© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

AI-powered Deaf Comparison System for Inclusive Communication Between Deaf and Hearing Individuals

MR. V.KEERTHISWARAN(1), MS.S.SANGEETHA(2)
Student (1), Assistant Professor (2)
Master of Computer Applications (1, 2)
Dr.M.G.R. Educational and Research Institute, Chennai (1, 2)

***Abstract*- This paper presents an AI-powered system designed to bridge communication gaps between deaf and hearing individuals. By employing comparison algorithms and intelligent interpretation models, the system can translate and interpret communication for inclusive dialogue. The project focuses on recognizing input patterns from both parties and delivering appropriate, comprehensible outputs to facilitate mutual understanding.**

I. INTRODUCTION

Communication between deaf and hearing individuals is often obstructed due to differences in language and interpretation methods. This system addresses the issue by utilizing artificial intelligence to compare and understand sign-based inputs and generate corresponding textual or voice outputs.

PROJECT OBJECTIVES
to create a real-time, AI-powered communication system that can interpret and translate sign language into written or spoken language and vice versa.

to use intelligent interpretation models to facilitate smooth two-way communication between hearing and deaf people. to put into practice pattern recognition algorithms that precisely recognise and handle user input signals like speech, text, and sign gestures.

to create an intuitive user interface that ensures accessibility and usability for both hearing and deaf users.

to use real-world testing scenarios to assess the system's performance in terms of user satisfaction, processing speed, and translation accuracy.

to support a variety of sign language variations and adjust to different speech accents and linguistic styles in order to guarantee the solution's inclusivity.

II. SCOPE

The creation and implementation of an AI-powered system that enables efficient communication between hearing and deaf people is included in the project's scope.

Sign language gestures, spoken words, and textual content are just a few of the input formats that the system can interpret and translate in real time into understandable outputs.

The following are the main areas of focus for the project:

Sign Language Recognition: The use of computer vision and machine learning methods to identify and decipher facial expressions and hand gestures related to sign language.
Converting spoken language into text and vice versa through the use of natural language processing and speech recognition technologies allows for two-way communication.

Low-latency processing is ensured in real-time translation to facilitate synchronous and seamless user interaction.

MATH

Recognition of Sign Language (Computer Vision & Classification)
Convolutional neural networks (CNNs) are used in this module to extract features from input video frames or images.

Convolution Process:

$(I * K)(x, y) = \sum i = 0\ m$
$\sum j = 0 n I\ (x + i, y + j) \cdot K\ (i, j)$ (I∗K)(x,y)= i=0 ∑ m

I(x+i,y+j)·K(i,j) where j=0 ∑n

The input image is I I.

The kernel/filter is denoted by K.

The spatial coordinates (x,y) are (x,y).

Activation Function (often ReLU):

$f(x) = \max (0, x)$
f(x)=max(0,x)
Softmax for Classifying Gestures in Multiple Classes:

$P(y = j \mid x) = ezj$
$\sum k = 1\ Kezk$
P(y=j∣x)=∑ k=1 K e z k

e z j

where K is the total number of gesture classes and zjz j is the output score for class j j.

III. REVIEW OF LITERATURE

From sensor-based approaches to AI-powered solutions, research on bridging communication between hearing and deaf people has advanced. While contemporary methods use CNNs and computer vision for non-intrusive sign language recognition, early systems, such as Kadous (2002), used instrumented gloves (Koller et al., 2015; Camgoz et al., 2018). Accurate speech-to-text conversion is now possible thanks to deep learning models like Wav2Vec 2.0 (Baevski et al., 2020) and Deep Speech (Hannun et al., 2014). Real-time, bidirectional communication is the goal of hybrid systems like Sign2Speech (Chai et al., 2020) and mobile tools (Abushaira & Doush, 2021). Accuracy, latency, and multilingual support issues still exist, though, underscoring the need for more flexible, inclusive, and intelligent communication systems like the one that has been suggested.

