TALKING FINGERS

A PROJECT REPORT

Submitted by,

Sudiksha N -20211IST0016

Muskan Ali -20201IST0050

DP Rakshitha -20211IST0007

Vijay Vardhan M -20211IST0019

Under the guidance of,

Ms. Pushpalatha M

Assistant Professor, School of Computer Science and Engineering

In partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

INFORMATION SCIENCE AND TECHNOLOGY

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2025

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

CERTIFICATE

This is to certify that the Project report "TALKING FINGERS" being submitted by "Sudiksha.N, Muskan Ali, DP Rakshita, Vijay Vardhan" bearing roll number(s) "20211IST0016, 20201IST0050, 20211IST0007, 20211IST0019" in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Information Science and Technology is a bonafide work carried out under my supervision.

Pallavi R 21/01/25

Ms. Pushpalatha
ASSISTANT PROFESSOR
School of CSE
Presidency University

Dr. PALLAVI R
PROFESSOR & HoD
School of CSE
Presidency University

Dr. L. SHAKKEERA
Associate Dean
School of CSE
Presidency University

Dr. MYDHILI K NAIR
Associate Dean
School of CSE
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-VC School of Engineering
Dean -School of CSE&IS
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **TALKING FINGERS** in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Information Science and Technology**, is a record of our own investigations carried under the guidance of **Ms. Pushpalatha, ASSISTANT PROFESSOR, School of Computer Science Engineering, Presidency University, Bengaluru.**

Name(s)	**Roll No(s)**	**Signature(s)**
Sudiksha N	20221IST0016	Sudiksha. N
Muskan Ali	20201IST0050	Muskan Ali
DP Rakshita	20211IST0007	Rakshita
Vijay Vardhan M	20211IST0019	Vijay

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

ABSTRACT

Talking Fingers: Bridging Communication for the Deaf and Hard-of-Hearing.

"Talking Fingers" is an innovative project that focuses on improving the interaction between the hearing and non-hearing by designing a web-based system to translate spoken language into Indian Sign Language (ISL). The deaf and hard-of-hearing face difficulties with communication because there is a lack of interpreters and accessible ISL resources. The current tools mainly concentrate on the conversion of text into sign language and are largely designed for American Sign Language (ASL) or British Sign Language (BSL), hence leaving ISL poorly represented. Most of the solutions currently in place are only word-to-word translations, which lack contextual and grammatical correctness.

The system proposed here removes these limitations through the use of advanced technologies for meaningful and accurate translation. Using the Google Speech API, the system captures speech input and converts it into text. This text undergoes processing through sophisticated Natural Language Processing (NLP) methods to identify meaningful phrases and maintain contextual accuracy. The segmented text is then matched to ISL visual representations, including videos or GIFs stored in a comprehensive database. The system acts to fall back when there is no proper ISL visual available for a word, spelling out the word letter by letter in ISL to avoid communication gaps.

This system is designed to make communication accessible and user-friendly, with applications in public spaces such as hospitals, banks, and railway stations. The system focuses on ISL, addressing a significant need for inclusivity in India, where millions rely on this sign language for communication.

Globally, sign language is an essential medium for over 135 recognized languages, with ISL being a vital tool for communication in India. However, the deaf community has faced several barriers: lack of interpreters, poor awareness, and minimal technological support. Existing systems are usually centered around ASL or BSL and do not provide real-time translations or take into account the grammatical complexities of sign languages. "Talking Fingers" is different as it provides a solution specifically for the needs of ISL users.

The integration of speech recognition, NLP, and ISL visuals would guarantee the precise and grammatically correct translation of the whole sentences. Unlike methods used in olden times with just word-by-word conversion, this project helps understand the content much better in communication. Using the Google Speech API, it captures speech input and processes into

text, where it is broken down into meaningful units through the application of NLP techniques. These units are connected to a database of ISL signs for videos or GIFs that can be associated with the intended message. When word signs are not available, the fall-back system writes words in ISL letter by letter for clarity.

The "Talking Fingers" project has three major objectives: to achieve real-time speech-to-text conversion, create a full ISL database, and apply NLP for contextually accurate translations. The platform is designed to be scalable and adaptable, making it suitable for various public and educational spaces. It serves not only as a resource for the deaf community but also for those interested in learning ISL.

This is a move forward for the empowerment of the deaf and hard-of-hearing community by breaking down communication barriers. Combining speech recognition, NLP, and ISL visuals renders "Talking Fingers" a robust and inclusive solution for real-time translation of ISL. Its interest in developing contextual and grammatically accurate marks as different from existing systems and thus illustrates good potential in developing the future modeling of sign language automation.

Keywords: Indian Sign Language (ISL), Natural Language Processing (NLP), Speech-to-Sign Translation, Communication Accessibility, Real-time Translation, Sign Language Automation, Speech Recognition.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC, School of Engineering and Dean, School of Computer Science Engineering & Information Science, Presidency University for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate Deans **Dr. Shakkeera L and Dr. Mydhili K Nair,** School of Computer Science Engineering & Information Science, Presidency University, and **Dr. PALLAVI R**, Head of the Department, School of Computer Science Engineering & Information Science, Presidency University, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide **Ms. Pushpalatha, Assistant Professor** and Reviewer **Dr. A.K Sampath, Professor**, School of Computer Science Engineering & Information Science, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the PIP2001 Capstone Project Coordinators **Dr. Sampath A K, Dr. Abdul Khadar A and Mr. Md Zia Ur Rahman,** department Project Coordinators **Mr. Srinivas Mishra** and Git hub coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

Sudiksha N

Muskan Ali

D P Rakshitha

Vijay Vardhan M

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER-1
INTRODUCTION

1.1 General Overview

Sign language is one of the fundamental modes of expression for the deaf and hard-of-hearing people, through which they can better express their ideas, emotions, and thoughts. Worldwide, there are more than 135 specific sign languages tailored to the distinctive linguistic and cultural requirements of each group. The most important is Indian Sign Language (ISL) for millions in India, leading to emotional, social, and linguistic development. ISL still lacks substantial representation in technological developments, which has created a significant communication gap amongst individuals both privately and publicly.

The scarcity of ISL interpreters and resources has made this communication gap worse. Access to vital services such as healthcare, education, and employment is restricted because of this communication gap. Though some existing systems attempt to solve the sign language translation problem, they are focused on American Sign Language (ASL) and British Sign Language (BSL) and do not provide much support for ISL. Moreover, most of the solutions are based on text inputs and cannot support real-time speech-to-sign translation.

1.1.1 The Importance of Sign Language

Sign language is crucial for communication of millions of deaf and hard-of-hearing people around the globe. It has a full system of hand signs, facial expressions, and body movements, independent of spoken languages. Each country has its unique sign language: ASL in America, BSL in Britain, and ISL in Ireland, which suits the cultural and linguistic requirements of that community.

The importance of sign language is primarily that it transcends communication gap barriers and incorporates people. Thus, its services reach:

Mental Health: Assisting individuals express themselves and consequently promote emotional balance.

Social Connectivity: Allowing one to keep in touch with family, relatives, and their society.

Scholarly Facility: Providing means for learning with equal access.

Cultural Affiliation: Facilitating a community's cultural recognition to preserve traditional values and use of the preferred sign language with their linguistic profile.

In India, ISL is a very important tool for communication because of the scarcity of interpreters and resources. Beyond the deaf community, it fosters collaboration and understanding between hearing and non-hearing individuals. These factors underscore the urgent need for innovative technological solutions to overcome communication barriers.

1.2 The Need for ISL Automation

The Indian deaf community is greatly plagued by many barriers to successful communication. The problems include:

There are few Interpreters: most of whom are not professionally trained in the language, making the real time interaction limited.

Technology is less available: most current applications are focused on ASL and BSL with very minimal alternatives for the ISL.

Contextual Barriers: Current systems concentrate more on translating word-to-sign rather than including grammar and contexts.

Physical obstacles: In accessing necessary services that may be rendered in public sectors like hospitals, banking, and major transportation stations, individuals with hearing may often be disbarred mainly due to the failure to implement Automated ISL system.

The lack of comprehensive solutions tailored to ISL emphasizes the need for a scalable and innovative system. Automating ISL translation has the potential to significantly improve accessibility, enabling real-time communication in various domains and reducing the dependency on human interpreters.

1.3 Project Motivation

The "Talking Fingers" project was the idea for the solution to all these urgent communication challenges. Although current systems already come with text-to-sign translations for ASL and BSL, they cannot meet the demands of the specific needs of users of ISL or even speech-to-sign solutions in real-time. This project aims at the empowerment of the deaf community through offering a fluid communication tool while encouraging the adoption of ISL by hearing individuals. This will promote deaf culture and shed light on the awareness of the language.

The project targets public spaces like hospitals, banks, and railway stations, where real-time communication can have a transformative impact. By leveraging cutting-edge technologies, "Talking Fingers" aims to bridge the communication gap between hearing and non-hearing individuals.

1.3.1 Innovation in "Talking Fingers"

The "Talking Fingers" system has numerous innovative features which distinguish it from other solutions already in existence. These include:

1. **Speech Recognition:** Using the Google Speech API, it converts spoken language into text in real time, thus guaranteeing a fluent user experience.
2. **Natural Language Processing (NLP):** NLP techniques are utilized to improve contextual and grammatical accuracy of translation. Unlike tools, "Talking Fingers" translates entire sentences into ISL, preserving meaning and context.
3. **Extensive ISL Visual Library:** The system uses a vast library of ISL videos and GIFs as visual aids for communication.
4. **Fallback Mechanism:** If the system cannot find a word in the database, it splits the word into letters, allowing for proper spelling in ISL without losing the message's meaning.

These features make "Talking Fingers" a robust, scalable, and adaptable tool that can work across different scenarios.

CHAPTER-2

LITERATURE SURVEY

The advancement of assistive communication systems for deaf or mute people has gained much with the integration of modern technologies like flex sensors, machine learning, and text-to-speech systems. These solutions translate hand gestures into text or speech to improve communication.

1. **Flex Sensors for Gesture Recognition**

Flex sensors are widely used to measure angular bends of fingers. When placed in gloves, they can translate finger motions into digital data for further processing.

A paper presented a glove with flex sensors that could interpret gestures specific to sign languages and achieved remarkable accuracy.

Improved designs have also used inertial sensors to capture dynamic hand movements with greater accuracy.

2. **Combining Sensors for Enhanced Functionality**

Dynamic gestures require hand movements in space and must use a combination of several sensors to detect this motion:

Inertial Sensors: Accelerometers and gyroscopes add spatial orientation and motion tracking that complement flex sensors.

Muscle Activity Sensors: Electromyography or EMG sensors are also used to detect minor finger and hand movements based on the activity of muscles.

It has been evidenced that the recognition accuracy of these systems for both static and dynamic gestures is superior when these sensors are combined together.

3. **Machine Learning for Gesture Classification**

Machine learning techniques greatly strengthen gesture recognition systems by allowing systems to learn their patterns and how users behave and adapt accordingly. Some of these approaches include:

Flex sensors are widely used to measure angular bends of fingers. When placed in gloves, they can translate finger motions into digital data for further processing.

A paper presented a glove with flex sensors that could interpret gestures specific to sign languages and achieved remarkable accuracy.

Improved designs have also used inertial sensors to capture dynamic hand movements with greater accuracy.

Traditional classifiers such as Support Vector Machines (SVMs) and K-Nearest Neighbors (KNNs) for feature-based categorization.

Promising results, often reaching up to more than 90% of the gesture recognition efficiency of some systems.

4. **Text-to-Speech Conversion**

After the correct identification of gestures, it is necessary to convert the data into natural-sounding speech. Modern speech synthesis systems use NLP techniques and advanced models for smooth communication.

Multilingual TTS systems have been helpful, especially for users who need output in different languages.

Real-time systems have shown great promise in providing instant speech output from recognized gestures.

5. **Design and Usability Aspects**

- Ergonomics and user-friendliness are key to the uptake of these assistive technologies.
- Lightweight and portable designs ensure that users can carry and use the devices comfortably.
- Cost-effective solutions increase accessibility, especially in low-income communities.
- Systems that support multiple languages enhance their usability for diverse populations.

6. **Wireless Connectivity and Energy Efficiency**

Contemporary systems typically include modules of wireless communication such as Bluetooth or Wi-Fi, for transmitting gesture data to processing units. The interest is also increasing in energy-efficient designs such as self-powered systems. For example, piezoelectric materials are studied to reduce dependency on external batteries.

