Translating Sign Language to Speech(PCSE25-65)

PROJECT SYNOPSIS

BACHELOR OF TECHNOLOGY

Computer Science & Engineering

SUBMITTED BY:

Tanya Sharma 2100290100174

Srishti Upadhyay 2100290100166

Vikas Kumar 2100290100186

PCSE 25-65

August 2023

KIET Group of Institutions, Delhi-NCR, Ghaziabad (UP)

Department of Computer Science and Engineering

Table of Contents

S. NO.	CONTENTS	PAGES
1.	Introduction	1
2.	Objective	2
3.	Literature Review	3–4
4.	Feasibility Study	5
5.	Methodology	6
6.	Facilities Required	7
7.	Expected Outcome	8
8.	References	9
9.	Conclusion	9

Introduction

Project Overview:

In a world marked by diversity, effective communication lies at the heart of human interaction. However, communication barriers can often hinder the exchange of thoughts and ideas between individuals who use different forms of expression. Sign language is a visual and tactile form of communication used by individuals with hearing impairments. It employs gestures, hand movements, and facial expressions to convey meanings, making it a vibrant language. Nevertheless, most of the population lacks the proficiency to understand or interpret sign language, leading to challenges in effective communication and social integration for deaf individuals. The "Translating Sign Language to Speech" project aims to bridge the communication gap between individuals who use sign language and those who don't understand sign language. By leveraging computer vision, machine learning, and speech synthesis technologies, the project aims to enable real-time translation of sign language gestures into spoken language. This translation process not only facilitates communication between sign language users and the wider community but also empowers individuals with hearing impairments to participate fully in a conversation, education, healthcare, and various other aspects of life.

In the following sections, the project's objectives, methodologies, technologies employed, anticipated benefits, and challenges will be explored in depth. The "Translating Sign Language to Speech" project emerges as an embodiment of innovation driven by empathy and a commitment to making the world a more inclusive place for everyone, regardless of their preferred mode of communication.

Technologies Used:

- Computer Vision: OpenCV for gesture recognition and image processing.
- Machine Learning: TensorFlow or PyTorch for building and training the recognition model.
- Natural Language Processing: To map gestures to spoken language.
- Text-to-Speech Engines: Google Text-to-Speech, Amazon Polly, etc.
- User Interface: Mobile app or web-based interface for real-time interaction

Rationale:

1. **Inclusivity:** The project promotes inclusivity by enabling effective communication between sign language users and the general population.
2. **Education:** It can serve as an educational tool for learning sign language.
3. **Accessibility:** The system can be used in various contexts, such as education, healthcare, and public service

Objectives

1. Gesture Recognition:

   Develop a system that can accurately recognize and classify sign language gestures in real-time using computer vision techniques like image processing and machine learning.

2. Translation:

   Implement a translation mechanism that maps recognized sign language gestures to corresponding spoken words or sentences.

3. Speech Generation:

   Convert the translated text into audible speech using text-to-speech (TTS) technology, making the communication accessible to a wider audience.

4. Real-time Interaction:

   Design an intuitive user interface that facilitates real-time interaction between sign language users and the system.

5. Accuracy and Robustness:

   Ensure the system's accuracy in recognizing gestures across different lighting conditions, hand positions, and variations in gestures

Literature Review

Sr No.	Journal	Year	Technique	Findings	Shortcomings
1	A Novel Natural Language Processing (NLP)-Based Machine Translation Model for English to Pakistan Sign Language Translation	2020	NLP	Quantitative results reveal a very promising Bilingual Evaluation Understudy (BLEU) score of 0.78. ·Comparative analysis shows that our proposed system works well for simple sentences.	System works well for simple sentences but struggles to translate compound and compound complex sentences correctly.
2	EasyTalk: A translator for Sri Lankan sign language using machine learning and artificial intelligence.	2020	RCNN CNN NLP ML	The model detects at an accuracy rate of 91% for all given test scenarios. The model was also tested against live video and still was able to identify the gestures	For the moment, the system is proposed to be a web application and soon will be made into a mobile application with faster responses and lower processing time. Further, with the introduction of 5G, the response times will be faster
3	ATLASLang NMT: Arabic text language	2021	Artificial Neural	The average BLEU score of ATLASLang MTS is 0,37.	The training could be more efficient if the

