROJECT

- Sign Language Recognition systems use computer vision and machine learning to interpret hand gestures and convert them into digital outputs.
- The main purpose of this system is to help people with hearing and speech impairments communicate easily.
- In this project, a webcam captures hand movements and detects gestures in real time.
- A hand-tracking technique extracts keypoints (hand landmarks) such as finger joints and palm positions.
- The keypoint data from multiple frames is used to train a Long Short-Term Memory (LSTM) model, which learns gesture patterns over time.
- When a gesture is performed, the trained model recognizes it and displays the corresponding text on the screen.

1.2 EXISTING SYSTEM

Existing sign language recognition systems use image processing, wearable sensor gloves, or basic machine learning models to detect hand gestures. Glove-based systems are accurate but expensive and inconvenient, while vision-based methods are sensitive to lighting and background conditions. Many systems support only a limited set of gestures and lack real-time performance, making them less suitable for practical use. These limitations create the need for a simple, real-time, camera-based recognition system using modern machine learning techniques.

1.3 PROBLEMS IN EXISTING SYSTEM

- Dependence on Special Hardware: Many existing systems use sensor gloves or wearable devices, which are expensive and uncomfortable for users.

- Sensitivity to Lighting Conditions: Vision-based systems that use cameras often fail in poor lighting or complex backgrounds, reducing accuracy.
- Limited Gesture Vocabulary: Most existing systems can recognize only a small set of gestures, which limits practical communication.
- Low Real-Time Performance: Some systems process gestures slowly and cannot recognize them instantly, making communication difficult.
- High Computational Cost: Advanced image processing techniques may require powerful hardware and high processing time.
- User Dependency: Some systems require specific hand positions, training, or calibration, making them difficult for new users.
- Lack of Accuracy: Incorrect detection of hand landmarks or motion can lead to wrong gesture predictions.
- Poor Scalability: Adding new gestures to existing systems is often difficult and requires retraining the entire model.

1.4 PROPOSED SYSTEM

The proposed system aims to develop a real-time Sign Language Recognition system using computer vision and deep learning techniques. The system uses a webcam to capture hand gestures and processes them using machine learning algorithms to recognize the corresponding sign.

In this system, a hand-tracking method is used to detect hand landmarks such as finger joints and palm positions. These landmarks are extracted from each video frame and converted into keypoint coordinates. The sequence of keypoints from multiple frames represents the movement of the hand while performing a gesture.

A Long Short-Term Memory (LSTM) deep learning model is used to analyze this sequential data. The LSTM model learns the patterns of different gestures during the training phase. When a user performs a gesture in front of the camera, the system processes the sequence of landmarks, and the trained model predicts the gesture.

After recognition, the system converts the detected gesture into text and displays it on the screen. This helps people with hearing and speech impairments communicate more easily with others.

1.5 ADVANTAGES OF PROPOSED SYSTEM

- Real-time recognition: The system detects and recognizes sign language gestures instantly using a webcam.
- No wearable devices required: It eliminates the need for expensive sensor gloves or special hardware.
- High accuracy: Deep learning models improve the accuracy of gesture recognition.

- Works in different environments: The system can operate under varying lighting and background conditions.
- Scalable system: New gestures can be easily added by updating the dataset and retraining the model.
- Improves communication: Converts gestures into text, helping hearing and speech-impaired people communicate easily.
- Cost-effective and user-friendly: Requires only a webcam and computer, making it simple and affordable to use.

2. LITERATURE SURVEY

Sign Language Recognition (SLR) has become an important research area in artificial intelligence, aiming to bridge the communication gap between hearing-impaired individuals and the general population. Over the years, researchers have explored various techniques ranging from traditional image processing to advanced deep learning models for real-time gesture recognition.

Early studies in SLR relied on image processing and handcrafted feature extraction methods. However, these approaches faced limitations in handling variations in lighting, background, and hand orientation. With the advancement of artificial intelligence, machine learning techniques were introduced to improve recognition accuracy. According to Xianwei Jiang, Suresh Chandra Satapathy, and Yu-Dong Zhang (2020), artificial intelligence plays a key role in improving sign language recognition by enabling systems to learn patterns from data rather than relying on predefined rules (Springer).

The introduction of deep learning significantly transformed the field of SLR. Junan Jiang (2024) highlighted that models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks have shown excellent performance in gesture recognition tasks. CNNs are effective for extracting spatial features from images, while LSTMs are suitable for learning temporal dependencies in sequential data (Darcy & Roy Press).