IV. METHODOLOGY

The system uses a camera module to record sign gestures and a microphone to record voice input. It makes use of speech recognition APIs for voice interpretation and machine learning for sign classification.

These inputs are assessed by the core comparison engine, which then provides an intelligible output format .

By combining AI-based interpretation models with visual and audio input processing, the suggested system enables real-time communication between hearing and deaf people. The following fundamental modules comprise the methodology's structure:

Information Gathering
Visual Input: The user's sign language gestures are recorded by a camera module. Audio Input: The hearing user's spoken words are captured by a microphone.

Preprocessing Image Frames:
To recognise gestures, video input is divided into separate frames.

Speech Signals: Google Speech-to-Text and other speech recognition APIs use noise reduction to process audio signals and turn them into text.

Feature Extraction Sign Gestures:
Computer vision techniques (e.g., OpenPose, Media Pipe) are used to extract motion vectors and key points from video frames.

Speech input: contextual embeddings or Mel-Frequency Cepstral Coefficients (MFCCs) are taken out for natural

V. SYSTEM IMPLEMENTATIONS

Python is the main programming language used to implement the system because of its extensive library and AI support. OpenCV is used to process video input and identify hand movements in real time for gesture capture. Voice input is handled by the Google Speech API, which effectively translates spoken words into text. Accurate classification is made possible by using TensorFlow to create and train deep learning models on a labelled dataset of sign gestures. Signs are interpreted by the trained models, which then translate them into words or phrases. The program is appropriate for desktop use since it uses Tkinter to create an easy-to-use graphical user interface. A responsive, real-time communication system between hearing and deaf people is ensured by this integrated approach.

VI. RESULTS AND DISCUSSION

A set of fundamental sign language gestures and spoken inputs were used to test the system's efficacy in real-time communication. The trained TensorFlow model's dependability was demonstrated by the gesture recognition accuracy rate of over 85%. In quiet settings, voice recognition using the Google Speech API demonstrated excellent precision, accurately translating spoken words into text. The smooth bidirectional communication made possible by the combination of these two modes allowed hearing and deaf users to understand one another. Sign language was successfully converted into speech and readable text by the system, and vice versa. Overall results support the system's potential as a useful assistive communication tool, even though performance somewhat decreased in noisy environments. Robustness in a variety of environments could be the focus of future developments.

VII. PROJECT METHODS

For a long time, the communication gap between hearing and deaf people has presented social and educational difficulties. Assistive technologies designed to close this gap have been made possible by recent developments in artificial intelligence (AI) and machine learning. Camera-based, vision-driven methods have replaced hardware-dependent solutions like sensor gloves in sign language recognition systems. The accuracy and resilience of gesture recognition have been greatly enhanced by deep learning methods, especially convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

Concurrently, end-to-end neural network models such as Deep Speech and Transformer-based architectures have advanced speech recognition technologies, allowing for more accurate and natural voice-to-text transcription. Real-time, bidirectional communication platforms that can translate sign language into speech and vice versa are made possible by combining these technologies.

But there are still issues like different sign language dialects, different signer styles, interference from background noise, and computational latency. In order to promote inclusive communication, this project combines cutting-edge AI models with easily accessible hardware and intuitive user interfaces.

tamper evidence in a trustworthy, query-friendly environment, it preserves system integrity.

One of the biggest social challenges still facing society is effective communication between hearing and deaf people. Conventional approaches depend on human interpreters, who can be expensive and unavailable at times. AI-powered systems that make use of computer vision and natural language processing (NLP) have been created in order to get past these obstacles.

Mechanical gloves and sensor-based devices have given way to camera-driven, vision-based systems that use deep learning for sign language recognition (SLR). While temporal models like Transformers or LSTMs capture motion dynamics, CNNs are frequently used to extract spatial features from hand gestures. Although there are still issues with signer variability and large annotated datasets, works like Koller et al. (2015) and Camgoz et al. (2018) show notable advancements in continuous and isolated sign language recognition.