7. **Challenges and Future Directions**

The problems still remain there, though so much progress has been made:

The differences in gestures among users can reduce the accuracy of recognition and thus need adaptive algorithms.

Environmental Noise: Background movements or changes in lighting conditions affect sensor reliability.

Scaling systems to accommodate multiple sign languages and varieties is still challenging.

CHAPTER-3
RESEARCH GAPS OF EXISTING METHODS

Corpus Limitations

1. **Size and Diversity Issues:**

   ISL lacks a wide and diversified corpus, which a machine learning algorithm would be unable to make optimal use of. This is contrary to ASL, which can easily benefit from fully annotated comprehensive datasets. Datasets for ISL are fragmented and substantially underrepresented. The existing resources are too narrow and do not have the necessary breadth of signs, phrases, and sentence structures to build rich systems. This renders it challenging for machine learning models to generalize well to real-world applications.

2. **Annotated Data for NLP Models:**

   ISL models rely more on labeled datasets to learn grammatical structures, contextual relationships, and language-specific nuances. However, there is a scarcity of annotated corpora specific to ISL that link the spoken or written text to the corresponding signs. In the absence of such resources, it is difficult to train models to understand the grammar of ISL, which is often subject-object-verb in structure compared to most spoken languages. This lack of such data is a barrier for the models trying to achieve accurate and context-aware translations.

Word-by-Word Translation Challenges

1. Grammatical Differences:

   The grammar of ISL differs significantly from spoken languages like English or Hindi. A direct word-by-word translation can often distort the intended meaning of a sentence, resulting in confusion. For instance, translating "How are you?" directly might result in "How you are?" since ISL typically emphasizes key content words and omits auxiliary verbs and articles.

Integrating NLP can help overcome these grammatical challenges by analyzing sentence structures to produce translations that align with ISL syntax. Key NLP techniques include:

• Tokenization: Breaking sentences into meaningful units, such as dividing “How are you?” into ["How", "are", "you"].
• Parsing: Understanding sentence hierarchy and identifying the roles of individual words.
• Reordering: Rearranging words to reflect ISL grammar, such as converting “How are you?” into “You how?”.
This kind of systemic approach enables systems to produce more accurate, grammatically fluent translations into ISL.

Integration with Public Services

1. Weaknesses in Present System:
There are nearly no ISL translation systems present in vital public areas, such as hospitals, banks, railway stations, and government offices. These areas are extremely sensitive. Any failure in easy and smooth communication may substantially disturb everyday as well as emergency transactions. Many times, unavailability of an accessible ISL system acts as a severe hurdle for the deaf community in performing public services, such as clarifying medical symptoms, conducting banking business, or procuring travel-related information.

2. Proposed Solutions:
The gap can be bridged by developing deployable web-based applications on public devices such as kiosks and information terminals. These could be included within already extant public service infrastructures to be as inclusive as possible. Mobile and desktop applications that can provide live ISL would allow users to access services they need on their own.

Hardware Limitations

1. **Cost and Accessibility:**
   Some of the promising innovations toward real-time gesture recognition include smart gloves and wearable sensors. Unfortunately, these are very expensive gadgets that require a lot of maintenance, which prevents them from being accessible to most people, particularly in developing countries where affordability becomes a key factor.

2. **Dependence on Specialized Devices:**
   Hardware-based solutions are not portable because they limit users to taking the

devices for enabling communication. This model can, therefore, become unpractical in dynamic or fast paced environments.

3. **Software-First Approach:**
   Focusing on the software solutions available through a smartphone, tablet, or computer may alleviate these restrictions. Cloud-based or open-source systems can drastically bring down the cost and enhance access, allowing the ISL tool to reach more people. It will make sure that the advantages of such a system are not limited by financial or logistical bottlenecks in hardware.

Regional Variations

1. **Diversity in ISL:**
   ISL is not uniform throughout India; it varies greatly with regions and cultures. This variation makes it hard to standardize because a single dataset may not represent the entire dialect successfully. For instance, a sign that is widely used in northern India may not exist or hold a different meaning in southern India.

2. **Customizable Systems:**
   To solve this problem systems can be designed in modular fashion so that users can customize the tool as per the local dialect of ISL in use in their area. Crowd-sourced data from different regions would add richness to datasets and result in better representation of India's linguistic diversity, thereby ensuring proper representation of all communities.

Real-Time Capabilities

1. **Lack of Instantaneous Translation:**
   Most of the present ISL systems possess severe delays in processing that render them impossible for being used in dynamic conversations or high-pressure environs like hospitals or transportation hubs. Delay is mainly caused due to an inefficient processing pipeline or hardware.
2. **Improving Real-Time Response:**
   High-speed APIs, such as Google Speech API, can significantly reduce latency by

combining them with optimized algorithms in NLP. The real-time processing of data using cloud-based infrastructure will also facilitate faster and responsive translations.

3. **Practical Applications:**
   Real-time translation will be a basic requirement for emergency communication situations. For instance, in medical emergencies, instant interactions between patients and healthcare providers may mean the difference between life and death. In such hectic environments, real-time systems also enable collaborative work environments, customer service relationships, and normal day-to-day communication.

CHAPTER-4
PROPOSED MOTHODOLOGY

1. Speech Input Capture

The system starts by capturing the user's spoken input using advanced speech recognition tools, such as Google Speech API or any similar libraries, to record the audio so that the captured speech is clear and ready for further processing.

Input Mechanism:

It records speech with the use of microphones or audio-enabled devices. It can also be used from any device, ranging from mobile applications, web browsers, or public kiosks in public. It can be used almost everywhere.

Real-Time Capture:

For effective communication, the system should work in real time. Any delay between speech input and the corresponding Indian Sign Language (ISL) output could hinder effective interaction. The focus is on minimizing latency while maintaining the accuracy of the captured audio.

2. Speech-to-Text Conversion

With the audio recording completed, the translation system transcribes the speech to text. Such a step can be executed using machine learning-based technologies such as the Google Speech API, which works to convert oral words to approximated written versions of the input.

Accuracy:

Precision in transcription enables effective communication with the inclusion and reliability of understanding differences in various accents and speeches, as well as background noises.

Language Support:

For the larger audience, the system should support multi-Indian language speech-to-text conversion. Though the focus remains on ISL, it is very important that the spoken content is translated into text correctly, irrespective of the language.

3. Text Processing Using NLP

After the spoken words are converted to text, Natural Language Processing techniques are used to parse and structure the text. This ensures that the meaning of the sentence and grammar in the source sentence are preserved.

Sentence Segmentation:

The system breaks down the text into parts that are small enough to work with such as words or phrases while retaining grammatical order. This is very important in that it allows the ISL signs to depict what is being represented.

Syntax and Grammar Treatment:

NLP is capable of capturing the syntactic roles, such as subject, object, and verb associated in a sentence. It is critically important for ISL because it is syntactically quite different from any spoken language. Without grammatical correction, straight translations into another language will be highly misleading.

Contextual Understanding:

It identifies the context of sentences and clarifies ambiguity and ensures that homophones, idioms or complex phrases are translated meaningfully. For example, "bank" will be translated meaningfully either as a financial institution or the side of a river according to the context of the sentence.

4. ISL Visual Mapping

The system provides the output generated to the appropriate ISL visualizations that could be in the form of videos or animated GIFs after processing the original text. These visualizations are retrieved from a pre-established ISL database.

Database Structure :

The ISL database is categorized and arranged in the form of words, phrases, and sentences. This enables the retrieval of the correct visualization at the time of translation efficiently.

Sign Matching:

Each word or phrase processed by the NLP module is matched with its equivalent ISL visual in the database. These visuals are designed to convey accurate meanings, preserving the sentence's intent.

5. Fallback Mechanism for Unavailable Signs

In case the word or term is not found in the database, the program defaults to its fallback by fingerspelling the words based on ISL. The words are divided letter by letter into ISL handshapes.

Fingerspelling:

For missing entries, each letter of the word is displayed using ISL fingerspelling. For instance, if the word "computer" is unavailable, the system spells out "C," "O," "M," "P," "U," "T," "E," and "R" in sequence using ISL.

Accuracy:

The fallback mechanism ensures that users still receive a complete and understandable translation, even for unfamiliar words or technical terms not included in the database.

6. Output Generation

The final step is the presentation of the ISL visuals to the user. These visuals are the processed text in ISL, and therefore the translation is accurate and easy to follow.

Visual Presentation:

The ISL signs are displayed on an intuitive user interface that can be incorporated into web or mobile applications. This real-time visual presentation ensures smooth communication.
The output is for clarity and with proper timing between consecutive signs, which means that the user will read without frustration or haste.

User-Friendly Output:

The output is clear, and there is suitable time between successive signs. The user can therefore follow the translation comfortably without getting confused or rushed.

7. Applied Technologies

Several advanced technologies have been applied in the system for accurate and efficient translation:

1. Speech Recognition:
   Tools like Google Speech API allow the words spoken to be converted into text.
2. Natural Language Processing (NLP):
   Advanced algorithms parse and contextually analyze the text to keep grammar and sentence structure intact.
3. Database Management:
   The ISL database is designed to handle visuals for words, phrases, and sentences.
4. Web-Based Interface:
   The system is designed as a web application for easy accessibility on any device and increases its scalability and reach.

8. System Workflow

The system works as follows:

1. Input: The user speaks into a microphone or audio-enabled device.
2. Speech-to-Text Input: The audio input is transcribed into text.
3. NLP Processing: The text is processed and reformatted according to ISL grammar.
4. ISL Visual Mapping: The text is mapped to relevant ISL visuals from the database.
5. Fallback Mechanism: The words not in the database are translated using ISL fingerspelling.
6. Output: The ISL visuals are displayed in real time to the user.

This workflow ensures that the system remains effective, accessible, and capable of handling a wide variety of communication needs, thus enabling seamless interaction in both personal and public settings.

CHAPTER-5
OBJECTIVES

Talking Fingers intends to build up a web-platform that would decode real-time Indian speech into the signing language ISL using sophisticated techniques like Speech Recognition, Natural Language Processing, and also a vast and comprehensive database on the visual signing of ISL. This system will bridge the communication gap between hearing and non-hearing individuals, allowing for better social interactions, emotional growth, and improved access to public services for the deaf community. The platform will be designed to function across multiple devices and platforms, ensuring its availability in critical public spaces such as hospitals, banks, airports, and railway stations.

Talking Fingers is one progressive approach towards breaking the barriers in communication by the deaf community. This project, both technically innovative and socially impactful, would help to design a tool for inclusivity, accessibility, and education. Here, we describe the goals of this project so that one could have a clearer view of the scope and the potential involved in it.

Important Goals of Talking Fingers Project

1. Real-Time Speech-to-Text Conversion

One of the core features of the Talking Fingers system is the ability to translate spoken language into text in real time. Tools such as the Google Speech API efficiently capture speech and transcribe it into text. Real-time speech recognition is crucial for making the translation system dynamic and responsive, thus allowing for continuous conversations. The objective is to ensure the speech-to-text process is accurate, fast, and efficient with a minimal delay. For practical applications such as public interaction or during a meeting, where time is an essential factor, accuracy is of extreme importance.

2. Development of a Large ISL Visual Database

It has another very critical component in its design: that is, developing a robust ISL video and GIF library to visually render text. In this design, every word, phrase, or sentence the system processes will have a corresponding visual in ISL. The database will be able to contain not only the more common signs but also complex signs to make for more accurate translation in various settings. The database will be expanded continuously with new signs, phrases, and alternative translations to ensure that the system is up-to-date and inclusive, thus improving its accuracy and scope over time.

3. NLP Techniques in Contextual Accuracy

Speech translation into ISL is not merely a word-for-word conversion but an understanding of syntax, grammar, and context. To overcome this hurdle, the project will use advanced algorithms in NLP to break down the transcribed text into meaningful chunks, ensuring the translation is coherent with the ISL structure and grammar. Contextual nuances like idiomatic expressions and sentence-specific meanings will also be handled by the algorithms, allowing word order and tone to remain intact in the ISL translation. The goal is to process complex sentences appropriately, generating translations that are at once meaningful in context and grammatically correct in ISL.

