Advances in Computer Engineering and Science (ACES)
(2026)
Volume -1, Issues – 1

Real Time Sign language Text to Speech Translator

P. Kavya[1],
M.Sc. Computer Science[1],
Department of Computer Science and Applications,
Vivekananda College of Arts and Sciences for Women (Autonomous),
Elayampalayam,Tiruchengode,
E-Mail: pugalkavya283@gmail.com

Mr.V. Arunkumar[2],
Assistant Professor/CS & CA[2],
Department of Computer Science and Applications,
Vivekananda College of Arts and Sciences for Women (Autonomous),
Elayampalayam.Tiruchengode,
E-mail: veere.arun@gmail.com

Abstract: **The communication barrier amongst the deaf and speech impaired individuals constitutes a major issue in the society. They use Sign language mostly and majority of the people have not been trained to understand them. The project is concerning the design and development of Real Time Sign Language Text to Speech Translator that can be utilized to translate hand gestures into readable text and audible speech in real time. The proposed system relies on computer vision and machine learning algorithms to recognize hand gestures captured by a camera. The system supports live video feed, uses hand gestures, extracts meaningful features and identifies the gestures using a trained deep learning model. Gestures that are known are translated into a corresponding text and a text to speech engine translates it into speech. The application will be used to provide a simple and convenient solution to the communication barrier between the hearing impaired and the rest of the society. Python and OpenCV are the primary programming language and images and models are processed and trained using TensorFlow, NumPy, and OpenCV libraries. This system was tested under real-time conditions and was found to be precise as far as gesture recognition with little delay is concerned. It is a comprehensive solution that promotes communication in the community, schools, and place of work.**

Keywords: **Sign Language Recognition, Text to Speech, Computer Vision, Deep Learning, OpenCV, Tensorflow, Real-Time System, Assistive Technology.**

1. Introduction

Sign language is non-verbal language meaning that it involves the use of signs like hand movements, facial expression and movements to convey a message. It is an extremely significant deaf mute communication tool[1]. Though this is a significant element, there is a huge obstacle to communication because of ignorance and absence of awareness of the sign language by the rest of the population. The Real Time Sign Language Text to Speech Translator aims to fill this gap by automatically translating sign language gestures into text and audio[2]. The system captures the live hand gestures with the help of a camera and then, processes the images and applies machine learning to recognize the signs appropriately. This gesture is then converted into text and speech is synthesized by a speech synthesis module. The project is connected with the use of the newest technologies such as computer vision and deep-learning to reach high precision and real-time features. The application can be used in the schools, hospitals, public service centers as well as in the normal conversations in order to facilitate inclusive communication. Studies indicate that real-time sign language translation system is capable of significantly increasing the accessibility of communication in case optimized in both speed and accuracy[3].

2. Literature Review

Image processing and machine learning have been utilized to carry out several studies on sign language recognition. Initial systems used sensor based gloves, to monitor hand movements, which were costly and uncomfortable in use[4]. As computer vision has developed, camera-based systems are gaining popularity because they are affordable and easy to use. Scientists have also examined different methods including convolutional neural networks (CNNs) to identify gestures since it achieves a high level of accuracy when it comes to classifying images. OpenCV is also popular in hand detection and image preprocessing, whereasTensorFlow and Keras are used to train deep learning models[5]. Research indicates that real time sign language translators may be used to a great effect in enhancing access to communication where it is optimized in speed and accuracy[6].

3. Methodology

The suggested Real Time Sign Language Text to Speech Translator is built following the Software Development Life Cycle (SDLC) methodology. First, the requirements for the system were analysed to determine the need for real-time gesture recognition, translation, and speech[7]. Images of hand gestures were captured using a camera and processed using OpenCV operations like resizing, background subtraction, and denoising[8]. A deep learning system was developed with TensorFlow to classify a set of gestures from the preprocessed images. In real-time mode, video frames are processed in real-time, and the model predicts the gesture[9]. The detected gesture is then translated into text and synthesised into speech using a text-to-speech converter. The system was evaluated for accuracy, speed and stability to provide real-time performance[10].

3.1 Working principle:

The Real Time Sign Language Text to Speech Translator operates on the principles of real-time image capture, gesture recognition and speech generation[11]. The system acquires real-time video frames of the user's hand gestures using a camera. The frames are analysed by OpenCV to locate the hand and extract relevant features. The extracted features are fed into a pre-trained deep learning model to recognise the sign language gesture[12]. The detected gesture is then translated into text and shown on the screen. Lastly, the text is converted into speech using a text-to-speech engine, facilitating real-time communication between hearing-impaired and hearing people.

3.1.1 Required Analysis:

In this phase, the functional and non-functional requirements of the Real Time Sign Language Text to Speech Translator are identified. The system must be capable of capturing hand gestures

in real time, recognizing sign language accurately, converting it into text, and generating speech output. Performance, usability, and real-time responsiveness are considered as important non-functional requirements. User needs such as ease of use and accessibility are also analysed[13].