Models like Deep Speech and Wav2Vec 2.0, which employ self-supervised learning to increase transcription accuracy even in noisy environments, have revolutionised speech recognition. Bidirectional communication platforms are made possible by integrating speech-to-text and sign-to-text technologies, which are essential for inclusive discourse.

Current systems are limited by problems such as ambient noise robustness, multi-dialect support, and latency, despite advancements. Emerging trends that seek to improve accessibility and user convenience include wearable technology and smartphone apps.

This project expands on these advancements by integrating high-accuracy speech transcription and dependable gesture recognition into an intuitive user interface. The goal is to enable seamless, real-time communication between the hearing and deaf communities.

REFERENCES

[1] Li, D., Yu, X., Opazo, C. R., & Li, H. (2019). A Comparison of New Large-scale Datasets and Methods for Word-level Deep Sign Language Recognition from Video. presents the WLASL dataset, which makes it easier to use deep learning models for large-scale word-level recognition of American Sign Language.

[2] Macêdo, D., Correia de Amorim, C., and Zanchettin, C.(2019). Convolutional Networks for Spatial-Temporal Graphs in Sign Language Recognition. suggests using a graph convolutional network technique to recognise spatial and temporal dependencies in sign language.

[3] Rokad, B., Bidani, S., Dennisan, A., & Kasukurthi, N. (2019). Deep Learning-Based Alphabet Recognition in American Sign Language. uses deep learning techniques to create a model for ASL alphabet sign recognition. Santhalingam, P. S., Hosain, A. A., Pathak, P., Kosecka, J., & Rangwala, H. (2019). Analysis of Multimodal Data for Sign Language Recognition. investigates the application of deep learning architectures for sign language recognition using skeletal and RGB video data.

[4] P.Gupta(2023). An AI model for real-time sign language translation is created by an Indian engineering student. explains the creation of an AI model that can translate words from American Sign Language into English.

[5] Waghmare, P. P., Dubewar, S., Deshpande, A. M., & Dhaybar,T.(2024).A Deep Learning Method for a Model that Combines Video Generation and Indian Sign Language Recognition. suggests a system that uses deep learning models to generate videos and recognise ISL.

[6] Tarafder, T., Aziz, R., Arora, M., Vashisth, H. K., & Alpana. (2023). Deep Learning-Based Hand Gesture Recognition in Indian Sign Language. uses deep learning techniques to investigate the recognition of hand gestures in Indian Sign Language.

[7] Xu, T., Brockman, G., McLeavey, C., Kim, J. W., and Radford, A. (2022). Whisper: Sturdy Speech Recognition with Wide-Scale Weak Supervision. presents Whisper, a machine learning model for transcription and speech recognition.

[8] D. Povey and associates (2020). Kaldi: A Speech Recognition Toolkit. offers a speech recognition toolkit that is open source for creating recognition systems.

[9] Bhashini (2024). Bhashini: An Instantaneous Indian Language Translation Tool. talks about Bhashini, a platform that improves digital inclusion by providing real-time translation between Indian languages.

WEBSITES REFERRED

[1] TensorFlow Documentation: https://www.tensorflow.org/

[2] Flask Documentation: https://flask.palletsprojects.com/en/2.0.x/

[3] MySQL Documentation: https://dev.mysql.com/doc/

[4] OpenCV Documentation: https://opencv.org/

[5] Bootstrap Documentation: https://getbootstrap.com/docs/5.1/getting-started/introduction/

[6] Scikit-learn Documentation: https://scikit-learn.org/stable/documentation.html

[7] Pandas Documentation: https://pandas.pydata.org/docs/

[8] NumPy Documentation: https://numpy.org/doc/

[9] Matplotlib Documentation: https://matplotlib.org/stable/contents.html

[10] Seaborn Documentation: https://seaborn.pydata.org/tutorial.html

[11] WampServer Documentation: http://www.wampserver.com/en/documentation.php