Recent literature emphasizes hybrid models that combine spatial and temporal learning. B. N. Madhukar et al. (2025) discussed that CNN-LSTM hybrid architectures achieve high accuracy (above 90%) in recognizing both static and dynamic gestures. These models first extract features using CNN and then process sequential information using LSTM, making them suitable for real-time applications (IJRASET).

A comprehensive survey by Yanqiong Zhang and Xianwei Jiang (2024) analyzed advancements in deep learning-based SLR systems. Their work highlighted the importance of large datasets, efficient preprocessing techniques, and model optimization for improving recognition accuracy. The study also emphasized challenges such as signer variability, complex gestures, and real-time performance constraints (ResearchGate).

Tangfei Tao, Yizhe Zhao, Tianyu Liu, and Jieli Zhu (2024) provided a detailed review of traditional and modern SLR methods.

They concluded that deep learning-based approaches outperform traditional methods due to their ability to automatically extract features and handle complex data variations. However, they also pointed out issues related to computational cost and data requirements (ResearchGate).

Recent research has focused on performance optimization in real-time systems. Abdullah Baihan, Ahmed I. Alutaibi, and Sunil Kumar Sharma (2024) proposed a hybrid CNN-Self Attention-LSTM model for sign language recognition. Their model achieved high accuracy of 98.7% by combining spatial, temporal, and attention-based features, demonstrating the effectiveness of hybrid deep learning architectures (Nature).

Samuel Ady Sanjaya and Hadinata Faustine Ilone (2023) conducted a systematic literature review on deep learning techniques for sign language recognition. Their study highlighted the importance of preprocessing, feature extraction, and model selection in achieving high performance. They also emphasized the need for standardized datasets and evaluation metrics (IJCS).

Recent advancements also include transformer-based models and multimodal approaches. Studies show that combining hand keypoints, body pose, and facial expressions improves recognition accuracy. Additionally, the use of MediaPipe for hand landmark detection has become popular due to its efficiency and real-time capabilities.

Another important trend is the development of lightweight models for real-time applications. Research indicates that traditional deep learning models require high computational resources, making them unsuitable for edge devices. New approaches focus on optimizing models to reduce latency and improve efficiency without compromising accuracy.

Overall, the literature shows that SLR has evolved from simple image processing techniques to advanced deep learning systems. Despite significant progress, challenges such as real-time performance, scalability, and robustness remain active areas of research.

3. METHODOLOGY

The performance-optimized real-time sign language recognition system is designed using a deep learning-based approach that integrates computer vision and sequential learning techniques. The system begins with capturing live video input using a webcam. The video stream is processed frame by frame to detect and track hand movements.

Each frame undergoes preprocessing to improve image quality and ensure consistency. Techniques such as resizing, normalization, and background removal are applied. A hand-tracking algorithm is used to identify key landmarks of the hand, including finger joints and palm positions. These landmarks are converted into numerical keypoints representing the structure of the hand.

Since sign language involves motion, the system collects sequences of keypoints over multiple frames. This sequential data captures the temporal dynamics of gestures. The extracted sequences are then used as input to the deep learning model.

A hybrid CNN-LSTM model is employed for gesture recognition. The CNN component extracts spatial features from the input data, while the LSTM processes temporal dependencies. This combination allows the system to understand both the shape and movement of gestures.

The model is trained using labeled datasets containing various sign language gestures. During training, optimization techniques such as backpropagation and gradient descent are used to minimize loss and improve accuracy. Data augmentation techniques are also applied to increase dataset diversity.

In real-time operation, the system continuously captures video, extracts keypoints, and feeds them into the trained model. The model predicts the gesture and converts it into text output. Performance optimization techniques such as model pruning and efficient data processing are used to reduce latency and improve speed.

The system is implemented using Python and deep learning frameworks such as TensorFlow or PyTorch. It is designed to run efficiently on standard hardware, making it suitable for real-time applications.

4.EXISTING METHODS

Existing sign language recognition systems can be categorized into traditional methods, machine learning approaches, and deep learning-based systems. Traditional methods rely on image processing techniques and handcrafted features. These methods include edge detection, contour analysis, and feature extraction techniques such as Histogram of Oriented Gradients. While these methods are computationally efficient, they lack robustness and fail in complex environments.

