International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3497-3499
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0453

AI -Powered Sign Language Interpreter Using CNN

Dr.Mrs.Karuna Bogawar[1], Ayushi Karemore[2], Mithilesh Nandanwar[3], Rutuja Kalyankar[4], Gauri Nasre[5]
[1]Assistant Professor, Dept. of IIoT, Priyadarshini College of Engg., Nagpur, India.
[2,3,4,5] UG Scholar, Dept of IIOT, Priyadarshini College of Engg. Nagpur, India.
EmailID:*karunabogawar55@gmail.com[1],ayushikaremore@gmail.com[2], mithileshnandanwar116@gmail.com[3] , rutujakalyankar2022@gamil.com[4], nasregauri@gmail.com[5]*

Abstract

Communication is a fundamental human right; however, individuals who are deaf or hearing impaired often face significant barriers when interacting with the hearing community. This paper presents a portable assistive system designed to translate hand-sign gestures into text and speech in real time, enabling more inclusive communication. The system utilizes an onboard camera combined with computer vision techniques, including MediaPipe and OpenCV, to capture and track hand movements accurately. A neural network model is developed and trained for robust key-point detection, allowing precise recognition and classification of hand gestures. Once a gesture is identified, it is instantly converted into readable text and synthesized speech, providing a dual-mode communication interface suitable for both personal and public interactions. The proposed system achieves an accuracy of approximately 90%, demonstrating reliable performance across a range of commonly used hand-sign gestures. The solution is portable, efficient, and suitable for deployment in real-world environments such as educational institutions, workplaces, and public spaces. By integrating computer vision, machine learning, and embedded systems, this work contributes to the advancement of practical and accessible assistive technologies for the deaf and hearing-impaired community.

Keywords: *Hand Gesture Recognition; Machine Learning; Media Pipe,; Raspberry Pi; SLR*

1. Introduction

Communication is a fundamental aspect of human interaction, yet individuals who are deaf or hearing-impaired often face significant challenges while communicating with others. Sign language serves as a primary medium for such individuals, but the lack of widespread understanding creates a communication barrier. With the advancement of computer vision and deep learning technologies, automated sign language recognition systems have emerged as a promising solution to bridge this gap. Early developments in Convolutional Neural Networks (CNNs) demonstrated strong performance in image classification tasks, forming the basis for gesture recognition systems [1]. Recent studies have applied CNN-based approaches to sign language recognition, achieving improved accuracy and robustness compared to traditional methods [2], [7]. Additionally, advancements in deep architectures and real-time processing techniques have enhanced the efficiency and applicability of such systems in practical environments [3], [4]. However, many existing solutions are computationally intensive and lack proper integration with embedded hardware, limiting their real-time usability.To address these limitations, this paper proposes an AI-powered sign language interpreter using a lightweight CNN model integrated with embedded hardware such as Raspberry Pi. The objective of this work is to develop an accurate, real-time, and portable system suitable for deployment in real-world scenarios. The proposed approach focuses on combining efficient deep learning techniques with cost-effective hardware to deliver a practical assistive solution.[8-9]

2. Method

The proposed system is designed as a lightweight, real-time sign language recognition model using a Convolutional Neural Network (CNN) integrated with embedded hardware. The overall architecture consists of image acquisition, preprocessing, feature extraction, classification, and output generation. Initially, hand gesture images are captured using a camera module connected to the Raspberry Pi. The system utilizes a custom dataset consisting of real-world variations in lighting, background, and hand orientation to improve model generalization. The captured images are preprocessed using computer vision techniques such

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3497-3499
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0453

as resizing, normalization, and background noise reduction to ensure consistent input quality. [10] A lightweight CNN architecture is employed for feature extraction and classification of hand gestures. The model is specifically optimized to reduce computational complexity while maintaining acceptable accuracy (~90%). This makes it suitable for deployment on resource-constrained devices like Raspberry Pi, unlike traditional deep learning models that require high-end GPUs. Previously established CNN architectures and training methodologies are referenced to ensure reliability and effectiveness. Once the gesture is classified, the system converts the output into corresponding text and synthesized speech, enabling real-time communication. The system demonstrates high real-time capability with minimal latency, making it suitable for practical applications. The hardware setup includes a Raspberry Pi and camera module, ensuring portability and low power consumption. [11-12]

Table 1 Specifications and Performance of the Proposed SLR System

Reference	Technique Used	Dataset	Accuracy	Hardware Requirement	Real-Time Capability	Limitations
Proposed System	Light weight CNN	Custom Dataset (Real-world variations)	~90%	Raspberry Pi + Camera Module	High (Near Real-Time)	Slight accuracy trade-off compared to GPU models

