A SYNOPSIS ON

Bidirectional Sign Language Translator

Submitted in partial fulfilment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by:

Vaibhavi Ghildiyal	**2021514**
Akhil Kumar	**2021066**
Ansh Mourya	**2021105**

Under the Guidance of
Dr. Ashish Garg
***Assistant* Professor**

Project Team ID: MP2025CSE27

Department of Computer Science and Engineering
Graphic Era (Deemed to be University)
Dehradun, Uttarakhand

September-2025

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the Synopsis entitled **"Bidirectional Sign Language Translator "** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering in the Department of Computer Science and Engineering of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the undersigned under the supervision of **Dr. Ashish Garg**, **Assistant Professor** Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Vaibhavi Ghildiyal	2021514
Akhil Kumar	2021066
Ansh Mourya	2021105

The above mentioned students shall be working under the supervision of the undersigned on the **"Bidirectional Sign Language Translator"**

Signature
Supervisor

Signature
Head of the Department

Internal Evaluation (By DPRC Committee)

Status of the Synopsis: Accepted / Rejected

Any Comments:

Name of the Committee Members: **Signature with Date**

1.

2.

Table of Contents

Chapter No.	Description	Page No.
Chapter 1	Introduction and Problem Statement	**4-5**
Chapter 2	Background/ Literature Survey	**6-7**
Chapter 3	Objectives	**8-9**
Chapter 4	Possible Approach/ Algorithms	**10-12**
	References	**13**

Chapter 1

Introduction and Problem Statement

1.1 Introduction

Communication is essential for human interaction, but people who are deaf or have speech difficulties often struggle because many hearing people do not understand sign language. Sign language uses hand gestures, body movements, and facial expressions, and is a main way for deaf people to communicate with others. Since most people do not know sign language, deaf individuals often find it difficult to communicate smoothly in schools, hospitals, workplaces, and daily life.

Our project aims to bridge the communication gap between hearing and deaf individuals by developing a two-way translation system. The solution offers two main features:

1. Sign Language to Text Conversion: The user records a video of someone using sign language and uploads it to the system. Our system recognizes the signs and translates them into written sentences. This helps users easily understand what the deaf person wants to communicate.
2. Speech to Sign Language Conversion: The second feature lets the user speak into a microphone. The system identifies the important words from the speech, matches them with the right sign language gestures, and shows pre-recorded videos. This helps deaf individuals understand what the hearing person is saying.

The system is easy to use, making it possible for both deaf and hearing people to communicate as smoothly as two hearing individuals. It can be used in education, healthcare, customer service, and daily life.

1.2 Problem Statement

Sign language is the main channel of communication for the deaf people, it is not commonly used by the general public. As a result of this communication gap between hearing people and deaf people, create obstacles in education, healthcare, workplace and daily life.

Although there are some systems that address this issue, but they are unidirectional in one area only dealing with speech/text- to-sign conversion or sign-to-text translation. These one-way or partial solutions are unable to deliver the two-way communication.

Thus there is a need for a bidirectional translation system that enable the hearing people to communicate using speech into sign language and enable deaf person to do the sign gestures translated it to text. Solving this problem will guarantee meaningful interaction, social integration of the deaf people.

Chapter 2

Background/ Literature Survey

2.1 Background

Communication is essential for human interaction, but people who are deaf or have speech difficulties often struggle because many hearing people do not understand sign language. Sign language uses hand gestures, body movements, and facial expressions, and is a main way for deaf people to communicate with others. Since most people do not know sign language, deaf individuals often find it difficult to communicate smoothly in schools, hospitals, workplaces, and daily life.

Traditional technique to reduce this communication gap are human interpreters, subtitles. Though these are not always feasible. But with the rapid growth in Artificial Intelligence, Computer Vision and NLP, it is possible to develop a system can real-time translate the sign language to text and speech to sign language, allowing both deaf and hearing person to interact easily as a normal conversation.