3.1.2 System Design

The system design is based on the requirements. It specifies the input, processing and output units[14]. Camera is used as the input, image processing and gesture recognition as the processing and text and speech generation as the output. Software components, such as OpenCV, machine learning models and text-to-speech, are specified to interact.

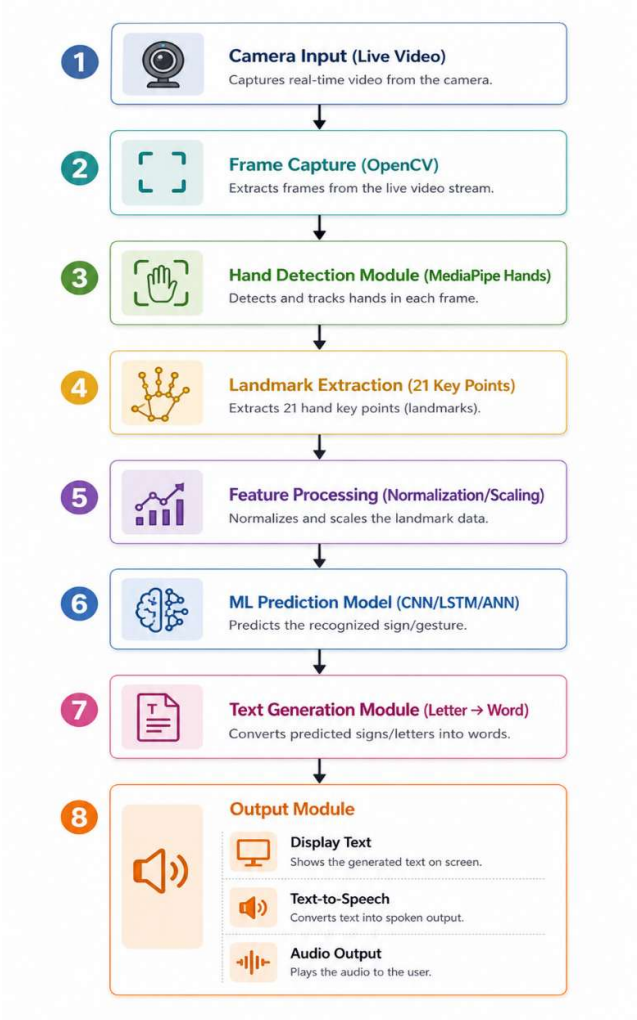

Figure 1 : Farmwork of translator system

3.1.3 Database Design

A database is used to store labels of gestures, model data, and other system configurations. It is stored in a structured manner for easy access and management. While real-time processing and operation is the focus of this system, the database provides support for model training, gesture mapping and future extensions.

3.1.4 Development

In the development stage, the system is developed in Python. OpenCV is used for image manipulation, TensorFlow is used for training and inference of gesture recognition model, and a text-to-speech library is used to output the speech. The system is built in modules which are then integrated together.

3.1.5 Testing

The system is tested under various scenarios to determine its accuracy, speed and stability. Unit and system tests are conducted to confirm proper functioning of gesture recognition, text processing and speech generation. Bugs discovered are resolved to enhance system functionality.

3.1.6 Implementation

The system is tested under various scenarios to determine its accuracy, speed and stability. Unit and system tests are conducted to confirm proper functioning of gesture recognition, text processing and speech generation. Bugs discovered are resolved to enhance system functionality.

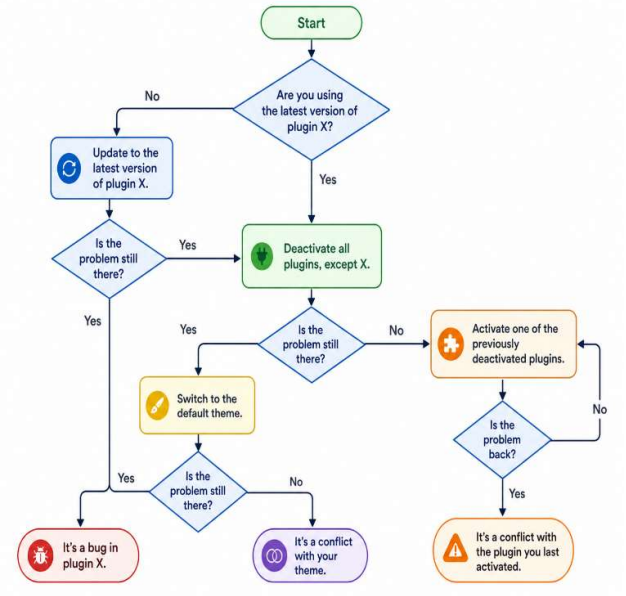

Figure 2 : Overflow of translator

3.1.7 Maintenance

Maintenance includes updates to the system to enhance accuracy, to support additional gestures and to rectify problems encountered while using the system[15]. Periodic upgrades enhance performance and flexibility to user needs.