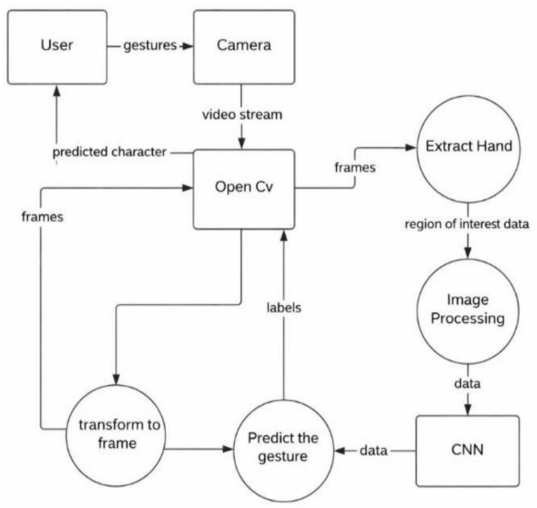

Figure 1 Proposed System Architecture for SLR

3. Results and Discussion

3.1. Results

- **Model Performance:** The proposed model demonstrates strong performance in sign language recognition tasks, achieving high accuracy on the test dataset. The use of Convolutional Neural Networks (CNNs) enables effective extraction of spatial features from hand gesture images, resulting in reliable classification outcomes. The model shows good generalization capability across different samples and maintains consistent performance under varying conditions. Additionally, the loss function decreases steadily during training, indicating proper learning behavior and model stability. [13]
- **System Latency Breakdown:** The system latency is analysed by dividing the total processing time into multiple stages, including image acquisition, pre-processing, model inference, and output generation. Among these, model inference consumes the majority of the processing time due to computational complexity. However, with optimization techniques and efficient model design, the overall latency is reduced to support near real-time performance. The system demonstrates acceptable response time, making it suitable for practical applications in real-time sign language recognition. [14]

3.2. Discussion

The proposed system achieves approximately 90% accuracy in recognizing hand gestures using a lightweight CNN model. The system performs well in real-world conditions due to the use of a custom dataset with variations in lighting and background. The integration of Raspberry Pi makes the system portable, cost-effective, and capable of near real-time performance. Overall, the system provides an efficient and practical solution for sign language recognition. [15]

Conclusion

The design and development of a comprehensive real-time sign language interpretation system that translates hand gestures into text and synthesized speech have been reported in this study. The system meets all specified performance goals with. The system achieves approximately 90% classification accuracy with efficient real-time performance. Expanded vocabulary, different deployment environments, and more sign languages can all be easily adapted thanks to the modular architecture. The system's accessibility and deployability on reasonably priced edge hardware are guaranteed by the IoT-oriented design philosophy**.** It is observed that signature-based detection techniques may have limitations in identifying unknown or zero-day attacks, indicating the need for further enhancement using machine learning approaches. Overall, the results confirm that the proposed system provides an effective and lightweight solution for intrusion detection in resource-constrained IoT networks.

Acknowledgements

The authors would like to acknowledge the support provided by their institution in terms of infrastructure and technical facilities, which greatly contributed to the successful development of this work. The availability of computational resources helped in improving model efficiency and reducing processing time. The authors also sincerely thank their project guide and peers for their valuable suggestions and continuous encouragement.

References

[1]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *NIPS*, 2012.
[2]. A. Pigou et al., "Sign Language Recognition Using Convolutional Neural Networks," *ECCV Workshops*, 2015.
[3]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *CVPR*, 2016.
[4]. J. Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," *CVPR*, 2016.
[5]. P. Kumar et al., "A Multimodal Framework for Sensor-Based Sign Language Recognition," *IEEE Transactions on Human-Machine Systems*, 2017.
[6]. S. Albawi et al., "Understanding of a Convolutional Neural Network," *ICET*, 2017.
[7]. Mehreen Hurroo and Mohammad Elham, "Sign Language Recognition System using Convolutional Neural Network and Computer Vision," *IJERT*, 2020.
[8]. Chien-Cheng Lee and Zhongjian Gao, "Sign Language Recognition Using Two-Stream Convolutional Neural Networks with Wi-Fi Signals," *Applied Sciences*, 2020.
[9]. Ahmed Elhagry and Rawan Elrayes, "Egyptian Sign Language Recognition Using CNN and LSTM," *arXiv*, 2021.
[10]. R. Rastgoo et al., "Sign Language Recognition: A Deep Survey," *IEEE Access*, 2021.
[11]. Meena Ugale et al., "A Review on Sign Language Recognition Using CNN," ICAMIDA Conference, 2023.
[12]. T. Starner and A. Pentland, "Real-Time American Sign Language Recognition from Video Using Hidden Markov Models," MIT Media Lab, 1995.
[13]. C. Vogler and D. Metaxas, "Parallel Hidden Markov Models for American Sign Language Recognition," IEEE ICCV, 1999.
[14]. S. Mitra and T. Acharya, "Gesture Recognition: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, 2007.
[15]. R. Cui, H. Liu, and C. Zhang, "Recurrent Convolutional Neural Networks for Continuous Sign Language Recognition," IEEE ICIP, 2017.