

Sign_Trans – Sign Language Translator in Text or Speech

Manasi Mundhe¹ Dhanashree Sandeep Mukhekar² Yogesh Suresh Mohite³

Department of Computer Applications

JSPM University, Pune

¹ manasimundhe4@gmail.com ² dhanashreemukhekar@gmail.com ³ yogeshmohite1725@gmail.com

Abstract: *People with hearing or speaking disabilities often have difficulty communicating with someone who does not understand Sign Language. Many individuals in India are using Sign Language daily. Currently there are no systems that have been implemented that will transform signed gestures to a format that others can read or hear. In this paper, we present a real-time Sign Language translator utilizing computer vision and deep learning techniques.*

This system employs a webcam to capture the user's signed gesture, applies MediaPipe to concurrently identify and track the hand (i.e. tracking of the hand position) and utilize trained machine learning models on the captured signed gesture to determine the interpretation of the signed gesture. Once the interpretation of the signed gesture is known, the signed gesture is converted to a textual occurrence and a speech occurrence (utilizing text-to-speech) will be produced concurrently for the end-user.

The model does not require any additional hardware (such as a data glove) to utilize the functionality of this model; thus, reducing the cost to produce the hardware and the ease of use of the mid-end to the end-user. The focus of the project is to provide means for interpreting Sign Language gestures using multiple devices to aid communication for the deaf/hard-of-hearing community and the hearing community.

Our experimental results indicate that the developed system functions approximately as accurate as expected to translate signed gestures. By providing an opportunity for individuals who use Sign Language to easily communicate with those who do not or do not have access to sign language, we can improve the accessibility of the population as a whole and improve the quality of life for individuals who are unable to hear and/or speak.

Keywords: : Indian Sign Language, Gesture Recognition, MediaPipe, Deep Learning, Text-to-Speech, Computer Vision.

I. INTRODUCTION

We all communicate with each other every day. For those who are deaf or cannot speak, communicating with those who do not know sign language can sometimes be difficult. Sign Language is a form of communication used by many deaf individuals in India, but many people are unaware of its existence and do not have the technology available to them to assist. Schools, hospitals, the workplace and other public places make it very difficult for people to communicate due to a lack of available interpreters. Most methods for translating sign language rely on complex and expensive systems that require a significant number of pieces of equipment, and most of them were designed specifically for translating foreign, such as American Sign Language; therefore, we need a new system designed specifically for Indian Sign Language.

This project's goal is to address this issue by creating a real-time Indian Sign Language translation system using both Artificial Intelligence and Computer Vision. The system will capture the user's hand gestures through a webcam and will use MediaPipe and Deep Learning algorithms to help determine what the gesture is meant to mean. Once the gesture is identified, it will then be converted into text and/or audio so that individuals who do not understand sign language can read or hear the message.



The ultimate goal of this initiative is to give people who have a hearing disability access to the tools necessary to communicate effectively in both verbal and written form with their peers and family members.

II. LITERATURE REVIEW

Machine learning and deep learning have allowed a number of researchers to create systems for gesture recognition. Some of them use convolutional neural networks (CNNs) for the detection of stationary hand gestures while others used recurrent neural networks (RNNs) with long-short term memory (LSTM) models for gesture recognition of mobile hand gestures. Recently, the discovery of MediaPipe for fast and accurate hand shape detection has enabled the temporal tracking of gestures. Gesture recognition in conjunction with text-to-speech technology is also being utilized by some products to enhance communication for the user.

Although positive strides have been made in developing gesture recognition systems, the vast majority of such systems only provide support for ASL. Compared to ASL, there are very few gesture recognition systems that have been developed to support Indian Sign Language, as well as the majority of gesture recognition systems relying heavily on the use of gloves or costly devices to limit the potential for most aspiring users.

These provide us with insight into gesture recognition technologies and ultimately the experience and expertise to develop our own gesture recognition system. Also, the learning of gesture recognition techniques will give us the knowledge base to build an enhanced gesture recognition system.

III. PROPOSED METHODOLOGY

The system described operates in several phases. The first phase captures a video from a webcam of the user's hand gestures, and the second phase processes this video through a series of actions to produce a clean version of the video without any extraneous information.

Next, the MediaPipe solution will identify the keypoints of the hand in the video and provide the position coordinates for these keypoints to a deep learning model. The deep learning model will determine what gesture is being performed and will return the identified gesture to the system.

