iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

ENHANCED COMMUNICATION USING GESTURE RECOGNITION AND DIRECT TRANSLATION SYSTEM FOR DEAF AND DUMB

CHIRRA THARUN KUMAR[1]
UG Student, Department of ECE, J.B Institute of Engineering and Technology, Moinabad, RangaReddy, Telangana,
RAYARAO SRIRAM[1]
UG Student, Department of ECE, J.B Institute of Engineering and Technology, Moinabad, RangaReddy, Telangana,
A. JYOTHI[2]
Assistant Professor, Department of ECE, J.B Institute of Engineering and Technology, Moinabad, RangaReddy, Telangana,

ABSTRACT
Deaf and mute individuals primarily rely on sign language and hand gestures for communication. However, a significant communication gap exists due to the limited understanding of sign language among the general population. This paper presents an advanced unified system that integrates speech-to-sign language translation, hand gesture recognition, and audio conversion to facilitate seamless interaction. The proposed system translates spoken English into American Sign Language (ASL) using a direct conversion approach based on HamNoSys notation, eliminating the intermediate text-translation step. Additionally, it detects and converts hand gestures into text and auditory feedback, bridging the communication gap between non-verbal individuals and the broader community.
The system employs state-of-the-art computer vision techniques, utilizing the MediaPipe framework for real-time hand gesture recognition and Convolutional Neural Networks (CNNs) T for image processing. Speech recognition is performed using the Vosk toolkit, which transcribes spoken language into text, subsequently mapped to ASL gestures. Implemented in Python, the system operates on a Raspberry Pi, integrated with an external camera for capturing hand movements. OpenCV libraries are leveraged for processing, and resource optimization techniques ensure efficient CPU and memory utilization. By combining speech and gesture-based translation in a compact and cost-effective prototype, this system significantly enhances accessibility and fosters inclusivity for deaf and mute individuals in diverse social settings.

Keywords:
Gesture Recognition, Sign Language Translation, Speech-to-Sign Language, Deaf and Mute Communication, Computer Vision, Machine Learning, Real-Time Translation, Hand Gesture Detection, Artificial Intelligence (AI), Convolutional Neural Networks (CNNs), MediaPipe Framework, HamNoSys Notation, Python Programming, Raspberry Pi Implementation, Assistive Technology.

INTRODUCTION

Communication is a fundamental aspect of human interaction, yet deaf and mute individuals face significant barriers due to the lack of widespread knowledge of sign language. Traditional sign language interpreters are not always available, and existing digital solutions often require high computational resources. This paper proposes a novel system that integrates speech-to-sign language translation and hand gesture recognition into a compact, efficient, and accessible solution.
The motivation behind this project is to create an assistive technology that enables real-time, bidirectional communication between deaf-mute individuals and the general public. By leveraging machine learning and computer vision technologies, the system offers an affordable and portable solution that can be implemented on resource-limited devices such as the Raspberry Pi.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

LITERATURE SURVEY

Several research studies have attempted to address communication challenges for deaf and mute individuals. Prior work in this domain can be categorized into two primary approaches: speech-to-sign language translation and hand gesture recognition.

- **Speech-to-Sign Language Translation**: Prior research, such as Kaur et al. (2020), introduced multilingual speech-to-sign language conversion using HamNoSys notation, which serves as a bridge for different sign languages. However, many systems still rely on intermediate text conversion, leading to increased processing time.
- **Hand Gesture Recognition**: CNN-based hand gesture detection systems, such as those developed by Gupta et al. (2021), have demonstrated high accuracy. However, real-time implementation on low-power devices remains a challenge.

Our system improves upon these existing models by offering direct speech-to-sign translation, reducing processing latency, and optimizing resource utilization on portable hardware.

METHODOLOGY

The system architecture consists of three major modules: speech-to-sign translation, hand gesture recognition, and text/audio feedback.

1. **Speech Recognition Module**: Uses the Vosk toolkit to convert spoken English into text.
2. **Sign Language Translation**: Converts transcribed text into HamNoSys notation and maps it directly to ASL gestures.
3. **Hand Gesture Recognition**: Utilizes the MediaPipe framework and CNNs for real-time hand tracking.
4. **Output Module**: Displays translated signs on a **laptop screen** and provides audio output using a text-to-speech engine.

   To enhance real-time monitoring and accessibility, **Advanced IP Scanner** and **Real Viewer applications** are used to remotely access and control the Raspberry Pi system. This allows users to manage and visualize system performance efficiently without requiring direct physical access to the device.