	into Arabic sign language neural machine translation		Network Neural Machine Translation n (NMT)	ATLASLang NMT gave an average score of 0,79, which is much closer to the ideal score	dataset were expanded. The system uses a limited sign database
4	Utalk: Sri Lankan sign language converter mobile app using image processing and machine learning.	2020	CV ML	Utalk can perform well in both static and dynamic sign classification. Utalk achieves high precision and recall values (over 0.90) for all the static signs	Limited dataset
5	Recognition of Amharic sign language with Amharic alphabet signs using ANN and SVM	2021	ANN SVM	This paper presents a system that translates Amharic sign language into text using digital image processing and machine learning algorithms. •The system can recognize the Amharic alphabet signs with an average accuracy of 80.82% and 98.06%, respectively. The system has four main stages: image preprocessing, segmentation, feature extraction and classification.	This work could not work with words, phrases or sentences for the study of sign languages. This project could not develop a system which will work like a two-way communicator to translate sign to text and vice versa.
6	2-way Arabic Sign Language Translator using CNNLSTM Architecture and NLP	2020	Natural Language Processing (NLP)	The CNNLSTM architecture used for sign to text translation is especially ideal for this task as it works with	work is limited to translating solo dynamic words and phrases,

			Deep Learning Neural Network (DLNN) Convolutional Neural Network (CNN) Long Short-Term Memory (LSTM)	an RGB input from a regular smartphone camera. •The translator endows the deaf with a choice between the 'Deaf Culture' and 'Normal' culture [21]. Communication via the mobile device would allow the deaf to explore and interact with more places and people, thus allowing them to have more social experiences	The model can be connected to a cloud database which holds a crowdsourced gesture library, would ensure that the model is robust to the sociolinguistic changes affecting sign language. the model is a desktop application with still images as output, which limits its utility in real-time scenarios
7	Translating Speech to Indian Sign Language Using Natural Language Processing	2022	NLP	The system accepts audio and text as input and matches it with the videos present in the database created by the authors. If matched, it shows corresponding sign movements based on the grammar rules of Indian Sign Language as output, if not, it then goes through the processes of tokenization and lemmatization	The training could be more efficient if. the dataset was expanded. The system uses a limited sign database The features of the system could be enhanced by integrating reverse functionality
8	Sign Language Recognition Using Gesture Recognition and Natural Language	2021	Literature Survey	The paper has tried to understand and analyze the approaches of various kinds and the developments which	After capturing the video of SL sentences, The video will be broken down into

	Processing			have taken place to make appropriate gesture recognition of the singer. The peculiarities in trying to create a robust system and NLP techniques have also been looked up to generate complete sentences.	images and individual words will be recognized. A system will be developed for detecting ISL and converting the detected words into a grammatically correct common language sentence.

Feasibility Study

A feasibility study is essential to assess the viability of the "Translating Sign Language to Speech" project from technical, financial, and operational perspectives. It helps determine whether the project is achievable, sustainable, and aligned with its objectives.

Here's a comprehensive feasibility study of the project:

1. Technical Feasibility:

- Expertise: The project requires expertise in computer vision, machine learning, natural language processing, and software development.
- Technology: The availability of advanced technologies like image recognition libraries, machine learning frameworks, and text-to-speech engines supports the project's technical feasibility.
- Resources**:** Access to hardware resources, such as cameras and computing power, is necessary for real-time gesture recognition.

2. Financial Feasibility:

- Budget: Funding is required for software development, hardware procurement (e.g., cameras), and potential integration with third-party APIs for text-to-speech capabilities.
- Revenue Generation: The project could be monetized through licensing agreements with institutions, organizations, or individuals who require the software for educational or communication purposes.

3. Operational Feasibility:

- Ease of Use: An intuitive user interface and straightforward interaction are essential for the system's operational feasibility.

- Maintenance: Regular updates, bug fixes, and continuous improvement are required to ensure the system's operational reliability.

4. Market Feasibility:

- Target Audience: Individuals with hearing impairments, educational institutions, healthcare providers, and organizations working with the deaf community.
- Competition: The project's uniqueness lies in its ability to combine gesture recognition and speech synthesis, giving it a competitive edge.

5. Environmental Feasibility:

- Hardware Impact: The use of cameras and computing equipment should be environmentally conscious.
- Energy Consumption: The project should aim for energy-efficient implementation to minimize environmental impact.

Methodology

1. Data Collection:

   Collect a comprehensive dataset of sign language gestures, including different signs and their meanings.

2. Gesture Recognition Model:

   Train a machine learning model, possibly a Convolutional Neural Network (CNN), to recognize sign language gestures based on the collected dataset.