Machine learning-based methods use classifiers such as Support Vector Machines and k-Nearest Neighbors. These systems improve accuracy compared to traditional methods but require manual feature extraction. They also struggle with dynamic gestures and temporal dependencies.

Deep learning-based methods have become the dominant approach in recent years. Convolutional Neural Networks are widely used for feature extraction, while Recurrent Neural Networks and LSTM networks handle sequential data. These systems achieve high accuracy but require large datasets and computational resources.

Glove-based systems represent another category of existing methods. These systems use sensors to capture hand movements accurately. Although they provide high precision, they are expensive and inconvenient for users.

Vision-based systems using cameras are more practical but face challenges such as sensitivity to lighting conditions, background noise, and occlusion. Many existing systems also support only a limited set of gestures, reducing their usability.

Another limitation of existing methods is the lack of real-time performance. Some systems process data slowly, making them unsuitable for real-world communication. Additionally, many systems are not scalable and require retraining when new gestures are added.

Overall, while existing methods have made significant progress, they still face challenges in accuracy, efficiency, and usability.

5. PROPOSED SYSTEM

The proposed performance-optimized real-time sign language recognition system aims to overcome the limitations of existing methods by integrating advanced deep learning techniques and efficient processing mechanisms. The system is designed to provide accurate, fast, and user-friendly gesture recognition.

The system uses a webcam to capture live video input and processes it using computer vision techniques. A hand-tracking algorithm extracts keypoints representing hand landmarks. These keypoints are used to create sequential data representing gesture movements.

A hybrid deep learning model combining CNN and LSTM is used for recognition. The CNN extracts spatial features, while the LSTM captures temporal dependencies. This combination improves accuracy and enables the system to recognize dynamic gestures effectively.

To optimize performance, the system incorporates techniques such as model compression, efficient data processing, and real-time inference. These optimizations reduce latency and improve speed, making the system suitable for real-time applications.

The system also includes a text output module that converts recognized gestures into readable text. This enables effective communication between users. Additionally, the system is designed to work under different lighting conditions and backgrounds, improving robustness.

Another key feature is scalability. The system allows new gestures to be added easily by updating the dataset and retraining the model. It also supports integration with other technologies such as speech synthesis and mobile applications.

The proposed system eliminates the need for expensive hardware and relies only on a webcam and computer. It is cost-effective, easy to use, and accessible to a wide range of users.

Overall, the proposed system provides a comprehensive solution for real-time sign language recognition by combining accuracy, efficiency, and usability.

6. SCREENSHOTS

OUTPUT SCREENS:

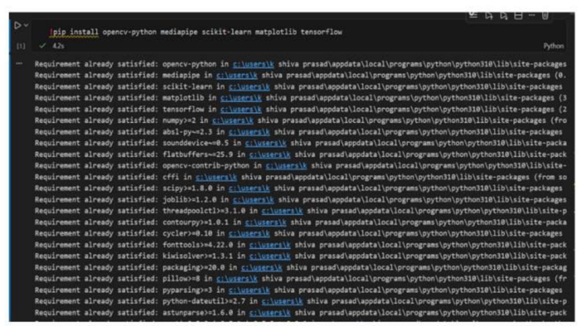

Fig 6.1:

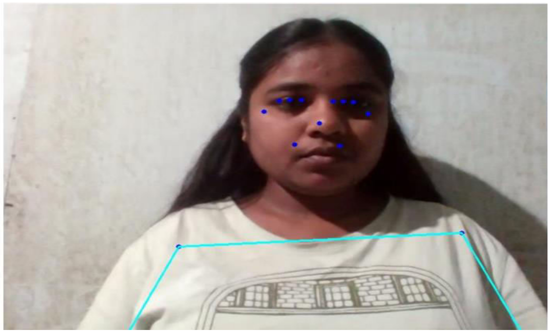

Fig 6.2 :

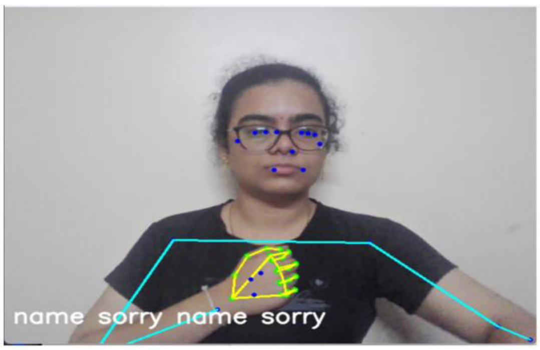

Fig 6.3:

Fig 6.4:

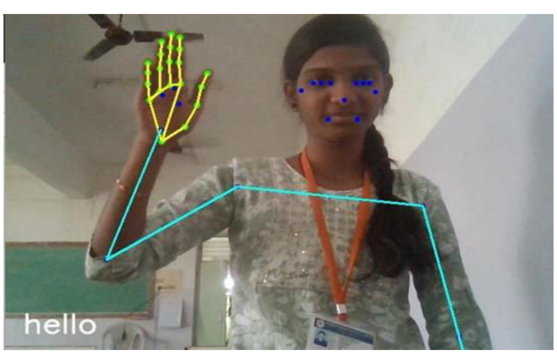

Fig 6.5:

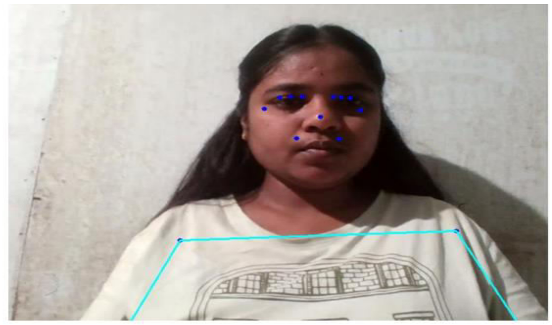

Fig 6.6:

Fig 6.7

Fig 6.8:

7. CONCLUSION

The Performance-Optimized Real-Time Sign Language Recognition system demonstrates an effective application of artificial intelligence, computer vision, and deep learning for assistive communication. The system successfully bridges the communication gap between hearing and speech-impaired individuals and the general public by converting hand gestures into meaningful text in real time. By integrating hand landmark detection with a hybrid CNN-LSTM model, the system achieves accurate recognition of both static and dynamic gestures while maintaining low latency for real-time performance.

The proposed system eliminates the need for expensive wearable devices and relies only on a standard webcam, making it cost-effective and accessible. Performance optimization techniques further enhance speed and efficiency, making the system suitable for real-world deployment. Additionally, the system's ability to work under varying lighting conditions and backgrounds improves its robustness and usability.

Overall, this project highlights how deep learning-based solutions can significantly improve accessibility and inclusivity. It provides a scalable and efficient framework that can be extended for broader communication applications, contributing to improved interaction between differently-abled individuals and society.

8. FUTURE SCOPE

The future development of the Performance-Optimized Real-Time Sign Language Recognition system offers several promising directions. One major improvement is the integration of speech synthesis, where recognized text can be converted into audible speech, enabling two-way communication between hearing and speech-impaired individuals and others.

Another important enhancement is the use of transformer-based deep learning models, which can further improve recognition accuracy and better handle long gesture sequences. The system can also be extended to support a larger vocabulary of sign languages, including regional and international sign variations.

Mobile and edge device deployment is another key area for future work. Optimizing the model for smartphones and embedded systems will make the solution more accessible and widely usable. Additionally, cloud-based deployment can enable large-scale usage and continuous model improvement through user data.

Incorporating multimodal inputs such as facial expressions and body posture can significantly enhance recognition accuracy and contextual understanding. Furthermore, the use of federated learning can help improve model training while ensuring user data privacy.

Overall, the system has strong potential for expansion into a comprehensive communication platform, contributing to inclusive and intelligent human-computer interaction.

9. REFERENCE

1. Jiang, X., Satapathy, S. C., & Zhang, Y. D. (2020). *Artificial Intelligence in Sign Language Recognition Systems*. Springer.

2. Zhang, Y., & Jiang, X. (2024). *Deep Learning Advances in Sign Language Recognition*. ResearchGate.

3. Tao, T., Zhao, Y., Liu, T., & Zhu, J. (2024). *A Comprehensive Review of Sign Language Recognition Techniques*. Elsevier.

4. Madhukar, B. N., et al. (2025). *CNN-LSTM Hybrid Model for Real-Time Sign Language Recognition*. IJRASET.

5. Baihan, A., Alutaibi, A. I., & Sharma, S. K. (2024). *Hybrid Attention-Based Deep Learning for Gesture Recognition*. Nature Scientific Reports.