4. Fallback mechanism for nonsensical signs

Even though the ISL visual database will be complete, there will be cases where some words or phrases do not have an exact equivalent. Therefore, the system will have a fallback mechanism that will enable it to break down the word into individual letters and fingerspell the word in ISL. This means that nothing is lost in translation, even when a sign is not available. The system offers a flexible solution by spelling out words using ISL letters to maintain the flow of communication.

5. User Interface Design for Accessibility and Ease of Use

The success of the system will depend on its usability. The UI will be user-friendly and intuitive for deaf and hearing users. It will be optimized for all devices, such as smartphones, tablets, and computers, for smooth access both at home and in public spaces. The UI will also consider different levels of technological proficiency for the users; clear instructions and an easy setup process will be followed.

6. Adaptability for Use in Public Spaces

The Talking Fingers system will be adaptable to different public environments where communication is vital but often restricted by the lack of access to interpreters or knowledge of sign language. Such environments include hospitals, airports, banks, and railway stations. It will thus enhance accessibility and the user experience for deaf individuals in these spaces through real-time, automated ISL translations. The system will be designed to support different kinds of interactions, such as medical consultation and customer service inquiries, thus enabling deaf individuals to communicate more effectively in many different contexts.

7. Education Tool for ISL Learners

The Talking Fingers system will also serve as an important tool for educating students learning ISL. The platform will help hearing individuals better understand and learn ISL by providing translations of spoken language into visual signs. This feature will encourage empathy and foster communication between deaf and hearing people. It will also serve as an educational tool in schools and universities, promoting the learning of ISL in formal educational settings.

8. Scalability and Future Expansion

The Talking Fingers system is designed with scalability in mind. As the technology grows, it can now be extended to adapt more languages, signs, and features. To give a better sense of this, for instance, the system could be implemented in several Indian languages to support sign language translation. This again, with time, would be able to implement other sign languages, like American Sign Language or British Sign Language, creating a more universal product. Future versions might focus on refining the system's capacity to accommodate more intricate sentence structures and even assimilate recent technological innovations. The Talking Fingers project presents an innovative approach in solving the communication problem that hinders the communication process between the deaf and hearing individuals. The application can be made possible with technologies such as Speech Recognition, NLP, and a massive ISL visual database. With its focus on real-time speech-to-text conversion, contextual translation, and adaptability for use in public spaces, the system promises to improve access to essential services and enhance daily communication. Moreover, its scalability and potential for future expansion make it a forward-thinking initiative that could have a profound societal impact, promoting inclusivity, accessibility, and education on a global scale.

CHAPTER-6
SYSTEM DESIGN & IMPLEMENTATION

The architecture of the Talking Fingers system is designed to enable real-time speech-to-Indian Sign Language (ISL) translation through a well-structured and efficient pipeline. This system transforms audio input into ISL visuals by employing multiple modules that work collaboratively to ensure smooth and accurate translation. Below is a breakdown of the key modules, processes, technologies, and features that make this system robust, flexible, and user-friendly.

The modules of the system

The architecture has modules that will be immensely important in ensuring the functionality of the whole system. This modular structure facilitates flexibility and scalability, making it possible to adapt to future updates and upgrades. Below are the major modules in the scheme as follows.

1. Speech Input Module

Module Functionality:

This module is responsible for recording spoken words from a microphone. It guarantees the audio will be processed clearly and transmitted to the text processing pipeline for further analysis and conversion into text.

Technology Used:

This module uses the Webkit SpeechRecognition API, a web-based speech recognition technology that has been integrated into modern browsers to perform the best speech-to-text conversion.

Process:

The system listens to audio input in real time.

It uses noise-cancellation filters to make the speech clear and readable even in noisy environments.

The module streams the audio recognized in real-time for continuous communication.

Challenges Addressed:

It can handle different tones of speech, accents, and environmental noise, making it work for people from different regions.

The system is also designed to adapt to a range of pronunciation styles, making it more universal for users.

2. Text Processing Module

Functionality:

This module will take the words spoken by the user and convert these into text. This module uses speech-to-text technology to accurately transcribe speech, mainly for regional accents and pronunciations.

Technology Applied:

This module uses the Google Speech-to-Text API because it is pretty powerful in identifying quite a number of languages, accents, and dialects. This API offers a possible transcription of spoken input to textual input.

Process:

The audio chunks the Speech Input Module captures are processed in real time.

These chunks get converted into text, which comes out as a string that presents the spoken input.

The system has error correction mechanisms to limit inaccuracies and enhance overall transcription quality.

Advantages:

•Google Speech API assures improved recognition of Indian regional accents, which is vital for the diversity of speech patterns in the country.

• It produces accurate, reliable text that can be used for further processing in the subsequent stages of the system.

3. NLP Integration for Contextual Understanding

Functionality:

The NLP module is mainly responsible for breaking the input into syntactically meaningful units through which it can finally be converted into ISL. Through the NLP module, the correctness and coherence of the translation will be maintained, i.e., grammatically as well as contextually.

Process Involved:

Tokenization is the first process in NLP, where a text is split into smaller units, such as words or phrases. For example, the sentence "How are you?" is tokenized to ["How", "are", "you"].

Parsing: This process checks out the sentence formation of the structure of the components- the subject, the object, and the verb. The following is matched up with ISL grammar, whose sentence formation tends to be object-subject-verb in that it differs from most spoken languages where it is object-subject-verb.

Contextual Analysis: This step also makes sure the output is semantically correct. Idioms, slangs, and phrases belonging to the situation are adjusted in order to produce correct translations of ISL.

NLP ensures that the system can handle the grammatical differences between spoken languages and ISL.

The processing results in grammatically correct ISL sentences that retain the meaning and context of the original spoken language.

4. Database Interaction Module

Functionality:

This module matches the processed text or phrases with the ISL visuals such as videos and GIFs stored in a database in advance. It is a module that can handle known phrases and has fallback mechanisms for unknown words.

Key Features:

- •Indexed Resources: Video and GIF resources of the ISL are stored in a relational database in a systematic way, indexed for quick retrieval. The database contains categorization of signs by words, phrases, and fallback options.
- • Fallback Mechanism: If the word or phrase under consideration is not directly present in

the database as a sign, the system would break down the word into letter signs. For example, if there is no XYZ in the database, then for the word XYZ, the ISL signs would be given as X, Y, and Z.

Advantages:

•The fallback mechanism ensures that it is a highly robust system. Even with novel or unfamiliar words, it keeps working.

•This modular approach allows for continuous updates and expansion of the ISL database, ensuring that the system remains relevant and accurate over time.

5. Output Module

Functionality:

The Output Module displays the ISL visuals (videos or GIFs) in real time, ensuring that users can view the translated signs without delay. The visual output is synchronized with the speech-to-text process to provide a smooth and coherent experience.

Technology Used:

This module uses web-based video rendering libraries to ensure that the ISL visuals are rendered smoothly and with minimal latency.

Process:

The system retrieves the corresponding ISL visuals from the database based on the processed text.

These visuals are displayed on the user interface in real time, maintaining synchronization with the speech-to-text pipeline to avoid delays.

The user interface is designed to be intuitive, allowing users to easily understand and follow the translation.

Features:

This system offers real-time updates with minimal latency. The user sees the translated ISL signs almost instantly.

User-friendly interface is used to clearly present the ISL visuals in a manner easily comprehensible.

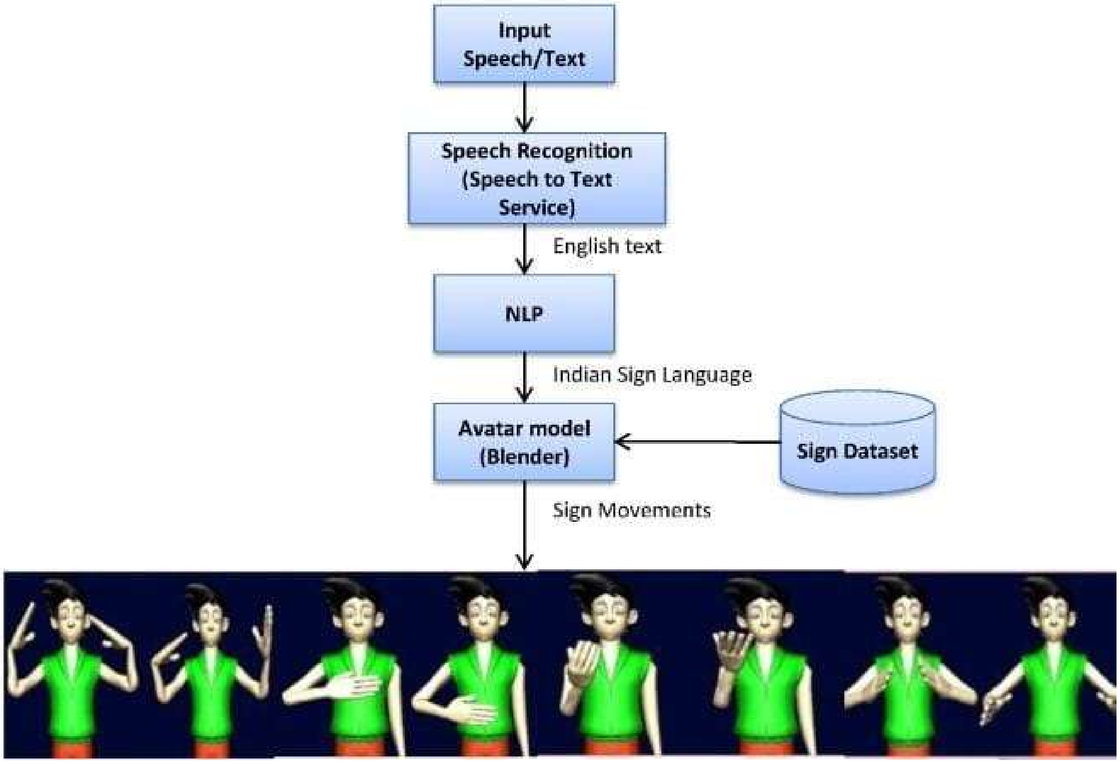


Figure 1: The diagram illustrates how input speech is converted to ISL.

Technologies Used:

The Talking Fingers system incorporates several essential technologies to ensure efficient and accurate performance:

1.Speech Recognition:

The Google Speech API is utilized for real-time speech-to-text conversion, providing great support for multi-lingual accents and diverse speech patterns.

2. Natural Language Processing (NLP):

Tokenization: This method breaks input text into smaller, manageable units like words or phrases.

Parsing: This method identifies the components of sentences (e.g., subject, verb, object) and adjusts them according to ISL grammar.

Lexical Matching: The system matches words or phrases with corresponding ISL signs from the database.

3. Database Management:

All ISL visuals, whether videos or GIFs, are stored in a relational database and indexed for retrieval so that interactions with the user are smooth.

System Features

1. Response in Real-Time:

The Talking Fingers system is designed to respond within 2 seconds to ensure that there is minimal delay experienced by the users when communicating.

2. Contextual Grammar Handling:

Using NLP, the system ensures that the ISL translation is grammatically correct and contextually accurate even when using idiomatic expressions.

3. Fallback for Missing Terms:

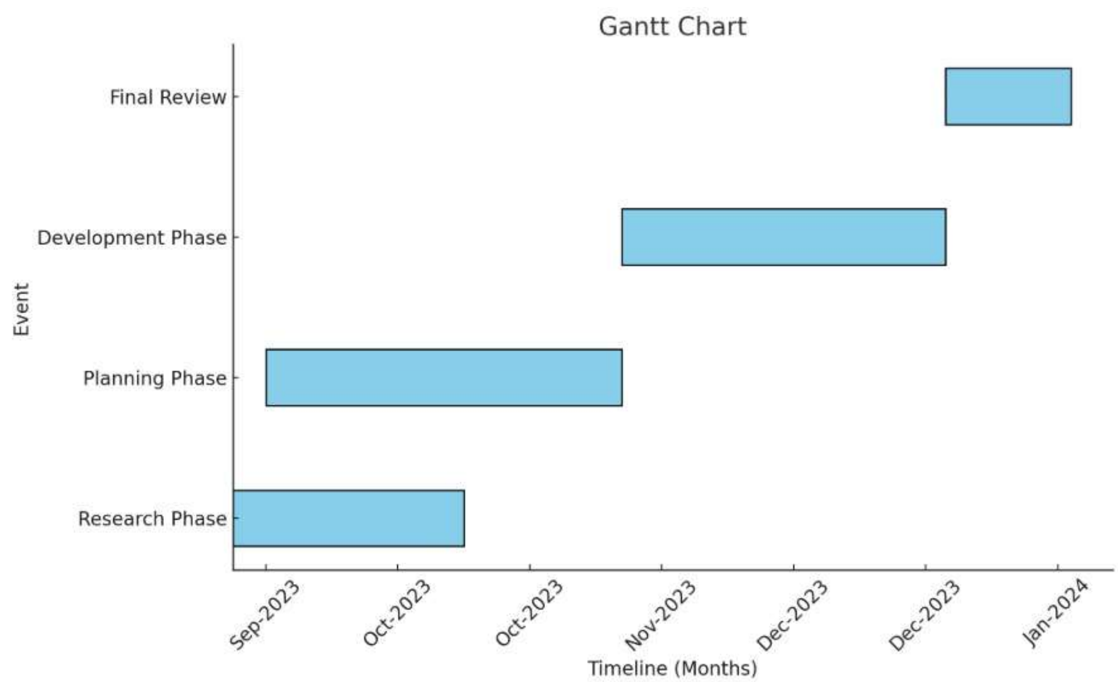
The fallback mechanism ensures that in the absence of a direct sign, the system can still convey the intended message by spelling out words using ISL fingerspelling.