This project is based on this idea by developing a bidirectional translation system:

1. Sign Language to Text Conversion- Use the record video and translate it into readable text using the computer vision and machine learning model.
2. Speech to Sign Language Conversion- Convert the speech into the text using NLP and use that text to exact the word and display the pre-recorded video of the sign Language based on the word exact from the text.

This dual functionality ensures ease of communication in real life situation.

2.1 Literatue Survey

Dongxu Li [1] introduced the World-Level American Sign Language(WLASL) dataset, consists of American Sign Language video dataset, consists of over 21,083 videos performed by 119 signers, and covers more than 2000 words. Their study show that larger datasets and advanced models significantly improve the sign recognition accuracy.

Mishra [2]developed an Audio-to-Indian Sign Language(ISL) translation system that convert the spoken or written input into sign language, uses NLP from text preprocessing and pre-recorded ISL video as output. Their work covered the shortage of interpreters and ISL tools in India and showed how a single direction of translation could be achieved from audio and text input to sign language output. The system did not work in the reverse direction.

Previous studies utilized Hidden Markov Models (HMMs), 3D Convolutional Neural Networks (CNN) for sign recognition. Although the approaches had potential, they tended to be based on small datasets and were not robust under real-world environments [1]. The advancement of deep learning models like CNNs, LSTMs, Transformers, and Graph Neural Networks (GNNs) made more dependable recognition and translation systems possible

Despite these advancements, most existing system are unidirectional(speech-to-sign or sign-to-text). This create a clear gap for a bidirectional, real-time system that will combine speech recognition, NLP and computer vision and provide both things in one system.

Chapter 3

Objectives

3.1 Functional goals:

1. Real time sign recognition: Software will recognize sign movements in real time using record video and translate them into text.

2. Speech to sign generation: The app will listen to the user voice through microphone and then convert the voice to the text and system identifies the important words from the speech and matches them with the right sign language gestures, and show pre-recorded videos according to the words.so that the deaf person can understand the sign language better.

3. Bii Directional communication: Enables two way communication between disabled person and a normal person.

4. Error correction and feedback: The app provides visual feedback during study capture. That is, in case a sign language is a not properly captured by the camera. The app will show error and give the feedback that the sign has not been captured properly so that the disabled person can repeat the gesture.

3.2 Advanced analytical objectives:

1. Translation accuracy: The app aims to provide a target accuracy of around 75%. We'll employ natural language processing technique to ensure that the converted text is accurate and grammatically correct.

2. Continuous learning and model optimization: The app will securely collect the user data for more training and improving the model.

3. Environmental Robustness: The app will be trained to give accurate translation during environmental factors like varying lightning conditions or in different backgrounds or in case of minor occlusions.

3.1 Usability and Scalability objectives:

1. Low latency: The training of the model will ensure that the latency for both STT and STS conversion is minimal. That is some less than 1000 milliseconds to enable a fluid and conversational experience.
2. Interactive GUI: Simple and accessible interface that is easy to use.

Chapter 4

Possible Approach/ Algorithms

For this project, we are planning to build a real-time, two-way communication app to help people using sign language and spoken language understand each other better. The app will have two main parts working together:

- A system that recognizes signs and converts them into text.
- Another system that takes spoken words and shows the corresponding signs visually.

4.1 Sign-to-Text Recognition

The core of our project is the sign recognition module. We decided to go with a **pose-based approach** because it is robust to variations in lighting, background, and signer appearance, which is crucial for a real-world application.

4.1.1 Vocabulary and Training Data : To keep the project manageable, we'll start with a focused set of 100–200 common ASL words. For training, we will use the WLASL dataset, which we chose because its diversity of signers is a key advantage for training a more generalizable model.

4.1.2 Feature Extraction and Pre-processing : Our plan is to convert each video into a sequence of numerical data. We will use mediapipe to extract the (x,y,z) coordinates of hand and body for each video.

4.1.3 The Prediction Model Architecture : Once we have the pre-processed landmark data, we need a model to classify it. We will build and compare two neural network architectures using PyTorch or TensorFlow.