3. Translation Mapping:

   Create a mapping between recognized gestures and their corresponding spoken words or phrases.

4. Text-to-Speech Integration:

   Integrate a reliable text-to-speech engine to convert the translated text into natural-sounding speech.

5. User Interface:

   Develop a user-friendly application with a camera interface that captures gestures and displays the translated speech output.

6. Testing and Optimization:

   Test the system with real users, gather feedback, and optimize the gesture recognition model and translation accuracy.

Challenges:

1. Gesture Variability: Sign language gestures can vary significantly based on regional dialects and individual styles.
2. Real-time Processing: Achieving real-time recognition and translation to ensure seamless communication.
3. Model Robustness: Ensuring the model's robustness in recognizing gestures under diverse conditions.

Facilities Required

1. Development Environment:

- Computing Equipment: High-performance computers or workstations equipped with sufficient processing power and memory.
- Software Tools**:** Development environments, integrated development environments, machine learning libraries and image processing libraries.

2. Data Collection and Training:

- Cameras and Capture Devices**:** High-quality cameras or capture devices to record sign language gestures for training and testing the recognition model.
- Diverse Dataset: Access to a diverse and comprehensive dataset of sign language gestures to train and validate the machine learning model.

3. Testing and Validation:

- Test Subjects: Individuals proficient in sign language who can assist in testing and validating the system's accuracy and performance.

4. Integration and User Interaction:

- Hardware Interfaces: Interfaces for integrating cameras or video sources with the software application.
- User Interface Development Tools: Software tools for designing and developing the user interface, ensuring user-friendliness and accessibility.

5. Ethical and Legal Considerations:

- Data Privacy Measures: Policies and protocols for handling user data responsibly and ensuring user privacy.

- Intellectual Property Considerations: Legal guidance to address potential intellectual property concerns related to image recognition and speech synthesis technologies.

6. Documentation and Collaboration:

- Documentation Tools: Software for creating project documentation, user manuals, and technical specifications.

7. Testing and Deployment Platforms:

- Test Devices: Devices on which the application can be tested to ensure compatibility across different platforms (mobile devices, computers, etc.).
- Deployment Platforms: Platforms for deploying the final application, whether as a mobile app or a web application.

Expected Outcome

1. Real-Time Gesture Recognition: Developing an accurate and efficient gesture recognition system that can interpret sign language gestures captured in real-time through camera input.

2. Translation to Spoken Language: Implementing a translation mechanism that maps recognized sign language gestures to their corresponding spoken words or phrases, enabling effective communication between sign language users and those who don't understand sign language.

3. Natural-Sounding Speech Synthesis: Integrating a reliable and natural-sounding text-to-speech (TTS) engine to convert the translated text into audible speech, making the communication process seamless and meaningful.

4. User-Friendly Interface: Creating an intuitive and user-friendly interface, possibly in the form of a mobile app or web application, that allows users to interact with the system easily and access translated speech.

5. Inclusivity and Accessibility: Empowering individuals with hearing impairments to communicate with a wider audience, breaking down communication barriers and fostering inclusivity in various contexts, such as education, healthcare, and social interactions.

6. Educational Tool: Serving as an educational tool for learning sign language and fostering awareness about the challenges faced by individuals with hearing impairments.

7. Versatility and Application: Enabling the system to be used across different scenarios, including educational institutions, healthcare facilities, public services, and personal interactions.

8. Improved Quality of Life: Contributing to an improved quality of life for individuals with hearing impairments by providing them with a means to effectively communicate with the broader community.

9. Technological Innovation: Demonstrating the innovative application of computer vision, machine learning, and text-to-speech technologies to solve a real-world societal challenge.

10. Empathy and Inclusion: Promoting empathy and inclusivity by encouraging society to acknowledge and address the needs of individuals with disabilities.

Conclusion

The "Translating Sign Language to Speech" project aims to create a technology-driven solution that facilitates effective communication between sign language users and the wider community. By utilizing computer vision, machine learning, and text-to-speech technologies, the project addresses a crucial need for inclusivity and accessibility while offering a unique and impactful application of advanced technology.

References

- A novel natural language processing (NLP)—based machine translation model for English to Pakistan sign language translation. Cognit. Comput. 12, 748-765 (2020).
- EasyTalk: A translator for Sri Lankan sign language using machine learning and artificial intelligence.
- ATLASLang NMT: Arabic text language into Arabic sign language neural machine translation.
- Utalk: Sri Lankan sign language converter mobile app using image processing and machine learning.
- Deep learning methods for sign language translation.