Bidirectional Sign Language Translator Using Machine Learning and 3D Avatar Technology

Mrs. Benita Mary BE, ME,
Assistant Professor, Computer Science and Engineering,
St. Joseph College of Engineering, Kanchipuram – 602117, Tamil Nadu,
Email Id – benitamary@gmail.com

Mr. Abishek J,
Student, Computer Science and Engineering,
St. Joseph College of Engineering,
Email Id – abishekjosephjegan@gmail.com

Mr. Appasamy I,
Student, Computer Science and Engineering,
St. Joseph College of Engineering,
Email Id – arunkkumar972@gmail.com

ABSTRACT:

Communication barriers between hearing-impaired individuals and the general population remain a major social challenge. Existing sign language translation systems are often unidirectional, limited in accuracy, or lack real-time responsiveness. This paper presents a Bidirectional Sign Language Translator using Machine Learning and 3D Avatar Technology, designed to enable seamless communication between sign language users and non-signers. The proposed system integrates MediaPipe for real-time hand landmark detection, a hybrid ensemble model combining Convolutional Neural Networks (CNN) and Random Forest classifiers for accurate gesture recognition, and Large Language Models (LLM) for contextual text refinement. The recognized gestures are converted into readable text and audible speech. Conversely, spoken or typed language is translated into sign language using a 3D animated avatar, enabling expressive visual communication. Built using Python, TensorFlow, FastAPI, and React, the system ensures real-time performance through webcam-based input processing. This approach reduces dependency on human interpreters and promotes accessibility, inclusion, and independence for hearing-impaired individuals.

INTRODUCTION

Sign language serves as a primary mode of communication for millions of hearing-impaired individuals worldwide. However, the lack of universal understanding of sign language creates communication gaps between signers and non-signers. Traditional translation systems often support only one-way communication and fail to provide real-

time interaction. The emergence of artificial intelligence and computer vision technologies offers new opportunities to bridge this gap. This project proposes a real-time bidirectional sign language translation system that enables both sign-to-text and text-to-sign communication using advanced machine learning models and 3D avatar technology.

OBJECTIVES OF THE PROJECT

The objectives of this project are designed to address the limitations of existing systems and provide a complete communication solution:

- To design a real-time bidirectional sign language translation framework.
- To accurately recognize hand gestures using machine learning techniques.
- To convert recognized gestures into readable text and audible speech.
- To translate text or spoken language into sign language using a 3D animated avatar.
- To reduce dependency on human interpreters.
- To enhance accessibility and social inclusion for hearing-impaired individuals.

SYSTEM METHODOLOGY

The system architecture consists of two primary modules: Sign-to-Text Translation and Text-to-Sign Translation. MediaPipe is used for real-time hand tracking and extraction of 3D hand landmarks. The extracted features are processed using a hybrid ensemble machine learning model combining CNN and Random Forest classifiers for high-accuracy gesture recognition. Large Language Models are integrated to refine predicted outputs into grammatically correct sentences. The translated text is converted into speech using text-to-speech engines. For reverse translation, text or speech input is processed and mapped to corresponding sign animations displayed through a 3D avatar developed using modern frontend frameworks.

INNOVATION AND DIFFERENTIATION

The proposed system introduces several innovative features that distinguish it from existing solutions. First, it enables real-time bidirectional communication rather than a single-direction translation. Second, the hybrid ensemble learning model improves recognition accuracy compared to standalone classifiers. Third, the integration of a 3D animated avatar provides expressive and visually accurate sign representation, enhancing user experience. Additionally, the use of Large Language Models ensures contextual correctness and sentence refinement.

TECHNOLOGIES USED

- MediaPipe for hand landmark detection.
- Convolutional Neural Networks (CNN) for feature learning.
- Random Forest classifier for ensemble prediction.
- Large Language Models (LLM) for contextual refinement.
- TensorFlow for deep learning implementation.
- FastAPI for backend service deployment.
- React for frontend development.
- Python as the primary programming language.
- Real-time webcam integration for live gesture capture.
- 3D Avatar technology for sign animation output.

CHALLENGES

Despite promising results, the system faces challenges such as varying lighting conditions affecting gesture detection, differences in signing styles among users, computational complexity in real-time processing, and the need for extensive datasets for improving model generalization. Future improvements can focus on expanding vocabulary, incorporating facial expression recognition, and optimizing performance for mobile deployment.

CONCLUSION

The Bidirectional Sign Language Translator using Machine Learning and 3D Avatar Technology represents a significant step toward inclusive communication. By integrating computer vision, deep learning, and intelligent avatar-based animation, the system provides a comprehensive real-time communication solution. The hybrid ensemble approach enhances recognition accuracy, while the 3D avatar ensures effective visual translation. This project contributes to reducing communication barriers, promoting accessibility, and empowering hearing-impaired individuals in educational, professional, and social environments.

REFERENCES

[1] S. Mitra and T. Acharya, “Gesture Recognition: A Survey,” IEEE Transactions on Systems, Man, and Cybernetics, 2025.

[2] C. Zhang et al., “Real-Time Hand Gesture Recognition Using CNN,” International Journal of Computer Vision, 2024.

[3] Google MediaPipe Documentation, 2024.

[4] TensorFlow Official Documentation, 2024.

[5] Recent Advances in Sign Language Translation Using Deep Learning, Journal of AI Research, 2024.