4. Result and Discussion

The developed system performed well in the experiments. It successfully identified static gestures in sign language and showed the text on the screen. The text-to-speech module accurately translated the text to audio. The system showed good performance in normal lighting conditions and was able to function in real-time with satisfactory accuracy. Our findings indicate that the system works for simple sign language translation.

4.1 Discussion

The findings show that a vision-based system is an effective approach to sign language recognition. The application of deep learning models resulted in better accuracy for gesture recognition. But the recognition accuracy was marginally reduced in low lighting or with complex backgrounds. Although the method was sensitive to light and background, the system provided reliable results and was easy to use. This system can be improved for more complex gestures and practical conditions through further optimization.

5. Future Enhancement

The system can be enhanced by handling dynamic gestures and formation of sentences. Facial recognition can be used to improve sign recognition. Support for multiple sign languages and languages for speech can be added. A mobile app version will enhance accessibility and convenience. Offloading processing to the cloud and using more sophisticated deep learning techniques can also be explored to enhance scalability and accuracy.

6. Conclusion

The Real Time Sign Language Text to Speech Translator effectively overcomes communication barriers for hearing- or speech-impaired people. The system uses computer vision, deep learning and text to speech to convert sign language gestures into text and speech in real time. This project has developed a low cost and assistive system to promote accessibility. The system's results demonstrate its effectiveness and practical use and lays the groundwork for future improvements and implementation.

Reference

[1] P. Duraisamy, A. Abinayasrijanani, M. A. Candida, and P. D. Babu, "Transforming sign language into text and speech through deep learning technologies," *Indian Journal of Science and Technology,* vol. 16, no. 45, pp. 4177-4185, 2023.

[2] P. Verma and K. Badli, "Real-time sign language detection using TensorFlow, OpenCV and Python," *International Journal for Research in Applied Science & Engineering Technology (IJRASET),* vol. 10, no. V, p. May2022, 2022.

[3] M. Prabhakar, P. Hundekar, S. Deepthi, S. Tiwari, and V. Ms, "Sign language conversion to text and speech," *Journal of Emerging Technologies and Innovative Research (JETIR),* vol. 9, no. 7, 2022.

[4] H. Kumar *et al.*, "Sign language detection and conversion to text using CNN and OpenCV," in *AIP Conference Proceedings*, 2022, vol. 2555, no. 1: AIP Publishing LLC, p. 040016.

[5] A. Pathak, N. Patil, P. Padhy, A. Jadhav, S. Rukhande, and L. Gadhikar, "Real-time sign language to speech converter using OpenCV and MediaPipe," in *2025 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, 2025: IEEE, pp. 32-36.

[6] G. Sharma, M. Dominic Savio, V. Pandiyaraju, P. Anandan, N. Paul, and R. Ghosh, "Indian Sign Language Recognition System Using OpenCV and TensorFlow," in *International Conference on Advances in Data-driven Computing and Intelligent Systems*, 2024: Springer, pp. 297-320.

[7] B. Natarajan *et al.*, "Development of an end-to-end deep learning framework for sign language recognition, translation, and video generation," *IEEE Access,* vol. 10, pp. 104358-104374, 2022.

[8] P. Sonsare, D. Gupta, J. Sayyed, P. Nandha, and S. Laira, "Sign language to text conversion using deep learning techniques," in *2024 2nd International Conference on Computer, Communication and Control (IC4)*, 2024: IEEE, pp. 1-6.

[9] T. J. Betri, "Gesture to Speech Transformation: A Real-Time Sign Language Alphabet Recognition System for Digital Inclusion of the Speech Impaired Community," *Citizen: Jurnal Ilmiah Multidisiplin Indonesia,* vol. 6, pp. 1-10, 2026.

[10] M. S. Parveen, R. Keerthana, S. Shanmathi, and M. Shajitha, "Sign Language Detection Using Open Cv," in *2024 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, 2024: IEEE, pp. 1-5.

[11] A. Pandey, "Real-Time Sign Language Recognition Using Deep Learning and Computer Vision," 2025.

[12] M. W. Mairaj, "Assistive Vocalizer Using Opencv For Sign Language Recognition And Translation."

[13] B. Bakariya, "Sign Language Recognition-Based Machine Learning Model for Hearing Disabilities Person," in *Applied Assistive Technologies and Informatics for Students with Disabilities*: Springer, 2024, pp. 113-133.

[14] S. S. Rajest and R. Regin, "Enhancing Communication: Real-Time Sign Language Detection with Deep Learning and Python," *Central Asian Journal of Medical and Natural Science,* vol. 5, no. 4, pp. 940-951, 2024.

[15] K. Sanjusaran, S. Shakthipriyan, and R. Supreetraju, "A REAL TIME INDIAN SIGN LANGUAGE RECOGNITION USING TENSORFLOW," *International Journal of Engineering Research and Sustainable Technologies (IJERST),* vol. 2, no. 4, pp. 26-33, 2024.