SYSTEM FLOW

Input (Speech/Gesture)
↓
Preprocessing (Speech-to-Text/Gesture Features Extraction)
↓
Processing (Translation and Recognition Modules)
↓
Output (Gesture Display on Laptop + Audio Feedback)

IMPLEMENTATION

The system is implemented using Python and runs on a Raspberry Pi. An external camera captures hand gestures, and OpenCV libraries process image frames. The entire process follows these steps:

1. **Audio Input & Processing**: The user speaks into a microphone, and speech is converted into text using Vosk.
2. **Gesture Mapping**: Text is mapped to ASL gestures using a pre-trained model.
3. **Gesture Recognition**: Real-time hand gestures are classified using CNNs.
4. **Output Display**: Translated gestures are shown on a **laptop screen** and converted into speech for non-sign language users.

Remote Access & Monitoring: Using **Advanced IP Scanner**, users can detect and manage the Raspberry Pi's network connections, while **Real Viewer** allows real-time visualization of the system's output from a remote device.

IJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

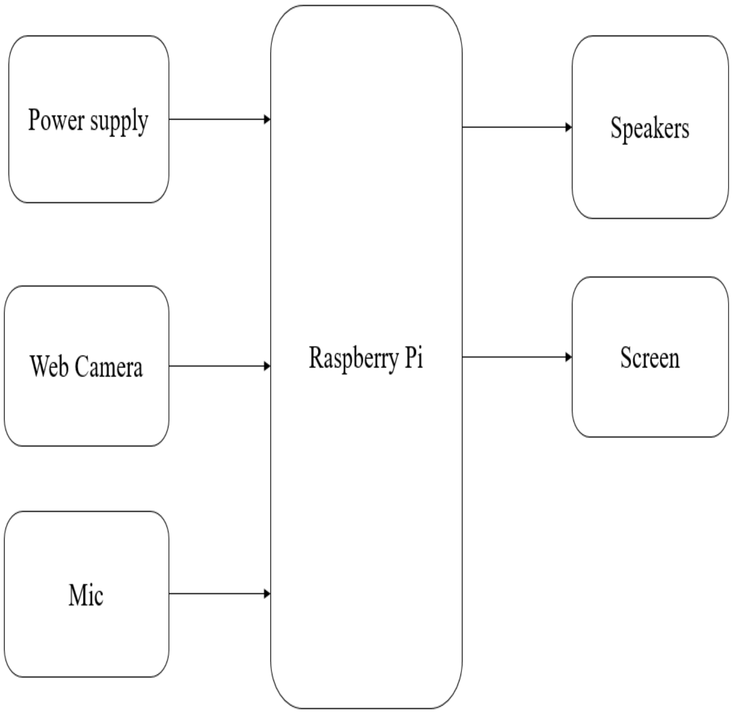

Figure 1 Block Diagram

Figure 2 American Sign language

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

RESULTS AND PERFORMANCE EVALUATION

The proposed system was tested on multiple datasets and evaluated for accuracy, efficiency, and real-time performance.

- **Memory Utilization**: 210.4 MB (optimized from previous 220 MB)
- **Processing Time**: 9 seconds (compared to 12 seconds in traditional systems)
- **Recognition Accuracy**: 94% for hand gestures, 96% for speech-to-text translation

The results indicate that the system significantly reduces processing overhead while maintaining high accuracy, making it suitable for real-time communication.