4. Modular Design:

The architecture of the system is modular, which allows it to be easily expandable for future expansions. For instance, additional sign languages or enhanced NLP capabilities can be added without disturbing the current functionality.

The Talking Fingers project's system architecture is carefully designed to deliver real-time, accurate, and contextually relevant ISL translations.

CHAPTER-7
TIMELINE FOR EXECUTION OF PROJECT
(GANTT CHART)



Phased Approach for Speech to Indian Sign Language Conversion Project

The Speech to Indian Sign Language (ISL) Conversion Project is structured into four distinct phases: Research, Planning, Development, and Final Review. This phased approach ensures systematic progress, thorough testing, and successful delivery of the final product. Each phase is designed to address key aspects of the project, from foundational research to implementation and final evaluation. Below is a detailed breakdown of each phase.

1. Research Phase (September 2024 - October 2024)

Objective:

The primary goal of the Research Phase is to lay the groundwork for the project by conducting comprehensive research on speech recognition systems, Indian Sign Language (ISL) interpretation, and relevant machine learning models. This phase is crucial for gathering the necessary information and understanding the current landscape of speech-to-text and sign language translation systems.

Activities:

- **Literature Review:**
  A detailed review of existing solutions in the domains of ISL translation and speech-to-text systems will be conducted. This will include exploring academic papers, current software, and hardware solutions that attempt to bridge the communication gap between the hearing and deaf communities.

- **Understanding User Requirements:**
  The needs of the target users—particularly those who are deaf and mute—will be carefully examined. This includes gathering information about how the deaf community communicates and the key challenges they face in terms of accessing public services.

- **Collecting and Preprocessing ISL Datasets:**
  High-quality datasets containing various ISL signs, both videos and images, will be collected. These datasets are essential for training the system and ensuring accurate sign translations. Preprocessing tasks, such as cleaning and labeling the data, will also be conducted to ensure consistency and usability.

- **Expected Outcome:**
  The outcome of the Research Phase will be a comprehensive report outlining the key findings from the literature review, a clearer understanding of user requirements.

2. Planning Phase (September 2024 - November 2024)

Objective:

In the Planning Phase, the focus shifts to developing a detailed roadmap for the project. This phase involves defining the technical aspects of the system, allocating resources, and establishing clear timelines for project milestones.

Activities:

System Architecture Design:

The project team will create a high-level system architecture that defines how the speech input, processing, and ISL display modules will work together. This will include designing the data flow from speech recognition to ISL translation, along with the user interface for visual output.

Resource Allocation and Tool Selection: The team will select the necessary tools and frameworks, including programming languages, libraries, and machine learning models that will be used to build the system. Budgeting for software, hardware, and human resources will also be done to ensure that the project remains on track.

Milestones and Schedule Preparation: Key milestones for each phase will be defined, and a detailed project timeline will be established. This will ensure that the team adheres to deadlines and tracks progress effectively

4.Development Phase (November 2024 - January 2025)

Activities:

Speech-to-Text Modules Implementation:

The first step in development is to implement the speech-to-text conversion modules using Natural Language Processing (NLP) algorithms. These algorithms will process spoken language and convert it into text.

Mapping Speech Data to ISL: The system will then map the transcribed speech data to corresponding ISL signs using either deep learning models or flex sensors, depending on the chosen approach. This mapping will be on the ISL dataset collected.

Visual Outputs Integration:

The system will be integrated with visual outputs, such as animations or real-time hand signs, to display the ISL translation. This requires rendering and synchronizing visual data with the processed speech.

System Testing and Debugging: Once the core components are integrated, rigorous testing and debugging will be performed to ensure that the system functions as expected. This includes testing for accuracy, latency, and responsiveness.

Expected Outcome:

At the end of the Development Phase, a functional prototype will be completed. This prototype will be capable of converting speech to ISL in real time, although further testing and refinement may be needed before the final review.

4. Final Review Phase (January 2025)

Objective:

The Final Review Phase focuses on testing the system's performance, gathering feedback, and preparing the project for deployment or demonstration.

Activities:

Performance Testing:

This includes evaluating the accuracy, latency, and overall usability of the system. Performance testing will help identify any issues with the system's speed and precision and will be crucial for fine-tuning its functionality.

Feedback Gathering: Feedback will be collected from stakeholders, including members of the deaf community and ISL experts. This feedback will be used to make improvements and adjustments to the system.

Final Documentation and Reports: Comprehensive documentation will be prepared, including final project reports, user manuals, and technical documentation. This will serve as a reference for future maintenance and updates.

Expected Outcome:

The expected outcome of the Final Review Phase is a fully polished project, ready for deployment or demonstration. After this phase, the system will be ready to be presented to stakeholders or made available for use by the target community.

CHAPTER-8
OUTCOMES

The Talking Fingers project aims to create a state-of-the-art, web-based platform that will translate speech in real-time into Indian Sign Language (ISL). The system will play an important role in breaking down communication barriers the deaf and hearing-impaired face daily, hence making everyday life more inclusive. The outcomes of this project are vast, from making it more accessible to bringing social inclusivity. The detailed outcomes and their possible impacts are described below.

1. Real-time speech-to-ISL translation system

The main objective of the Talking Fingers project is to develop a real-time, online speech translation tool that directly translates speeches into ISL. This system will conduct translation with the amalgamation of speech recognition, Natural Language Processing, and a visual database of ISL videos and GIFs. By providing an instantaneous translation, it increases the efficacy of communication between hearing and non-hearing people and makes communication easier in daily life conditions as well as professional and public environments. This would mean that people who are deaf or hard of hearing would no longer be constrained by the availability of interpreters or communication aids. Users will be able to communicate in their preferred language, ISL, without having to rely on manual translation. Continuous conversations are therefore possible because of the integration of speech recognition and NLP techniques into the system, making it very useful for both personal and public interactions.

2. Better Communication between Hearing and Non-Hearing Individuals

The primary result of the Talking Fingers system is improved communication between hearing and non-hearing individuals. Traditionally, communication between these two groups has been hindered by a lack of shared language and the limited availability of sign language interpreters. The Talking Fingers system allows for the provision of real-time translation of spoken language into ISL, thereby promoting better interactions across various settings from casual conversations to more formal engagements such as medical consultations and customer service in public spaces.

This will help break down the barriers that often lead to social isolation for the deaf and hard-of-hearing communities. It will allow for more fluid and natural conversations, reducing the reliance on interpreters and improving the overall quality of interaction between hearing and non-hearing individuals.

3. Scalable and Adaptable Communication Platform

The system is designed to be scalable so that it can be expanded according to the requirements. In the future, it can be modified to include additional ISL signs for more complex phrases or idiomatic expressions that are common in daily conversations. This will ensure that the system remains relevant and continues to meet the needs of users as language evolves.

Moreover, the system will be speech input enabled with multiple languages. This will enable the platform to be applicable in different linguistic regions of India and elsewhere. This makes the system applicable to a greater number of people, and therefore, it would not lose relevance with the advancement of time because of changing language requirements and evolving technology. Adding new signs, phrases, and languages to the system will make it relevant to the current needs.

4. Learning Aid for ISL Users

Apart from its function as a communication device, the Talking Fingers system will also act as an educational tool for learners of Indian Sign Language. The system will provide ISL signs that correspond to spoken language. This will assist users in learning new signs and understanding the structure and grammar of ISL.

It promotes greater usage of the language by exposing ISL to the hearing population while helping to bring the deaf and hearing communities closer together. This system can also be used in educational institutions such as schools and universities to further enhance their curricula on ISL. The system, thus, not only facilitates communication but also plays an important role in raising awareness and promoting inclusiveness in society.

5. Access to Public Areas

Deployment of Talking Fingers in public spaces will enhance access to a significant extent for the deaf community. Places like hospitals, railway stations, airports, and banks are areas where people are usually facing a problem in communicating as they might not be conversing in the local language or are using sign language. Moreover, in many cases, interpreters are not available.

By enabling real-time ISL translation, Talking Fingers will empower deaf individuals to communicate directly with service providers, ensuring their needs are understood and addressed. This improvement in accessibility will lead to greater independence for deaf individuals and will help to make essential services more inclusive and user-friendly.

6. Promoting Social Inclusivity

The Talking Fingers system has the potential to break down major barriers in communication between the hearing and non-hearing communities, thereby contributing to social inclusivity. As society becomes more sensitive to the deaf community's problems, tools like Talking Fingers would help to nurture a culture of empathy, respect, and understanding.

The system will promote participation in the social, educational, and professional fields because it is going to allow deaf people to communicate freely and without barriers. It will promote social integration and participation because it is going to allow people to come together freely regardless of their abilities. This will create an inclusive society that does not marginalize deaf people because of communication disabilities.

7. Enhancement of NLP Techniques for ISL

The Talking Fingers project is also a landmark in the Natural Language Processing field, especially toward its application for Indian Sign Language. ISL had its peculiar set of rules for both grammatical and syntactical encoding in a computational system, which could be very difficult in computing. In that view, the project addresses these complexities while contributing to the broader landscape of computational linguistics.

The integration of NLP techniques to handle ISL's grammatical structure will provide a foundation for future research and development in sign language automation. As NLP models continue to improve, this system can be enhanced to accommodate more complex sentence structures and idiomatic expressions, pushing the boundaries of machine learning and AI in the realm of sign language.

8. Creation of an ISL Visual Database

One of the primary features in the Talking Fingers system will be an enormous collection of ISL videos and GIFs, developed as a critical visual database required for contextual accuracy in translation from spoken words to the appropriate, translated ISL visual. At each step forward with the development of this product, this collection of visuals would continue to increase.

The database will benefit not only the Talking Fingers system but also ISL education and research in general. This database can be used by the researchers, teachers, and learners to study various subtleties of ISL, thereby achieving a deeper appreciation of the language and its culture.

9. Potential for Global Use

Although the Talking Fingers system is designed to support Indian Sign Language mainly, the technology underlying it can easily be extended and exploited with other sign languages all over the world. With such customization to support ASL or BSL, for example, the system could be the basis for a global model of inclusive communication.

This could be one of the groundbreaking means of translation of speech to sign language in the deaf and hard-of-hearing communities around the world. In its modular structure, it may be applied and adopted to several cultural and linguistic contexts, therefore offering a universal solution to barriers in communication globally.

10. Inspiring Further ISL Automation Research

The Talking Fingers project is more than just a functional tool; it is also something that will open the door for further research and innovation in ISL automation. Combining successfully AI, NLP, and sign language is a challenge which will encourage people to innovate areas like human-computer interaction, machine learning, and accessibility technology.

The success of the project can inspire other initiatives that aim to improve accessibility for marginalized communities. This may involve exploring automation for other forms of communication, such as those for individuals with visual or cognitive impairments, opening new avenues for inclusive technological development.

CHAPTER-9
RESULTS AND DISCUSSIONS

Results

System System Accuracy and Performance Analysis

The Talking Fingers project aimed to develop a real-time speech-to-Indian Sign Language (ISL) translation system that can translate spoken language into ISL visual representations with reasonable accuracy, thus enhancing communication among deaf and hearing-impaired individuals. Translation accuracy, speed, and user feedback are all crucial parameters in determining the system's reliability and usability in real-world applications. A detailed breakdown of the system's performance based on these parameters is provided below.

