www.ijarp.com ISSN 2456-9992 Page: 01-05

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 02

AI SIGN LANGUAGE TO TEXT AND SPEECH CONVERTOR

***Aarushi Yadav*[*1] *Anu Yadav*[2]*, Ms. Pushpanjali*[3]

[1,2]UG Student of Department of Bachelor of Computer Application, Shri Ramswaroop Memorial College of Management, Lucknow, Uttar Pradesh, India.

[3]Assistance Professor, Department of Bachelor of Computer Application, Shri Ramswaroop Memorial College of Management Lucknow, Uttar Pradesh, India.

Article Received: 15 January 2026, Article Revised: 03 February 2026, Published on: 23 February 2026

***Corresponding Author: Aarushi Yadav**

UG Student of Department of Bachelor of Computer Application, Shri Ramswaroop Memorial College of Management, Lucknow, Uttar Pradesh, India.

DOI: https://doi-doi.org/101555/ijarp.9109

ABSTRACT:

People with hearing and speech disabilities mainly depend on sign language for communication. However, most individuals are not familiar with sign language, which often creates communication problems in daily life. In many situations, interpreters are required, but they may not always be available. This research work presents an AI-based system that converts sign language gestures into text and speech. The system captures hand gestures using a camera and processes them with computer vision and deep learning techniques. Recognized gestures are displayed as text and also converted into speech output. The system works in real time and helps reduce communication gaps, especially in educational and social environments.

KEYWORDS: Sign Language Recognition, Gesture Detection, Computer Vision, Deep Learning, Text-to-Speech.

I.INTRODUCTION:

Sign language is commonly used by people who are unable to hear or speak. Despite this, a large part of the population does not understand sign language, which makes communication difficult in many everyday situations. In places such as schools, hospitals, and public offices, this problem becomes more serious and often leads to dependence on interpreters. While interpreters are helpful, they are not always present, and continuous assistance is not practical.

With the development of artificial intelligence, it has become possible to process images and videos efficiently. Computer vision techniques allow systems to understand hand movements captured through cameras. Using these technologies, sign language gestures can be recognized automatically. This paper focuses on designing a system that converts sign language gestures into text and speech, making communication easier for hearing- and speech-impaired individuals.

II. BACKGROUND AND RELATED WORK

A. Problems with Traditional Interpretation

Traditional sign language interpretation depends on trained professionals. This approach requires continuous human effort and is not suitable for all situations. In many real-life cases, interpreters are unavailable, which limits effective communication.

B. Gesture Recognition Studies

Gesture recognition has been studied for many years in computer vision. Earlier methods used manually designed features, but recent systems rely on deep learning models that learn gesture patterns directly from data.

C. AI in Assistive Applications

Artificial intelligence is widely used in assistive technologies. AI-based systems help individuals with disabilities perform tasks independently and reduce the need for constant human support.

III. OBJECTIVES

The main objective of this project is to develop a system that converts sign language gestures into text and speech. The specific objectives are:

1. To capture hand gestures using a camera
2. To recognize sign language gestures in real time
3. To convert detected gestures into text
4. To generate speech output from the text
5. To provide a simple communication solution

IV. METHODOLOGY

A. System overview

The proposed system works by capturing live video and processing it frame by frame. Each frame is analyzed to detect hand gestures, which are then converted into text and speech.

B. Video Capture

A webcam is used to capture real-time video. OpenCV is used to read video frames and perform basic preprocessing operations.

C. Gesture Detection and Recognition

A deep learning model is applied to recognize hand gestures from the video frames. The model predicts the corresponding sign based on learned gesture patterns.

D. Feature Analysis

Hand-related features such as movement, position, and shape are analyzed to improve recognition accuracy.

E. Text and Speech Conversion

After a gesture is recognized, it is mapped to predefined text. The text is then converted into speech using a text-to-speech module so that it can be heard by others.

TECHNOLOGIES USED

Programming Language

- Python

Libraries and Tools

- OpenCV for image and video processing
- TensorFlow or Py Torch for gesture recognition
- NumPy for numerical calculations
- Text-to-Speech library for audio output

SYSTEM FEATURES

- Real-time recognition of sign language gestures
- Conversion of gestures into readable text
- Speech output for better communication
- Easy-to-use interface

- Works well in indoor environments

RESULTS AND ANALYSIS

A. Testing Environment

The system was tested using a standard webcam in indoor conditions. Different scenarios such as movement, multiple students, and varying lighting were considered.

B. Performance Analysis

The system successfully detected students, identity cards, and mobile phones. Pose-based uniform detection improved accuracy compared to basic bounding box methods. Alerts were generated immediately when violations were detected.

C. Observations

- The need for human interpreters was reduced
- The response time of the system was fast
- Minor errors occurred when gestures were performed very quickly

FUTURE SCOPE

The system can be improved in the future by adding support for full sentence recognition, including facial expression analysis, extending support to more sign languages, deploying the system on mobile devices, and training the model with larger datasets.

CONCLUSION:

This paper presented an AI-based sign language to text and speech conversion system. The system uses computer vision and deep learning techniques to recognize hand gestures in real time. By converting gestures into text and speech, the proposed system helps reduce communication barriers for hearing- and speech- impaired individuals. The results show that the system performs well in controlled environments and can be further enhanced for real-world use.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the institution and faculty members for their guidance and support. Special thanks are given to the project guide for providing helpful suggestions during the completion of this research work.

REFRENCES

1. T. Starner and A. Pentland, “Real-Time American Sign Language Recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997.
2. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2023.
3. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
4. G. Bradski, “The OpenCV Library,” Dr. Dobb’s Journal of Software Tools, 2000.