



Sign Language to Text and Speech Conversion: A Review

Prof. Hole P P^a, Andhale Sneha^b, Berad Asmita^b, Ramnani Divya^b

^a HOD and Professor, Computer Engineering, Sau. Sundarbai Manik Adsul Polytechnic, Chas, Ahilya Nagar, Maharashtra 414005, India

^b Student, Computer Engineering, Sau. Sundarbai Manik Adsul Polytechnic, Chas, Ahilya Nagar, Maharashtra 414005, India

ABSTRACT:

The rapid evolution of technology in everyday makes it important for the development of effective communication help for the people with hearing and vocal disabilities. Sign language, a crucial medium for the deaf and mute community, remains underutilized due to the general population's lack of understanding. Though the world is developing and getting digital, there are still few issues that haven't completely resolved it. It is difficult for normal people to learn sign language. This causes a communication drift between normal people and deaf and mute people. Researchers are trying to bridge this gap by implementing various techniques.

This paper provides a comprehensive review of recent research that took place in sign language conversion. We analyse the variety of different technological advancements that took place in this field, from vision-based recognition systems generated by deep learning to use of noteworthy machine learning algorithms. This review gives prominence to the strengths, features and limitations of various approaches, the difficulty of capturing the hand formations of sign language, and the significance of precise execution and underscores the significance of bridging communication gaps and fostering inclusivity through technological advancements.

Keywords: Sign language, Gesture recognition, Computer vision, Machine learning, CNN.

Introduction:

Communication plays important role in an individual's life. It is a fundamental aspect of human life which helps connect people, exchange information, ideas and way to express their emotions and feelings. However, not all individuals are privileged with this ability. [5]According to the World Health Organization (WHO), over 1.5 billion people globally experience hearing loss constituting nearly 20% of the world's population. For individuals with hearing and vocal disabilities, sign language is a primary means of communication.

Despite its importance, sign language is not widely understood by the general population, creating a significant communication gap. Many individuals with disabilities rely on interpreters, fostering dependency and limiting their autonomy.

To address this, researchers have developed automated systems capable of translating sign language into text and speech, leveraging technologies like computer vision, machine learning, and sensor-based methods. This paper provides insights of different research that took place in the recent times by various researchers and experts on Sign Language Recognition. The common shared goal of these research is to bring together normal people and deaf-dumb people. To create a system that converts the hand gestures into text and speech. This would help to build an environment comfortable and safe for deaf-dumb people to express themselves.

Literature Review:

Bharat A Manoj proposed a system for Conversion of Sign Language to Text and Speech and Prediction of Gesture. This proposed system uses an Arduino Uno board, flex sensors and an Android Application. The flex sensors are fixed on a glove for detecting the gestures and hand formations of a wearer. The detected gesture is further sent as a corresponding message to the mobile device using GSM (Global System for Mobile) module. It also sends the values to ThingSpeak cloud from where the values can be recovered for further predictions. This device also uses Particle-swarm optimization method and Back-propagation algorithm to fetch accurate solutions and provide predictions by reducing errors. However, its reliance on wearable hardware limits user convenience and scalability. [1]

Akash Kamble, et.al. developed a Conversion of Sign Language to Text system. The emphasis of this system is to predict Indian Sign Language in Real-Time. This proposed system is based on the combination of computer vision, machine learning and natural language processing (NLP) techniques. It consists of use of Media Pipe Library and multi-layered LSTM (Long Short-Term Memory) model. Media Pipe Library is used for detection and recognition of key points of hand gestures in real-time which are further used to extract features. These extracted features are treated as data points which

are compared with the database and LSTM model is implemented for converting the recognized gestures into corresponding text. The system provides the result with an accuracy of 96.66%. However, it lacks speech synthesis integration, which would make it more comprehensive for users.[2]

Kaniti Kavya Sushma Mounika et.al. proposed a model for Audio-To-Sign Conversion and Hand Gesture Recognition with An Air Board for Deaf and Dumb Using Deep Learning. The suggested model is a desktop-based application, created and implemented using python programming language. It has three key characteristics: (i) Audio to Sign Conversion using Speech recognition library of python. (ii) Hand Gesture Recognition using Convolutional Neural Network (CNN) model and (iii) Air Board uses OpenCV a python module for capturing live frames of hands, this feature can detect Hindi letters. The average accuracy rate of this model is better compared to other models. This model can be further improved in future by detecting the hand gestures even with the noisy background and expand its languages of conversion. [3]