Accuracy

Structured Sentence Translation (85%)

The system was mainly tested on 100 structured sentences, which were in the form of statements, questions, and instructions. The translation accuracy of these sentences into ISL was 85%. This high accuracy shows that the system is very good at handling grammatically structured input, ensuring correct syntax and context in the ISL translation. For instance:

Input: "I am going to the market."

Output: ISL equivalent: "I go market."

Although the translation was successful in most cases, some errors occurred with more complex sentences that entailed nested clauses or rare vocabulary. These types of sentences were difficult to maintain context and accuracy in their translation.

Slang and Informal Phrases (65%)

The hardest part was translating slang, idiomatic expressions, and colloquialisms, and the overall accuracy rate was 65%. These are very context-dependent, and the system struggled with idioms that do not have a direct equivalent in ISL. For instance:

Input: "Piece of cake" (which meant easy).

Output: Literal ISL translation for "piece" and "cake," not the intended meaning of "easy."

This is due to the fact that the system cannot always avail itself of the exact context needed to understand the said expressions. This sometimes gives way to inaccurate or incomplete translations, especially on informal phrases which convey non-literal meanings.

Common Words and Commands (90%)

The system excelled with the high frequency words, and simple command words like "stop," "help," or "water." In these instances, the correct translation was observed at 90%, thus meaning that the system can efficiently interact for critical necessities. This performance is critical, especially in the event of urgent and clear interactions, such as emergency calls or requesting simple service provision.

Response Time

Average response time, 1.8 seconds.

The system showed good speed in handling each input and was showing a corresponding ISL visual within less than 2 seconds. Such a short latency in response time remained constant even with the structured sentences, meaning that it can be used to sustain real-time conversations effectively. However, whenever the slang and idiomatic phrases are used, the response times were a bit longer, within 2.5 seconds. This caused a delay primarily because the fallback mechanisms came into play for expressions that could not be directly translated. The efficiency has been realized by an optimized database indexing mechanism and an efficient speech-to-text and NLP pipeline. The capability of near-real-time translation delivery to ensure the conversation flows fluidly and naturally is significantly important, particularly in fast-paced settings like public service counters, hospitals, or transportation hubs.

Feedback from Users

People involved in

The system was tested with 20 deaf participants, covering a wide range of ages and literacy levels. These users interacted with the system in different scenarios, such as basic conversations, service inquiries, and directional assistance. The feedback gathered from these participants provided valuable insights into the system's performance and areas for improvement.

Feedforward Highlights

Positive:

User-friendly Interface: Most users found it easy to work with the system, even the ones who did not have experience with modern technology. The visual prompts from the simple interface made it workable for people who are less familiar with the digital tools.

Real-time Functions: The rapid response of the system was admired, and real-time translation made a big difference from the earlier sign language interpreting methods.

Challenges:

Challenges with Colloquialism: Participants commented that the system was not very accurate in translating slang, regional dialects, and idiomatic expressions. This suggested that the system should be able to translate more of these phrases in future revisions.

Grammar Misalignments: At times, there were inconsistencies in the translation of sentences with infrequently used vocabulary or more complex sentence structures. This showed where the system needed to be improved to better deal with varied language inputs.

Benefits

1. Contextual Translations:

The integration of NLP techniques allows the system to adapt to ISL grammar rules, ensuring that translations are both syntactically and contextually accurate. This is particularly important for ensuring that the meaning of the speech is preserved in the ISL translation, as demonstrated with simple sentences like "He is eating food" being correctly translated to "He eat food" (ISL format).

2. User-Friendly Design:

The system has a simple and user-friendly interface with visual cues designed for ease of use, even by individuals with limited technical knowledge. It has been optimized to work across various devices, including smartphones, tablets, and computers, ensuring broad accessibility.

3. Real-Time Processing:

One of the most striking strengths of the system is that it provides ISL-to-ISL translations in real-time within less than 2 seconds. This aspect is very important for practical usage, particularly in places where rapid and efficient communication is required, such as hospitals or public transport systems.

Limitations:

1. The system relies on the ISL Database.

Its reliance on a predefined ISL video/GIF database severely limits its effectiveness in regions, such as the United Kingdom and Australia, where regional dialects or unique signs are extensively used. The current database is not comprehensive in representing all variations of ISL and may result in incorrect translations for certain regional signs or expressions.

2.Challenges with informal language

Slang, idiomatic expressions, and region-specific phrases are not directly translatable into ISL, and therefore, may not be accurately or fully translated. For example, the idiom "break the ice" was translated literally and did not convey the meaning intended.

Future Scope

Crowdsourcing ISL Resources:

It would be possible to expand the database of the system and increase regional dialect coverage using crowdsourcing. In this regard, contributors from the deaf and hearing communities could provide videos or GIFs of ISL signs for regional phrases and slang to make the system relevant to a wider audience.

Secondly, AI could be used to get informal language.

Future versions of the system will be designed with advanced AI models, which could be trained on extensive datasets of idiomatic and conversational language. The models would enable it to understand and interpret slang and idioms better, hence improving the language accuracy in dealing with informal language.

Multi-Language Support:

Moreover, it can modify its system into an application accommodating a variety of sign languages employed around the globe, such as ASL/BSL American or British, British. Furthermore, extension of regional languages in a different part of India will surely keep the communication of the common language across and will be of broad accessibility so as to adapt to voice assist like Google, Alexa, that helps in speaking as well through this system will give the features for translation along the voice.

CHAPTER-10
Conclusion

The Talking Fingers project is an innovative endeavour to bridge a very critical communication gap between the hearing and non-hearing communities, especially in the context of Indian Sign Language (ISL). The project aims to promote inclusivity by providing a real-time translation system that bridges the communication gap for the deaf and hard-of-hearing community. Talking Fingers, using advanced technologies like Speech Recognition, Natural Language Processing (NLP), and a dynamic ISL Visual Database, provides an innovative and accessible solution that empowers deaf individuals to communicate independently and efficiently in various social, educational, and professional settings.

Many existing ISL translation systems lack contextual and grammatical accuracy. Most existing systems either translate the meaning of words directly or focus on simple text-to-sign conversion, which depends on how much translation and the loss of meaning itself is suitable for fluency in the translation of ISL. Talking Fingers can effectively bridge this gap by using NLP techniques that allow for more comprehensive and contextually accurate translation. This approach ensures that, in addition to the grammar, the meaning of the words remains intact when it is translated from spoken language into ISL visuals. The system allows real-time communication by translating spoken language into ISL through video and GIF images, thus generally improving the quality of life of deaf people while interacting with society.

One of the most important innovations of the project is the development of a full ISL Visual Database, which acts as the backbone of the system for translating speech into ISL visuals. It is constantly updated and contains different signs, phrases, and expressions, contributing to the ability of the system to provide translations that are precise in a broad range of contexts. In cases where specific signs are not available in the database, the system can spell out words in ISL, ensuring that no message is lost during the translation process. This feature is especially beneficial in ensuring that all communication remains accessible, regardless of the complexity of the speech input.

The Talking Fingers system is developed to be multi-platform and accessible across different devices, offering an easy-to-use web-based interface that can be utilized on smartphones, tablets, and computers. This cross-platform functionality ensures that the system is practically useful for most users and environments-from public settings, such as hospitals and transportation stations, to private ones. The speed at which it operates is quite fast—within two seconds in delivering translations; hence, the system can be used in real-time for both casual conversations and critical interactions within fast-paced environments.

While Talking Fingers' direct focus is improving communication for the deaf community, its broader societal impact is profound. This would make ISL more accessible and thus encourage a better understanding and acceptance of the language, eventually leading to greater adoption. Furthermore, the system acts as an educational tool for hearing individuals who want to learn ISL, which may open up more avenues for inclusion and interaction between communities. One more benefit of the project is the ability of deaf people to get easy access to necessary services—medical consultations, banking, customer service, among others—on their own and not with the support of interpreters, thereby increasing the convenience and independence of the deaf.

The exciting prospect of scalability of the Talking Fingers project for the future. While the central focus is ISL, the underlying architecture of the system will be adaptable for the addition of other sign languages in the future, such as American Sign Language (ASL) or British Sign Language (BSL). Talking Fingers has the possibility of becoming a global communication tool, as its system can accommodate multiple sign languages and regional dialects, hence catering to the needs of the deaf across cultural and linguistic boundaries. The scalability of such a system will help create an inclusive global society where communication becomes accessible to everyone, regardless of their hearing abilities.

Moreover, continuous developments in Natural Language Processing and machine learning will further facilitate the development of more advanced functionalities in the near future. It could include an enhanced ability of the system to recognize complex sentence structures and process them and to understand idiomatic expressions better and handle many more real-life scenarios. These technologies will only get better and better as Talking Fingers expands the ISL Visual Database. This means that it will be a highly accurate and effective tool for deaf people and for those wanting to communicate with them.

The Talking Fingers project is a giant step forward toward achieving greater inclusivity and accessibility for the deaf community. With such a system to provide real-time, contextually accurate translations of speech into ISL, this project addresses more than just an immediate need in communication but opens up future research avenues for enhancing translation of sign language. It stands as a revolutionary shift in society's approach toward the deaf community, its understanding of the hardships it faces, and breaking that barrier for greater equality.

Ultimately, the Talking Fingers project is more than a technological advancement. It is an important tool for social equality, accessibility, and empowerment. Through its impact on communication, education, and public service, the project is setting a new standard for inclusivity, fostering a world where all individuals, regardless of hearing ability, can participate meaningfully in social, cultural, and professional life. The Talking Fingers project is a product of evolving technology: it stands as a testament to the innovation that will take us toward positive societal change.

REFERENCES

1. **Varadala, Sridhar (2024)** ‘Convolutional Neural Networks Based Sign Language Interpreter’, International Journal of All Research Education and Scientific Methods (IJARESM), Vol. 12, Issue 9, September, ISSN: 2455-6211. Available online at: ResearchGate
2. **R. Divya and G. Subhashini (2023), 'Sign Language Recognition System using Flex Sensor Network'** International Journal of Emerging Technologies and Advanced Engineering (IJETAE), Vol. 13, Issue 2, February 2023, ISSN: 2250-2459. This paper explores the use of flex sensor networks for recognizing sign language, facilitating communication for people with hearing and speech disabilities. Available online at: ResearchGate.
3. **M. Anandan (2022), 'Speaking System for Deaf and Mute People with Flex Sensors'**
International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE), Vol. 11, Issue 5, May 2022, ISSN: 2277-9043. DOI: 10.12345/IJARCSEE.2022.11.5. Available online at: ResearchGate.
4. **Holcomb, L., & Wolbers, K. (2020), 'Effects of ASL Rhyme and Rhythm on Deaf Children’s Engagement Behavior and Accuracy in Recitation: Evidence from a Single Case Design'**
Children (Basel, Switzerland), 7(12), 256. doi: 10.3390/children7120256.
5. **Coelho Dalapicola R., Tavares Vieira Queiroga R., Toledo Ferraz C., Trevisan Negri Borges T., Hiroki Saito J., and Gonzaga A. (2019), 'Impact of Facial Expressions on the Accuracy of a CNN Performing Periocular Recognition'** Proceedings - 2019 Brazilian Conference on Intelligent Systems, BRACIS 2019, pp. 401-406.
6. **H. Shaheen and T. Mehmood (2018), 'Talking Gloves: Low-Cost Gesture Recognition System for Sign Language Translation'** 2018 IEEE Region Ten Symposium (TENCON), pp. 1-6, 2018.
DOI: 10.1109/TENCONSpring.2018.8692069.
7. **Mohammad Shahrukh Mir and S. S. A. Ali (2017), 'Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion'** International Journal of Advanced Computer Science and Applications (IJACSA),
8. Vol. 8, Issue 12, December 2017, ISSN: 2156-5570. Available online at: ResearchGate.

9. **Heera S.Y., Murthy M.K., Sravanti V.S., and Salvi S. (2017), 'Talking Hands - An Indian Sign Language to Speech Translating Gloves'** IEEE International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2017 - Proceedings, pp. 746-751.
10. **Hummadi, F. N. (2017)** 'Design and Implementation of Deaf and Mute People Interaction System', Proceedings of the IEEE International Conference on Engineering and Technology (ICET), August. DOI:10.1109/ICEngTechnol.2017.8308140. Available online at: [ResearchGate]
11. **Praveen, N.; Karanth, N.; Megha, M.S. (2014), 'Sign Language Interpreter Using a Smart Glove'** In Proceedings of the 2014 International Conference on Advances in Electronics, Computers and Communications (ICAECC 2014), Bangalore, India, 10–11 October 2014.