After the gesture is identified, the gesture will be displayed as text on the user's display. This text will then be converted to sound using the speech synthesis engine (Text-To-Speech). The system is using the Text-To-Speech engine to produce audible sound from the text for the user to hear.



IV. SYSTEM IMPLEMENTATION

The proposed system has incorporated several technologies in its development: Python, OpenCV, MediaPipe, TensorFlow, and Keras. The OpenCV library provided access to the user's webcam video feed; MediaPipe enabled hand landmark localization through image processing techniques. For creating the hand gesture recognition model, TensorFlow and Keras were utilized. The proposed system functions by continually capturing the user's webcam feed, using image processing techniques to find the location of the hand landmarks and extract information as to how the hands move to represent gestures. The representation of the gesture will then be used to identify what gesture was made. The corresponding text to the identified gesture will be shown via the user interface of the proposed system, and a text-to-speech engine will produce an audio representation of the identified gesture.

The user-friendliness of the proposed system minimizes the need to have previous technological experience as an end-user. The proposed system interface includes three main parts; the webcam live feed, area for displaying identified text,



and an area for producing audio of the identified text. This simple structure fosters ease of interaction by the end user and makes the proposed system suitable for use for interaction.

Technology	Purpose
Python	Core Programming
OpenCV	Image Processing
MediaPipe	Hand Tracking
TensorFlow/Keras	Deep Learning
gTTS	Speech Generation
Flask/Streamlit	Web Interface

V. RESULTS AND DISCUSSION

The system test of the developed application utilized various gestures from the Indian Sign Language, performed under normal conditions of the environment, demonstrated its ability to recognize the user's hand gesture input and convert it into on-screen text and/or audio output, in real time. The observed response time of the application for the identified hand gesture inputs was suitable for basic communication.

MediaPipe greatly enhanced the user's ability to track the hand gestures while also minimizing extraneous background interference when compared to traditional hand segmentation methods. As a result it was able to provide greater speed and more stable performance than the previously identified hand segmentation algorithms.

The text output produced by the application was legible, as was the generated audio output via the Text-to-Speech engine. The application was able to perform adequately when sufficient lighting was available and the user's hands could be seen. In contrast, the accuracy of the application was somewhat diminished when there was insufficient lighting and/or when the background density contained numerous moving objects.

Observed Results

Parameter	Observation
Gesture Recognition	Working properly
Detection Speed	Fast
Text Conversion	Accurate
Speech Output	Clear
Cost Requirement	Affordable

VI. ADVANTAGES AND LIMITATIONS

Advantages

- The system works in real time and provides quick gesture recognition.
- No external devices such as sensor gloves are required.
- The application is affordable because it uses a standard webcam.
- Both text and speech outputs are generated for easier communication.
- The interface is simple and easy to use.
- The project specifically focuses on Indian Sign Language.



Limitations

- Performance may reduce in low lighting conditions.
- Complex backgrounds can affect hand detection accuracy.
- Similar hand gestures may sometimes produce incorrect predictions.
- The current system supports limited ISL gestures only.
- Dynamic sentence-level translation still requires improvement.

VII. CONCLUSION AND FUTURE SCOPE

The Indian Sign Language to Text/Speech Translator we created is very useful. It helps people who cannot hear and the rest of us communicate with each other easily. The Indian Sign Language to Text/Speech Translator uses a computer to recognize hand gestures and convert them into words and speech that we can understand. This process happens quickly. We discovered that by using tools like MediaPipe and OpenCV and TensorFlow we can build a system that facilitates communication between people. The Indian Sign Language to Text/Speech Translator is affordable and easy to use. We do not need any machines for it to work. This makes the Indian Sign Language to Text/Speech Translator easy to use every day.

The Indian Sign Language to Text/Speech Translator can be used in places such as schools and hospitals and offices where people need assistance with communication.

In the future we can improve the Sign Language to Text/Speech Translator. We can add Indian Sign Language gestures to the system so it can understand them better. We can also make the Indian Sign Language to Text/Speech Translator speak in languages. We can put the Sign Language to Text/Speech Translator on a phone so people can use it anywhere they go. We can also add features like recognizing expressions and responding to people. This will make the Indian Sign Language, to Text/Speech Translator more effective and useful.