Sasikala Dhamodaran, et.al. in Implementation of Hand Gesture Recognition using OpenCV developed a model by combining different domains. The model is developed using different python libraries and modules, it uses open-source framework from Google Media Pipe. Collecting input from the webcam. This proposed model provides user with different functionalities such as “Volume Control”, “Brightness Control”, “Virtual Mouse”, “Arrow-key Control” and “Zoom-In and Zoom-Out Controls”. The major benefit of this model is touch-free working. While innovative in its touch-free interaction design, the system is not tailored for comprehensive language translation, limiting its broader applicability.[4]

Approaches Used in Sign Language Recognition:

Vision-based approach: Vision-based approach is the most preferred approach for sign language recognition. This method looks at pictures to recognize hand shapes. It involves use of computer vision, camera for detection of hand gestures. Different Image processing algorithms like object detection, feature extraction, morphological image processing and Gaussian Image processing algorithms are some of the frequently used algorithms in this technique for tracking the formations made by the signer. Another technique in this approach is Video Analysis in which instead of single picture, it examines a sequence of pictures (video) to capture how signs change over time. This approach has a benefit that it the signer does not require any hardware to wear.

Sensor-based approach: In Sensor-based approach a signer needs to wear a glove which has a number of sensors attached to it. This glove tracks the position of hands and convert the gestures into digital signals that a computer system can interpret. It has a motion sensor too. Though this approach provides easy way of detection of gestures, The studies revealed while reliable, this approach is less practical due to the need for additional hardware, which can be inconvenient for everyday use.

Kinect: Kinect is a motion/action sensing device that can track and detect different body movements and gestures. This approach uses depth camera and then use algorithm for recognition of signs. This approach is considered a better choice as compared to other approaches the reason for that is its execution is not affected by the surrounding light conditions.

Machine Learning and Deep Learning: In consists of machine learning techniques like feature extraction and classification for extracting important details like position and speed of hands and the further classify these features. Deep Learning include powerful neural networks like Convolutional Neural Network (CNNs) and Recurrent Neural Networks (RNNs) used for analyzing the visual data and converting it into sequential output.

Hybrid approach: Hybrid approach combines two or more machine learning techniques to create a robust system. Like combining Vision based approach with Sensor based approach. This can result in better accuracy of recognizing of hand gestures.

Conclusion :

In conclusion, advancements in sign language recognition technologies mark significant progress toward bridging communication gaps for the deaf and mute community. This review highlights key approaches, including vision-based, sensor-driven, and machine learning techniques, while acknowledging their limitations, such as hardware dependency and restricted language support. Future developments should focus on reducing these barriers by creating robust, cost-effective, and multilingual systems capable of delivering high accuracy in diverse real-world environments.

Based on the evaluation it can be concluded that the literature review conducted of the papers works towards a common goal of trying to make it easier for hearing impaired people to live in our day to day with less difficulties and express themselves comfortably. In future, systems would be developed with less limitations and drawbacks. By fostering greater inclusivity, these systems have the potential to empower individuals with hearing and vocal disabilities, allowing them to participate more actively in society and improve their quality of life.

REFERENCES :

1. Bharat A. Manoj (2020). “Conversion of Sign Language to Text and Speech and Prediction of Gesture.” International Journal of Recent Technology and Engineering (IJRTE) (Vol. 8, Issue 6, pp. 4191–4194).
2. Kamble, A. (2023). “Conversion of Sign Language to Text.” International Journal for Research in Applied Science and Engineering Technology (Vol. 11, Issue 5, pp. 1963–1968).

3. Kaniti Kavya Sushma Mounika, Juttuka Satya Hemalatha, Ghantasala Divya Sri, Mylavarapu Kalyan Ram. (2023). "Audio-To-Sign Conversion and Hand Gesture Recognition with An Air Board for Deaf and Dumb Using Deep Learning." International Journal of Research Publication and Reviews (Vol. 4, Issue 4, pp. 1542–1550).
4. Dhamodaran, S., Phukan, P. P., Singh, M., & Nandakum, S. (2024). "Implementation of Hand Gesture Recognition using OpenCV." In International Journal of Research Publication and Reviews (Vol. 5, Issue 5, pp. 8039–8050).
5. World Health Organization (WHO). Hearing Loss Statistics. Available at: <https://www.who.int>.