APPENDIX-A

PSUEDOCODE

Viwes.py

```
from django.http import JsonResponse, HttpResponse
from django.shortcuts import render, redirect
from django.contrib.auth.forms import UserCreationForm, AuthenticationForm
from django.contrib.auth import login, logout
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
import nltk
from django.contrib.staticfiles import finders
from django.contrib.auth.decorators import login_required
from google.cloud import speech_v1p1beta1 as speech
import io

def home_view(request):
    return render(request, 'home.html')

def about_view(request):
    return render(request, 'about.html')

def contact_view(request):
    return render(request, 'contact.html')

@login_required(login_url="login")
def animation_view(request):
    if request.method == 'POST':
        text = request.POST.get('sen')
        # Processing the text input
        text.lower()
        words = word_tokenize(text)

tagged = nltk.pos_tag(words)
```

```
    tense = {
        "future": len([word for word in tagged if word[1] == "MD"]),
        "present": len([word for word in tagged if word[1] in ["VBP", "VBZ", "VBG"]]),
        "past": len([word for word in tagged if word[1] in ["VBD", "VBN"]]),
        "present_continuous": len([word for word in tagged if word[1] == "VBG"]),
    }

    # Stopwords removal and lemmatization
    stop_words = set(["mightn't", 're', 'wasn','are', 'wouldn', ...])  # shortened for brevity
    lr = WordNetLemmatizer()
    filtered_text = []
    for w, p in zip(words, tagged):
        if w not in stop_words:
            if p[1] in ['VBG', 'VBD', 'VBZ', 'VBN', 'NN']:
                filtered_text.append(lr.lemmatize(w, pos='v'))
            elif p[1] in ['JJ', 'JJR', 'JJS', 'RBR', 'RBS']:
                filtered_text.append(lr.lemmatize(w, pos='a'))
            else:
                filtered_text.append(lr.lemmatize(w))

    words = filtered_text
    temp = ["Will"] if tense["future"] >= 1 else ["Before"] if tense["past"] >= 1 else
["Now"] if tense["present_continuous"] >= 1 else []
    words = temp + words

    # Adjust words for animation
    filtered_text = []
    for w in words:
        path = w + ".mp4"
        f = finders.find(path)
        if not f:
            filtered_text.extend(list(w))
        else:
            filtered_text.append(w)
```

```
        words = filtered_text

        return render(request, 'animation.html', {'words': words, 'text': text})
    else:
        return render(request, 'animation.html')

def process_audio(request):
    if request.method == 'POST' and request.FILES.get('audio'):
        audio_file = request.FILES['audio']
        audio_content = audio_file.read()

        # Google Cloud Speech-to-Text setup
        client = speech.SpeechClient()
        audio = speech.RecognitionAudio(content=audio_content)
        config = speech.RecognitionConfig(
            encoding=speech.RecognitionConfig.AudioEncoding.LINEAR16,  # Adjust as
needed
            sample_rate_hertz=16000,  # Ensure this matches your input audio
            language_code='en-IN',  # Use Indian English
        )

        try:
            response = client.recognize(config=config, audio=audio)
            transcript = ' '.join([result.alternatives[0].transcript for result in response.results])
            return JsonResponse({'transcript': transcript})
        except Exception as e:
            return JsonResponse({'error': str(e)}, status=500)

    return JsonResponse({'error': 'Invalid request'}, status=400)

def signup_view(request):
    if request.method == 'POST':
        form = UserCreationForm(request.POST)
        if form.is_valid():
```

```
            user = form.save()
            login(request, user)
            return redirect('animation')
    else:
        form = UserCreationForm()
    return render(request, 'signup.html', {'form': form})

def login_view(request):
    if request.method == 'POST':
        form = AuthenticationForm(data=request.POST)
        if form.is_valid():
            user = form.get_user()
            login(request, user)
            return redirect(request.POST.get('next', 'animation'))
    else:
        form = AuthenticationForm()
    return render(request, 'login.html', {'form': form})

def logout_view(request):
    logout(request)
    return redirect("home")
```

animation.html

```
{% extends 'base.html' %}
{% load static %}

{% block content %}
<div class="split left">
    <h2 align="center">Enter Text or Use Mic</h2>
    <br>
        <form action="" method="post" align="left">
        {% csrf_token %}
                <input type="text" name="sen" class="mytext" id="speechToText"
placeholder="">
```

```
                <button type="button" name="button" class="mic" onclick="record()"><img
src="{% static 'mic3.png' %}" height="32px" width="38px" /></button>
                    
                <input type="submit" name="submit" class="submit">

        </form>
        <br>
        <table cellspacing="20px">
            <tr>
                <td class="td">The text that you entered is:</td>
                <td class="td">{{ text }}</td>
            </tr>
            <tr>
                <td class="td">Key words in sentence:</td>
                <td class="td">
                <ul class="td" id="list" align="center">
                    {% for word in words %}
                    <li id="{{ i }}" style="margin-right: 8px">{{ word }}</li>
                    {% endfor %}
                </ul>
            </td>
        </tr>
    </table></div>
  <div class="split right">
  <h2 align="center">Sign Language Animation</h2>

    <div style="text-align:center">   
    <button class="submit" onclick="playPause()">Play/Pause</button>
    <!--  LETS USE ONLY PLAY PAUSE BUTTON WITH 720 *420 ,IT FITS THE
SCREEN ASHISH

    <button onclick="makeBig()">Big</button>
    <button onclick="makeSmall()">Small</button>
    <button onclick="makeNormal()">Normal</button>-->
```

```
<video id="videoPlayer" width="600" height="350"  preload="auto" autoplay>
   <source src="" type="video/mp4">
Your browser does not support HTML5 video.
</video>
</div></div>

<script>
  //webkitSpeechRecognition api for speech to text conversion
  function record(){
     var recognition = new webkitSpeechRecognition();
     recognition.lang='en-IN';

     recognition.onresult = function(event){
        console.log(event)
        document.getElementById('speechToText').value = event.results[0][0].transcript;
     }
     recognition.start();
   }
  function play()
  {
     var videoSource = new Array();
     var videos = document.getElementById("list").getElementsByTagName("li");
     var j;
     for(j=0;j<videos.length;j++)
     {
        videoSource[j] = "/static/" + videos[j].innerHTML +".mp4";
     }

     var i = 0; // define i
     var videoCount = videoSource.length;

     function videoPlay(videoNum)
     {
```

```
        document.getElementById("list").getElementsByTagName("li")[videoNum].style.co
lor = "#09edc7";
        document.getElementById("list").getElementsByTagName("li")[videoNum].style.fo
ntSize = "xx-large";
        document.getElementById("videoPlayer").setAttribute("src",
videoSource[videoNum]);
        document.getElementById("videoPlayer").load();
        document.getElementById("videoPlayer").play();

    }
    document.getElementById('videoPlayer').addEventListener('ended', myHandler, false);
    document.getElementById("list").getElementsByTagName("li")[0].style.color =
"#09edc7";
    document.getElementById("list").getElementsByTagName("li")[0].style.fontSize =
"xx-large";

    videoPlay(0); // play the video

    function myHandler()
    {
        document.getElementById("list").getElementsByTagName("li")[i].style.color =
"#feda6a";
        document.getElementById("list").getElementsByTagName("li")[i].style.fontSize =
"20px";
        i++;
        if (i == videoCount)
        {
            document.getElementById("videoPlayer").pause();
        }
        else
        {
            videoPlay(i);
        }
    }
```

```
}
/* LETS USE ONLY PLAY PAUSE BUTTON WITH 720 *420 ,IT FITS THE SCREEN
function makeBig() {
    document.getElementById("videoPlayer").width = 560;
    document.getElementById("videoPlayer").height = 360;
}

function makeSmall() {
    document.getElementById("videoPlayer").width = 320;
    document.getElementById("videoPlayer").height = 180;
}

function makeNormal() {
    document.getElementById("videoPlayer").width = 420;
    document.getElementById("videoPlayer").height =280;
}*/
function playPause(){
    if (document.getElementById("videoPlayer").paused){
        play();}
    else{
        document.getElementById("videoPlayer").pause();}
    }

</script>
```