6. Sanjaya, S. A., & Ilone, H. F. (2023). *Systematic Review on Deep Learning for Sign Language Recognition*. IJCS.

7. Wang, M., & Deng, W. (2018). *Deep Face and Gesture Recognition: A Survey*. arXiv.

8. Guo, G., & Zhang, N. (2019). *Deep Learning for Facial and Gesture Analysis*. ScienceDirect.

9. Singh, R., & Kumar, A. (2022). *Real-Time Gesture Recognition Using CNN and LSTM Models*. IEEE Access.

10. Liu, Z., et al. (2023). *Efficient Deep Learning Models for Real-Time Human Action Recognition*. ACM Computing Surveys.

11. Mahesh Ganji. (2025). Enhancing Oracle Cloud HR Reporting Through AI-Driven Automation. Journal of Science & Technology, 10(6), 28–36. https://doi.org/10.46243/jst.2025.v10.i06.pp28-36

12. [17] Todupunuri, A. (2025). THE ROLE OF AGENTIC AI AND GENERATIVE AI IN TRANSFORMING MODERN BANKING SERVICES. American Journal of AI Cyber Computing Management, 5(3), 85–93. https://doi.org/10.64751/ajaccm.2025.v5.n3.pp85-93

13. [18] Todupunuri, A. (2025). Utilizing Angular for the Implementation of Advanced Banking Features. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283395

14. [19] Sushma Babburi. (2025). Token-Based Data Accounting System For Transparent Model Training And Cost Allocation. American Journal of AI Cyber Computing Management, 5(4), 463–474. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp463-474

15. [20] Snigdha Gaddam. (2025). SOFTWARE STACK PREPARED FOR AI TRANSITIONING FROM MODULES TO MODELS. American Journal of AI Cyber Computing Management, 5(4), 451–462. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp451-462

16. [21]Gaddam, S. INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING.

17. [21]Bajarang Bhagwat, V. (2023). Optimizing Payroll to General Ledger Reconciliation: Identifying Discrepancies and Enhancing Financial Accuracy. JOURNAL OF ADVANCE AND FUTURE RESEARCH, 1(4).

https://doi.org/10.56975/jaafr.v1i4.501636

18. [22]Suhasnadh Reddy Veluru, Sai Teja Erukude, and Viswa Chaitanya Marella. 2025. Multimodal Detection of Fake Reviews using BERT and ResNet-50. In 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). IEEE, 877–882.

19. [23]Doragacharla, V. R. (2026). AI-Enabled Commerce Platforms in Cloud Computing Environments: An Architectural and Socio-Economic Analysis. Journal of Computational Analysis & Applications, 35(1).

20. [24]Jay Bharat Mehta. (2025). AUTONOMOUS PATCH VALIDATION FOR ZERO-DAY EXPLOITS IN ENTERPRISE CLOUDS. International Journal of Applied Mathematics, 38(4s), 1270–1285. https://doi.org/10.12732/ijam.v38i4s.685

21. [25]Reddy, S. K. (2025). Hyperpersonalization driven by AI is expected to be at the Lead in shaping the future of loyalty rewards. Journal of Emerging Technologies and Innovative Research.

22. [26]Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.

23. [27]Poojari, R. (2026). Privacy-Preserving Generative AI in Healthcare Systems Using Federated Learning Approaches. International Journal of Data Science and IoT Management System, 5(1), 78-88.

24. [28]Kalae, U. K. (2021). Enhancing data analytics and reporting efficiency using Power BI and SQL in cloud computing environments. Journal of Computational Analysis and Applications, 29(6), 2021. https://doi.org/10.48047/jocaaa.2021.29.06.48

25. [29]Kalae, U. K. (2025). Optimizing cost-effective cloud data pipeline orchestration across multiple cloud providers. Journal of Information Systems Engineering and Management, 10(63s), e726–e741.

26. [30]Saikumar, B. (2023). Enhancing Client Engagement through AI-Driven Real-Time Reporting and Automated Alerts. International Journal of Enhanced Research in Science, Technology & Engineering, 12(11), 111–117. https://doi.org/10.55948/ijerste.2023.1115

27. [31]Saikumar, B. (2024). Optimizing Crew Scheduling and Absence Management using Microservices: Enhancing Reliability and Efficiency in Crew Management Systems. International Journal of Enhanced Research in Management & Computer Applications, 13(11),50–55. https://doi.org/10.55948/ijermca.2024.0116