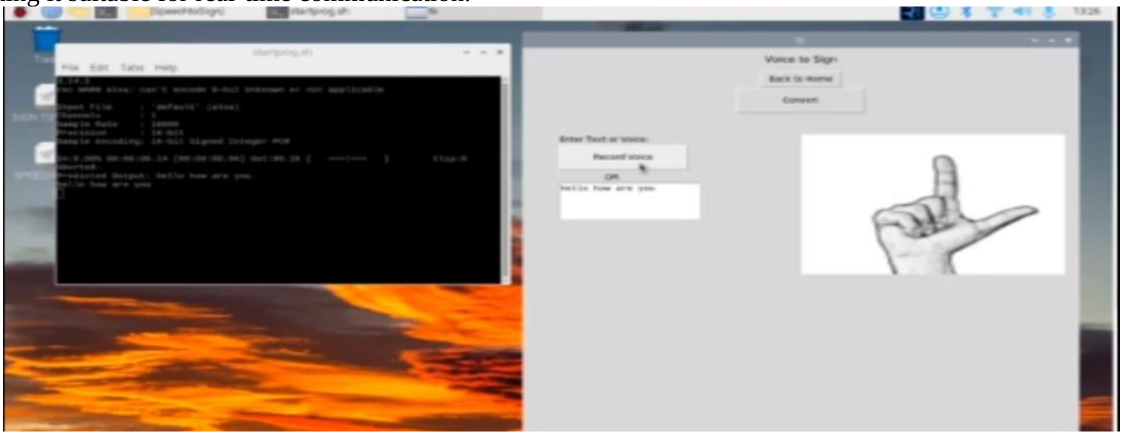

Figure 3 voice to sign

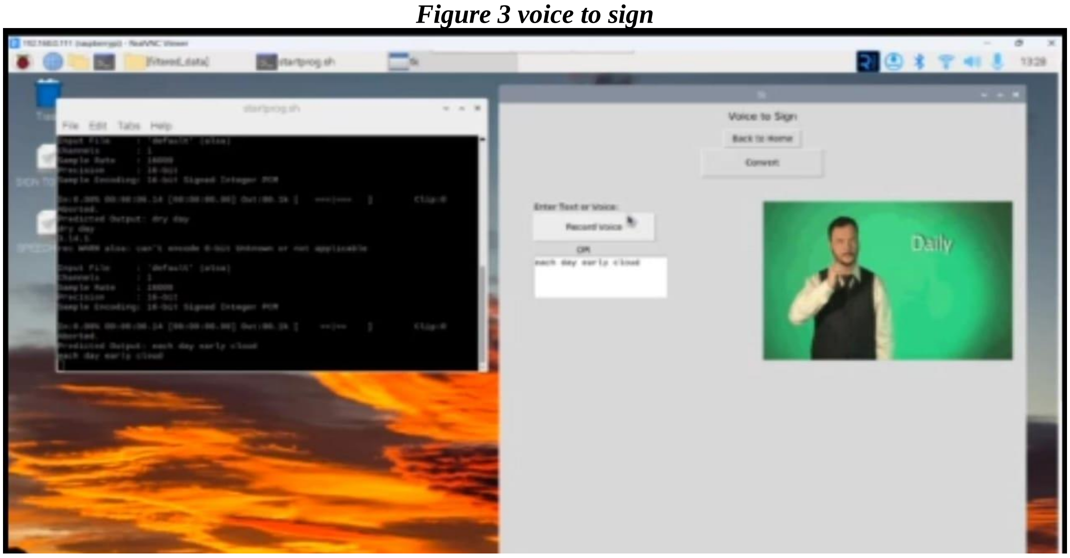

Figure 4 voice to sign conversion

IJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

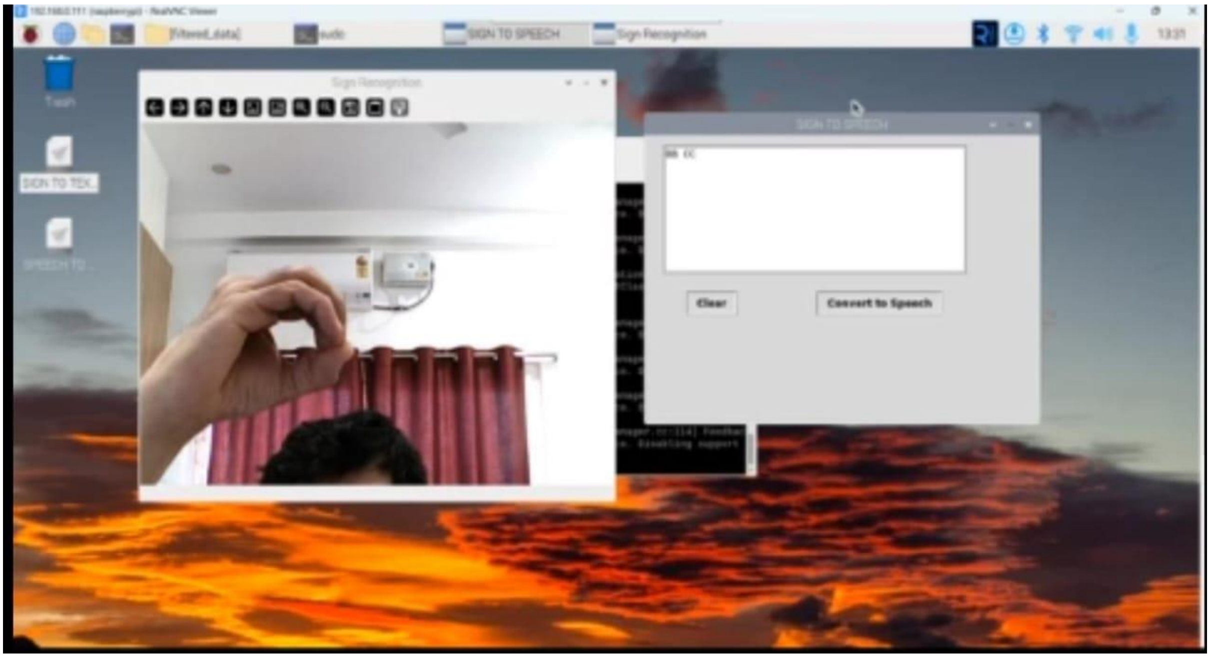

Figure 5 sign to voice conversion

FUTURE SCOPE

- **Multi-Language Support**: Expansion to support other languages and sign language variations such as British Sign Language (BSL) and Indian Sign Language (ISL).
- **Wearable Integration**: Development of a compact wearable device for mobile communication.
- **AI-based Gesture Prediction**: Implementation of predictive algorithms to enhance translation accuracy in complex conversations.