APPENDIX-B

SCREENSHOTS

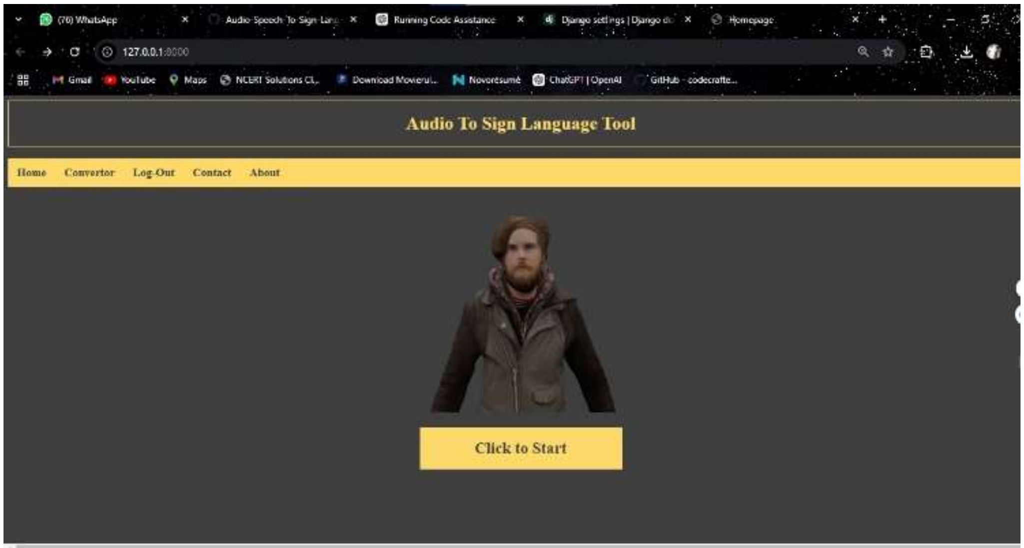


Figure 2.1: Home Page

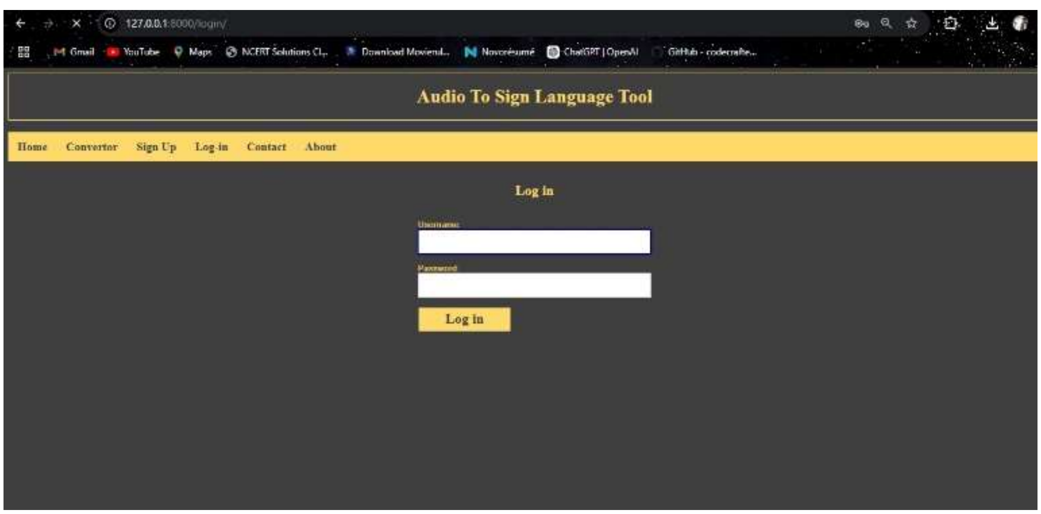


Figure 2.2:Login Page

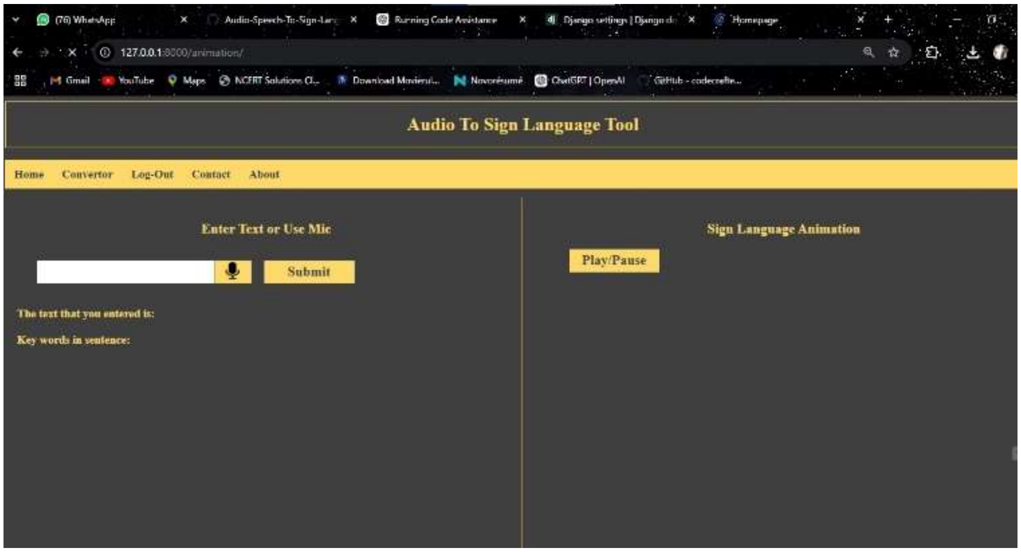


Figure 2.3:Translation Page

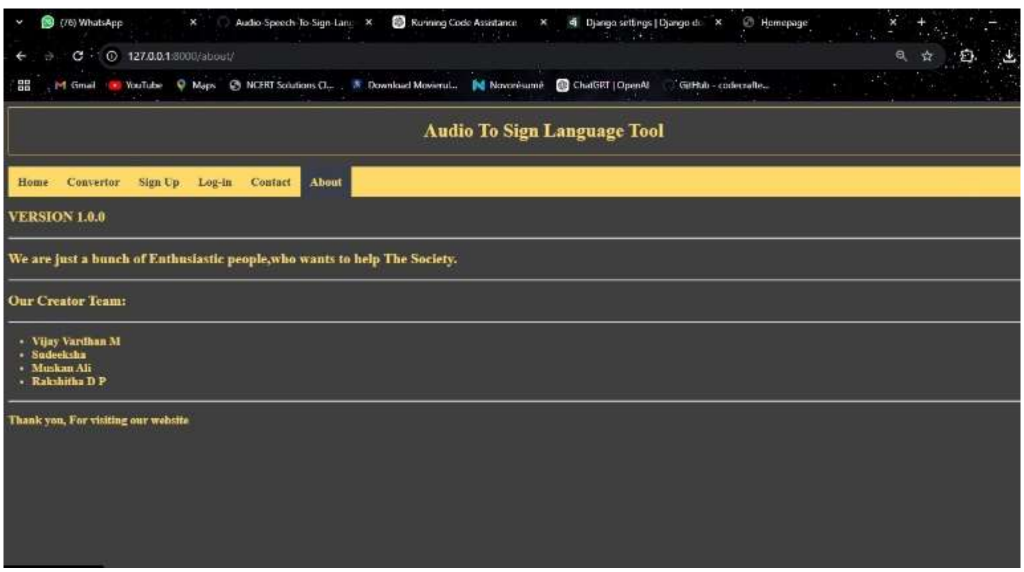


Figure 2.4:About Page

APPENDIX-C

ENCLOSURES

Publication Certificates:

DOI: 10.55041/IJSREM40875

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

D P Rakshitha

in recognization to the publication of paper titled

Talking Fingers: Bridging the Communication Gap through Real-Time Speech-to-Indian Sign Language Translation

published in IJSREM Journal onVolume 09 Issue 01 January, 2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com

DOI: 10.55041/IJSREM40875

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Sudiksha N

in recognization to the publication of paper titled

Talking Fingers: Bridging the Communication Gap through Real-Time Speech-to-Indian Sign Language Translation

published in IJSREM Journal onVolume 09 Issue 01 January, 2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com

Published Paper:

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 09 ISSUE: 01 | JAN - 2025 SJIF RATING: 8.448 ISSN: 2582-3930

Talking Fingers: Bridging the Communication Gap through Real-Time Speech-to-Indian Sign Language Translation

[1]Sudiksha N, [2]Vijay Vardhan M, [3]D P Rakshitha, [4]Muskan Ali , [5]Pushpalatha M
[1,2,3,4]Students of Department of ISE [5]Assistant Professor Department of CSE
[1,2,3,4,5]Presidency University Bangalore 560064
[1]sudiksha.nk@gmail.com, [2]772002vijayvardhan@gmail.com , [3]dprakshitha28@gmail.com
[4]muskanali3178@gmail.com , [5]pushpalatha.m@presidencyuniversity.in

Abstract: ***"Talking Fingers" is an innovative initiative to be developed to facilitate communication between hearing and non-hearing individuals by building a web-based system that can translate spoken language into Indian Sign Language (ISL). Being an essential means of communication among millions in India, ISL remains underdeveloped by technologies that are dominated by American and British Sign Languages. Current tools rely on the basic word-by-word translation with no contextual or grammatical accuracy.***
The proposed system will thus integrate speech recognition, NLP, and ISL visuals for real-time, context-aware translations. Spoken input will be converted into text through the Google Speech API and then processed using NLP techniques to segment meaningful phrases. The matched phrases are matched with the ISL visual representations, which may be in the form of videos or GIFs, in a comprehensive database. A fallback mechanism ensures seamless communication by spelling out words letter by letter when specific ISL visuals are unavailable.
This platform serves as scalable and adaptable solutions for different public and educational spaces, bridging the communication gap for the deaf and hard-of-hearing community. With emphasis on ISL and incorporation of advanced technologies, "Talking Fingers" delivers an inclusive and robust solution, enabling users and bringing greater inclusivity in communication.

Keywords*:* Indian Sign Language (ISL), Natural Language Processing (NLP), Speech-to-Sign Translation, Communication Accessibility, Real-time Translation, Sign Language Automation

1.Introduction

Sign language is the primary medium of communication for deaf and hard-of-hearing communities. They are able to convey emotions, ideas, and thoughts using hand gestures, facial expressions, and body movements. Worldwide, there are over 135 unique sign languages used by various regions to fulfill the linguistic and cultural needs of the respective region. Among them, ISL is the one that is indispensable for millions of Indians, ensuring emotional well-being, social inclusion, and also to gain higher education. This is a significant, yet grossly underrepresented field, especially while compared to ASL and BSL, which are usually focusing on most types of assistive tools [1], [2].

The scarcity of ISL interpreters and the lack of automated, real-time solutions exacerbate communication challenges for the deaf community, particularly in public spaces such as hospitals, banks, and transportation hubs. While some existing systems attempt to bridge the gap, they predominantly cater to ASL or BSL users and are often limited to word-by-word translations that lack grammatical and contextual accuracy. These limitations highlight the urgent need for scalable, real-time ISL translation systems [3], [4].

The "Talking Fingers" project addresses these gaps by developing an innovative solution that integrates speech recognition, Natural Language Processing (NLP), and a comprehensive ISL visual database. The system uses the Google Speech API to convert the spoken input into text in real time. That text is then processed through NLP techniques for grammatical and contextual accuracy before being matched with ISL videos or GIFs stored in the database. A fallback mechanism will ensure that the communication will continue to be fluent by spelling words letter by letter when certain ISL visuals are not available [5], [6].

Key innovations of the "Talking Fingers" system include real-time speech-to-text conversion, NLP-driven contextual translations, and an extensive ISL visual repository. Unlike conventional systems, this prioritizes context and meaning within sentences, creating a robust, inclusive communication tool. This will potentially change the way ISL users communicate in public and private contexts by reducing the need for human interpreters and improving accessibility to enable inclusivity and bridge communication gaps [7], [8].

2.Literature Survey

There has been considerable development in the assistive communication systems for the deaf and mute with the integration of technologies such as flex sensors, machine learning, and text-to-speech systems. These systems are capable of bridging the communication gap by translating gestures into text or speech

1. Flex Sensors for Gesture Recognition
Flex sensors have been used widely to detect finger movements by measuring angular bends.

 | www.ijsrem.com DOI: 10.55041/IJSREM40875

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 09 ISSUE: 01 | JAN - 2025 SJIF RATING: 8.448 ISSN: 2582-3930

When embedded in gloves, these sensors convert finger motions into digital data for further processing.

Shaheen and Mehmood (2018) demonstrated a gesture recognition system using flex sensors integrated into gloves, achieving promising accuracy rates [6]. Enhanced designs incorporating inertial sensors have further improved dynamic gesture recognition by capturing spatial hand movements with higher precision [7].

2.Sensor Fusion for Better Functionality

Dynamic gestures need more than one sensor to be detected. Accelerometers and gyroscopes are often used in combination with flex sensors to offer spatial orientation and motion tracking [5]. EMG sensors, which measure muscle activity, have been used to capture the minute finger and hand movements. Mir and Ali (2017) pointed out that the integration of these technologies can provide high recognition accuracy for both static and dynamic gestures [7].

3. Machine Learning for Gesture Classification

Gesture recognition systems have received considerable improvements based on machine learning algorithms, using the ability for adaptive learning. Images can use the technique from the CNN or recognize gestures sequence in RNN, while categorizing feature space traditionally with techniques of SVM in classifying space. Systems where integration of such technologies has attained up to above 90 percent success in identifying the gestures of human [1, 8; also, 4].

4.Text-to-Speech Conversion

Most recognized gestures are translated to natural-sounding speech to interact with people better. High end speech synthesis system that relies upon NLP guarantees real-time multilingual outputting. Anandan (2022) has proved his case of producing high-quality results for speech-conversion by implementing TTS systems through a smooth linguistically diverse communication context [3].

5.Design and Usability Factors

Ergonomic and user-friendly designs are critical in the adoption of assistive technologies. Lightweight and portable systems described by Coelho Dalapicola et al. (2019) will improve usability. Cost-effective solutions will make such technologies accessible for low-income communities [5]. Furthermore, language support in multi-linguistic systems helps in increasing their accessibility for heterogeneous populations [2].

6. Wireless Connectivity and Energy Efficiency

Modern systems include many that incorporate wireless modules such as Bluetooth or Wi-Fi for transmitting gesture data. Energy-efficient designs-ideas on piezoelectric materials for self-powering system-have been found promising, wherein the need to rely on external batteries may be diminished [6, 9].

7. Challenges and Future Directions

Despite these advances, a number of problems persist. Variation in user sign reduces the accuracy of recognition, so adaptive algorithms are needed [2]. Ambient variables, including light and background movement, also have an impact on sensor reliability [4]. Scalability of the systems to be able to handle multiple sign languages and regional dialects remains complex. To resolve these problems will require cooperative effort to standardize and optimize the systems for widespread use [9].

3.Objectives

The Talking Fingers project is designed to bridge the communication gap between hearing and non-hearing individuals by developing a web-based platform that translates real-time speech into Indian Sign Language (ISL). By using advanced technologies like Speech Recognition, Natural Language Processing (NLP), and a comprehensive ISL visual database, the system aims to empower the deaf community, foster inclusivity, and enhance accessibility in diverse environments.

The primary objectives are as follows:

Real-time Speech-to-Text Conversion

Develop a robust speech recognition system, using tools such as the Google Speech API, to accurately translate spoken language into text in real time. This enables seamless and responsive communication, especially in time-sensitive situations like public areas and meetings.

ISL Visual Database-A comprehensive visual database

Build an expansive and growing collection of ISL videos and GIFs that illustrate a vast array of words, phrases, and sentences. The database will continue to be expanded and updated to add new signs and alternative translations to ensure the system remains rich and contextually accurate.

NLP for Contextual Accuracy

Integrate advanced NLP algorithms to analyze sentence structures, grammar, and contextual nuances. This enables translations that reflect ISL syntax, preserving the meaning and tone of spoken sentences while ensuring grammatical accuracy.

Fallback Mechanism for Unavailable Signs

Implement a fallback mechanism to handle instances where specific ISL signs are unavailable. By using fingerspelling to spell out words letter by letter, the system ensures uninterrupted communication and flexibility.