4.1.4 Training and Evaluation: We will split the WLASL data into training, validation and testing sets. The model performance will be test on the unseen data and measure the accuracy, precision of the model.

4.1.5 Generating Text

Finally, when our model recognizes a sign with enough confidence, the app will show the text as output to the user.

4.2 Speech-to-Sign Translation

The other part of the app converts spoken words into signs. We're focusing on making this simple and understandable.

4.2.1 Capturing Speech

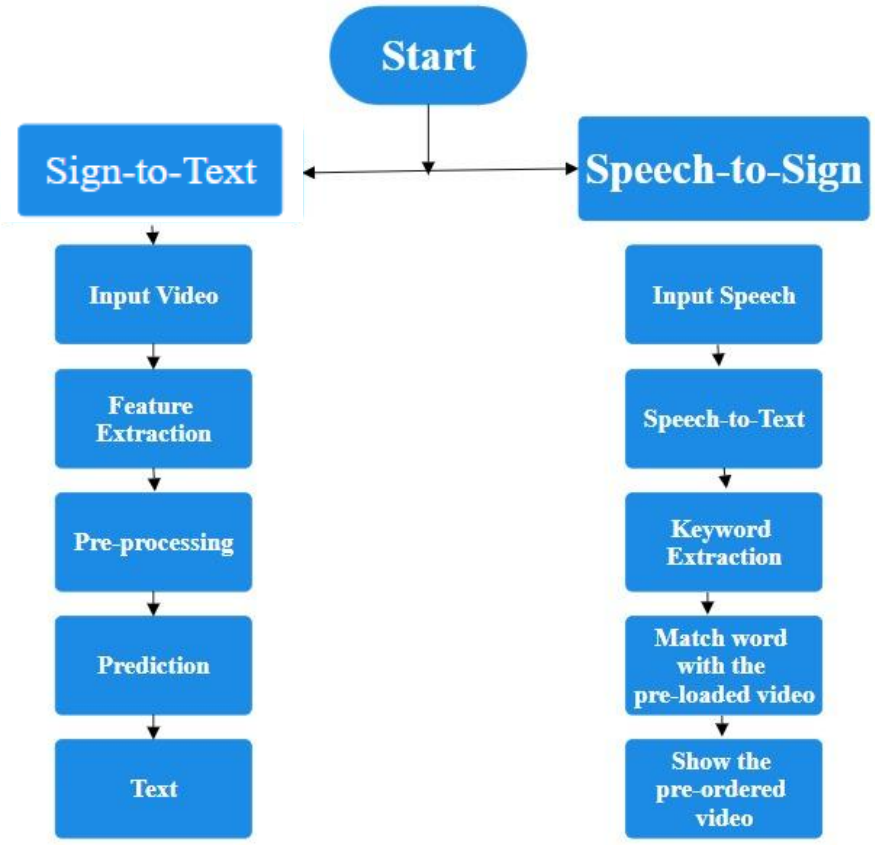
We will use a standard speech-to-text API (like the ones built into modern browsers) to turn what the person says into text.

4.2.2 Finding Keywords

Then, we will look for keywords in the text that match our vocabulary of 100–200 signs. For example, if someone says, "Can I have some water, please?" we'll just pick out "water". We are not trying to translate full sentences perfectly.

4.2.3 Showing the Sign

For showing signs, we will use pre-recorded videos. Each word in our vocabulary will have a short video clip. When a keyword is detected, the app will play the corresponding video.



4.3 Putting It All Together

Both parts will be integrated into a single app. The interface will switch smoothly between the camera for signing and the text/video display for speech. The goal is to make the conversation as natural and easy as possible.

References

[1] Li, D., Rodriguez, C., Yu, X., & Li, H. (2020). Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV).

[2] M. Mishra, R. Raut, and A. Wanjari, *"Audio to Indian Sign Language Translation,"* International Journal of Research in Engineering, Science and Management, vol. 5, no. 7, pp. 66–69, 2022.

[3] Vaswani, A., et al. (2017). Attention Is All You Need. In Advances in Neural Information Processing Systems (NeurIPS), 30.