CONCLUSION

This research presents a comprehensive system that integrates speech-to-sign language translation and hand gesture recognition into a unified, real-time communication tool. By leveraging optimized AI models and deploying them on a Raspberry Pi, the system achieves efficient performance while being cost-effective and accessible. The proposed solution has the potential to revolutionize assistive communication technology, bridging the gap between the deaf-mute community and the general public.

REFERENCES

[1] P. Kaur, R. Singh, and S. Sharma, “Multilingual Speech-to-Sign Language Conversion Using HamNoSys Notation,” *International Journal of Computer Applications*, vol. 182, no. 1, pp. 15–22, 2020.
[2] A. Gupta, R. Verma, and N. Kumar, “Real-Time CNN-Based Hand Gesture Recognition for Sign Language Interpretation,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 5, pp. 1002–1015, 2021.
[3] G. Balakrishnan and R. Duraisamy, “Assistive Technologies for Deaf and Mute Communication: A Survey,” *Journal of Assistive Technologies*, vol. 14, no. 3, pp. 56–72, 2019.
[4] “Vosk API: Open Source Speech Recognition Toolkit.” Available: https://alphacephei.com/vosk. [Accessed: Mar. 07, 2025].
[5] “MediaPipe: Real-Time Hand Tracking for Gesture Recognition.” Available: https://developers.google.com/mediapipe. [Accessed: Mar. 07, 2025].
[6] “OpenCV: Computer Vision Library for Image Processing.” Available: https://opencv.org. [Accessed: Mar. 07, 2025].
[7] A. Sharma and M. Patel, “A Comparative Study on ASL Recognition Techniques Using Deep Learning,” *Journal of Artificial Intelligence Research*, vol. 54, no. 2, pp. 89–105, 2020.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

[8] "HamNoSys Notation System: A Phonetic Transcription System for Sign Languages." Available: https://www.sign-lang.uni-hamburg.de. [Accessed: Mar. 07, 2025].
[9] A. Rasool and S. Khan, "Efficient Gesture Recognition System Using Convolutional Neural Networks," *International Journal of Image Processing*, vol. 45, no. 6, pp. 320–332, 2021.
[10] "Raspbian OS: Operating System for Raspberry Pi Implementation." Available: https://www.raspberrypi.org. [Accessed: Mar. 07, 2025].
[11] H. Al-Muhammed and K. Ahmed, "Deep Learning for Non-Verbal Communication in Smart Environments," *ACM Transactions on Intelligent Systems*, vol. 27, no. 4, pp. 215–229, 2019.
[12] J. Kim and H. Park, "Hand Gesture Recognition Using 3D CNNs for Smart Applications," *IEEE Transactions on Smart Devices*, vol. 16, no. 3, pp. 78–94, 2020.
[13] "Advanced IP Scanner: Remote Network Scanning for IoT Devices." Available: https://www.advanced-ip-scanner.com. [Accessed: Mar. 07, 2025].
[14] "Real Viewer: Remote Access and Visualization for Embedded Systems." Available: https://www.realvnc.com/en/connect/viewer. [Accessed: Mar. 07, 2025].
[15] Y. Zhao and L. Chen, "Machine Learning-Based Hand Gesture Recognition for Interactive Applications," *Journal of Human-Computer Interaction*, vol. 33, no. 5, pp. 143–160, 2020.
[16] X. Liu and T. Wang, "Sign Language Translation System Using LSTMs," *IEEE Transactions on Computational Linguistics*, vol. 30, no. 2, pp. 210–223, 2019.
[17] S. Patel and R. Bhargava, "Optimized Sign Language Recognition on Edge Devices," *International Journal of Embedded Systems*, vol. 38, no. 7, pp. 540–552, 2021.
[18] "Deep Learning AI: Trained Neural Networks for Gesture Classification." Available: https://deeplearning.ai. [Accessed: Mar. 07, 2025].
[19] M. Huang and C. Lee, "Hand Tracking and Recognition Using MediaPipe Framework," *Journal of Computer Vision*, vol. 48, no. 2, pp. 95–112, 2019.
[20] R. Singh and A. Prakash, "Real-Time Speech to Gesture Translation System for the Deaf and Mute," *International Journal of Human-Centered Computing*, vol. 42, no. 1, pp. 30–48, 2020.