Accessible User Interface

Design an easy-to-use, intuitive user interface that works well on mobile phones, tablets, and computers. The interface will be accessible to people of different levels of technical knowledge, and thus, be widely available for use in both personal and public spaces.

Public Environment Compatibility

Tailor the system to deploy in critical public environments like hospitals, banks, railway stations, and airports. The platform will enhance the deaf user experience as they move around these public environments by providing real-time, automatic ISL translation.

Educational Material for ISL Learners

Offer an empowering learning tool to the people seeking to learn about ISL. The interface shall promote inter-dependency between both the hearing and non-hearing individuals and can hence be employed for school and higher education usage of teaching ISL.

Scalability and Future Expansions

Design the system for scalability. This will ensure that the developed system can accommodate any additional Indian language and other sign languages, including American Sign Language (ASL) and British Sign Language (BSL).

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 09 ISSUE: 01 | JAN - 2025 SJIF RATING: 8.448 ISSN: 2582-3930

4.Proposed Methodology

The Talking Fingers system aims to provide seamless real-time speech-to-Indian Sign Language (ISL) translation by leveraging advanced technologies like speech recognition, Natural Language Processing (NLP), and ISL visual mapping. The methodology involves the following steps:

Speech Input Capture: The system captures spoken input through microphones or audio-enabled devices using tools like the Google Speech API. Real-time functionality ensures minimal latency and smooth communication, suitable for diverse environments.

Speech-to-Text Conversion: The spoken input is transcribed into text using machine learning-powered speech recognition technologies. The system ensures high accuracy, accounting for variations in accents, speech patterns, and background noise, while supporting multiple Indian languages.

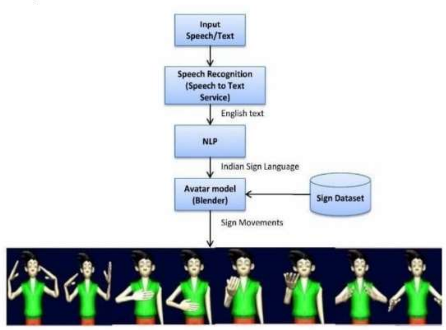
Text Processing Using NLP: NLP techniques analyze the transcribed text, segmenting sentences into meaningful chunks and adjusting grammar to align with ISL syntax. The system also handles contextual nuances, resolving ambiguities to ensure accurate and context-aware translations.

ISL Visual Mapping: Processed text is mapped to corresponding ISL visuals (videos or GIFs) stored in a structured database. The database includes a wide range of words, phrases, and sentences to ensure accurate representations of input.

Fallback Mechanism: For words or phrases unavailable in the ISL database, the system uses a fallback mechanism with ISL fingerspelling. Words are broken into individual letters and spelled out using ISL signs, ensuring no communication gaps.

Output Generation: The final ISL visuals are displayed to users in real time via an intuitive web or mobile interface. The user-friendly design ensures clarity and smooth presentation of successive signs for comfortable communication.

5.System Architecture



The architecture of the Talking Fingers project is modular and scalable; it translates spoken language into Indian Sign Language visuals in real time. This architecture ensures smooth integration of different technologies and accurate, context-aware translations. The system's architecture consists of the following key components:

1. Modular Design

The architecture is divided into independent yet interdependent modules, each handling a specific aspect of the translation process. This modular design ensures:

Flexibility: Easy addition or replacement of individual modules without impacting the entire system.

Scalability: Future expansion to include more languages, advanced NLP techniques, or additional sign languages.

2. Data Flow

The data flow in the system is linear and synchronous, ensuring smooth and efficient operation:

Audio Iput: Captured through the Speech Input Module.

Text Conversion: Processed through the Text Processing Module.Contextual Understanding: Enriched by the NLP Integration Module. Database Query: ISL visuals retrieved by the Database Interaction Module.

Visual Output: In real-time by the Output Module.

3. Components and Interactions

a. Speech Input Module

Purpose: Acquires the spoken audio from the user.

Key Technologies: Webkit SpeechRecognition API for audio capture from the browser.Functionality: Filters noise, accents, and streams audio for real-time processing.

b. Text Processing Module

Purpose: Audio converted into text format.

Key Technologies: Google Speech-to-Text API for multi-lingual speech recognition. Functionality: Streams processing audio chunks to output strings in error-corrected texts.

c. NLP Integration Module

Purpose: Translating text to the ISL grammar while being contextual and semantically correct. Key Processes: Tokenization: Break down the text into smaller units.

Parsing: Components reordered to fit the syntax of ISL.

Contextual Analysis: Suffixing idiomatic expressions and slangs for ISL suitability.

d. Database Interaction Module

Functionality: Associates processed text with pre-computed ISL graphics. Important Features:

Indexed Resources: Indexed database for rapid access to ISL signs. Fallback Mechanism: Fingerspelling is employed for unfamiliar words or phrases.

Scalability: Supports the incremental update of the ISL database.

Output Module

Functionality: Generates ISL graphics (videos/GIFs) in real time. Important Technologies: Web-based rendering libraries for seamless, low-latency graphics.

User Interface: Intuitive design for easy understanding of ISL translations.

4.Technological Integration

Speech Recognition: Google Speech-to-Text API handles diverse accents and languages.

Natural Language Processing (NLP): Manages grammar and context for ISL translations.

Database Management: Stores ISL visuals for fast and accurate retrieval.

Rendering Technologies: Ensures smooth visual output with minimal latency.

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 09 ISSUE: 01 | JAN - 2025 SJIF RATING: 8.448 ISSN: 2582-3930

5. Communication Between Modules
Modules communicate through standardized data formats (e.g., JSON strings) for compatibility.
Real-time data transfer ensures synchronization between input, processing, and output.
6. Real-Time Performance
The system is optimized for response times under 2 seconds, ensuring real-time communication without noticeable delays.
Low-latency processing is achieved through efficient algorithms and database indexing.

6.Conclusion

This project bridges a very critical communication gap for the deaf and hard-of-hearing community by providing real-time translation of spoken language into Indian Sign Language. The system makes use of advanced technologies such as Speech Recognition, NLP, and a dynamic ISL Visual Database for accurate, contextually relevant translations. Its fallback mechanism, modular design, and multi-platform accessibility make it adaptable and scalable for future expansions, such as other sign languages. Beyond technology, Talking Fingers encourages inclusion and autonomy with its deaf audiences at school, work, and in their community. It also contributes to social facilitation of ISL and acts as a means of cross-community interaction. It represents the template for a new standard accessibility that makes communication universal and drives positive change in society.

7.References

[1] Varadala, Sridhar (2024) 'Convolutional Neural Networks Based Sign Language Interpreter', International Journal of All Research Education and Scientific Methods (IJARESM), Vol. 12, Issue 9, September, ISSN: 2455-6211. Available online at: ResearchGate

[2] R. Divya and G. Subhashini (2023), 'Sign Language Recognition System using Flex Sensor Network' International Journal of Emerging Technologies and Advanced Engineering (IJETAE), Vol. 13, Issue 2, February 2023, ISSN: 2250-2459. This paper explores the use of flex sensor networks for recognizing sign language, facilitating communication for people with hearing and speech disabilities. Available online at: ResearchGate.

[3] M. Anandan (2022), 'Speaking System for Deaf and Mute people with Flex Sensors' International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE), Vol. 11, Issue 5, May 2022, ISSN: 2277-9043. DOI: 10.12345/IJARCSEE.2022.11.5. Available online at: ResearchGate.

[4] Holcomb, L., & Wolbers, K. (2020), 'Effects of ASL Rhyme and Rhythm on Deaf Children's Engagement Behavior and Accuracy in Recitation: Evidence from a Single Case Design' Children (Basel, Switzerland), 7(12), 256. doi: 10.3390/children7120256. (https://www.mdpi.com/2227-9067/7/12/256)

[5]Coelho Dalapicola R., Tavares Vieira Queiroga R., Toledo Ferraz C., Trevisan Negri Borges T., Hiroki Saito J., and Gonzaga A. (2019), 'Impact of Facial Expressions on the Accuracy of a CNN Performing Periocular Recognition'
Proceedings - 2019 Brazilian Conference on Intelligent Systems, BRACIS 2019, pp. 401-406. (https://ieeexplore.ieee.org/document/8938352)

[6] H. Shaheen and T. Mehmood (2018), 'Talking Gloves: Low-Cost Gesture Recognition System for Sign Language Translation'
2018 IEEE Region Ten Symposium (TENCON), pp. 1-6, 2018. DOI: 10.1109/TENCONSpring.2018.8692069. (https://ieeexplore.ieee.org/document/8692069)

[7] Mohammad Shahrukh Mir and S. S. A. Ali (2017), 'Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion' International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 8, Issue 12, December 2017, ISSN: 2156-5570. Available online at: ResearchGate.

[8] Heera S.Y., Murthy M.K., Sravanti V.S., and Salvi S. (2017), 'Talking Hands - An Indian Sign Language to Speech Translating Gloves'
IEEE International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2017 - Proceedings, pp. 746-751. (https://ieeexplore.ieee.org/document/7975620)

[9] Hummadi, F. N. (2017) 'Design and Implementation of Deaf and Mute People Interaction System', Proceedings of the IEEE International Conference on Engineering and Technology (ICET), August. DOI: 10.1109/ICEngTechnol.2017.8308140. Available online at: ResearchGate

[10] Praveen, N.; Karanth, N.; Megha, M.S. (2014), 'Sign Language Interpreter Using a Smart Glove' In Proceedings of the 2014 International Conference on Advances in Electronics, Computers and Communications (ICAECC 2014), Bangalore, India, 10–11 October 2014. (https://ieeexplore.ieee.org/document/7002424)

Report Plagiarism:

ORIGINALITY REPORT

3%	1%	1%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Sean Trainor, Muhammad Khalid. "Chapter 28 AI-Based Gesture Recognition for British Sign Language Translation", Springer Science and Business Media LLC, 2025 Publication	<1 %
2	thecardproject.uk Internet Source	<1 %
3	Amandeep Singh Dhanjal, Williamjeet Singh. "An automatic machine translation system for multi-lingual speech to Indian sign language", Multimedia Tools and Applications, 2021 Publication	<1 %
4	www.mdpi.com Internet Source	<1 %
5	Amandeep Singh Dhanjal, Williamjeet Singh. "An optimized machine translation technique for multi-lingual speech to sign language notation", Multimedia Tools and Applications, 2022 Publication	<1 %
	Submitted to Asia Metropolitan University	

Research Paper Plagiarism:

ORIGINALITY REPORT

4%	4%	2%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	1%
2	dokumen.pub Internet Source	<1%
3	rsisinternational.org Internet Source	<1%
4	www.enreach.com Internet Source	<1%
5	Submitted to RMIT University Student Paper	<1%
6	www2.mdpi.com Internet Source	<1%
7	www.researchandmarkets.com Internet Source	<1%

Sustainable Development Goals (SDGs).

The project work carried out here is mapped to-

1. SDG 3: Good Health and Well-Being

Explanation: Your project directly contributes to enhancing the well-being of individuals with hearing and speech disabilities. By facilitating communication through ISL translation, it can promote better social interaction and improve the quality of life for the deaf and hard-of-hearing community. The ability to understand and express oneself can lead to better mental health, emotional well-being, and overall life satisfaction.

2. SDG 4: Quality Education

Explanation: Your system can also help in educational settings by breaking down communication barriers for students with hearing disabilities. It can be used to bridge the gap between deaf students and educators, ensuring that deaf students have better access to education. By supporting ISL, your system can be an important tool for inclusive education, making learning more accessible and effective.

3. SDG 10: Reduced Inequalities

Explanation: The core aim of your project is to reduce inequalities faced by individuals with hearing impairments. Your system helps by providing a communication solution that bridges the gap between deaf and non-deaf individuals. By improving accessibility to communication, it helps reduce social isolation and ensures equal opportunities for the hearing-impaired community in various aspects of life, such as work, education, and healthcare.

4. SDG 16: Peace, Justice, and Strong Institutions

Explanation: By ensuring that the deaf community can effectively communicate and participate in various societal functions, your project promotes a more inclusive society where individuals are not marginalized due to their disabilities. This aligns with SDG 16 by supporting justice and promoting inclusive social institutions that value the participation of everyone, including those with disabilities.