**MIT School of Engineering**

**Department of Computer Science and Engineering**

**Project Synopsis**

**Group ID: LY-AIA 306**

**Project Title: Sign Language To Speech Recognition**

**Group Members:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Enrollment Number** | **Roll No.** | **Name of student** | **Email Id** | **Contact Number** |
| MITU21BTCS0177 | 2213721 | DARSH THAKUR | thakurdarsh21  @gmail.com | 7666248123 |
| MITU21BTCS0743 | 2213708 | YUVRAJ SINGH | yuvraj270404  @ gmail.com | 8084602004 |
| MITU21BTCS0652 | 2213490 | SYED SHARIQ | syedshariqpvt09  @gmail.com | 9356513926 |
| MITU21BTCS0388 | 2213698 | OMKAR PAWAR | omkarpawar131313@gmail.com | 9860053854 |

**Problem Statement:**

In today's interconnected world, communication between individuals from diverse linguistic backgrounds has become increasingly important. Communication barriers between individuals who use sign language and those who do not understand it pose significant challenges in daily life, limiting interactions in personal, social, and professional environments. While sign language is a powerful and expressive mode of communication, it is not universally understood, leading to exclusion and difficulties for deaf and non-verbal individuals in accessing essential services, education, and employment opportunities.

Current solutions, such as human interpreters or text-based alternatives, have limitations in terms of cost, availability, and scalability. Moreover, these approaches often lack the convenience and immediacy required for seamless real-time communication in dynamic environments. There is a critical need for an accessible, accurate, and real-time solution that bridges the communication gap between sign language users and non-signers, enabling inclusivity and independence.

This project aims to develop an intelligent system that translates sign language gestures into spoken words or phrases in real-time. Leveraging advancements in machine learning and computer vision, the system will capture hand and facial gestures using a camera, process the data through an algorithm trained on sign language datasets, and output corresponding spoken language through a speech synthesis module. By providing an intuitive and portable solution, this project seeks to empower sign language users, fostering greater accessibility, social integration, and equality.

**Abstract:**

Communication barriers between sign language users and non-signers pose significant challenges in fostering inclusivity and accessibility. This project aims to address these challenges by developing a real-time sign language to speech recognition system that converts sign language gestures into spoken words or phrases. The system employs advanced computer vision, deep learning, and gesture classification techniques to enable accurate recognition of hand and facial gestures.

Using a camera to capture the user's gestures, the system processes the input through a deep learning model trained on a comprehensive sign language dataset. The recognized gestures are then converted into speech using the Google Text-to-Speech (GTTS) library, ensuring a cost-effective, portable, and efficient solution for real-time communication.

This project focuses on bridging the communication gap between sign language users and non-signers, empowering individuals with hearing or speech impairments to interact seamlessly in personal, social, and professional settings. By eliminating the need for human interpreters and providing a scalable, automated alternative, this system promotes accessibility, independence, and social integration for sign language users.

**Literature Survey:**

**1. Introduction to Sign Language Recognition Systems**

Several researchers have worked on developing systems to bridge the communication gap between sign language users and non-signers. Early solutions often relied on wearable sensors or data gloves to capture hand movements. However, these methods were limited in accessibility and convenience due to the need for specialized hardware.

**2.** **Computer Vision for Gesture Recognition**

Advances in computer vision have significantly improved the accuracy and feasibility of gesture recognition systems. Studies highlight the use of cameras to capture hand gestures and facial expressions, eliminating the need for physical sensors. Algorithms like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been widely used for gesture classification. OpenCV-based real-time gesture tracking is also frequently cited for its efficiency.

**3. Deep Learning in Sign Language Recognition**

Deep learning has become a critical component in sign language recognition. Techniques like CNNs are commonly used for image and video analysis to identify hand movements, while Long Short-Term Memory (LSTM) networks help process sequences of gestures. Existing works, such as those using pre-trained models like YOLO, have demonstrated real-time performance with high accuracy.

**4. Gesture Classification Techniques**

Gesture classification is a fundamental task in sign language recognition. Recent studies emphasize the importance of using datasets that include a wide range of gestures with variations in lighting, angles, and skin tones. Machine learning models like Random Forests, Support Vector Machines (SVMs), and deep learning-based classifiers have been compared for their effectiveness in recognizing hand shapes and motion patterns.

**5. Speech Synthesis for Output**

The integration of speech synthesis has been explored in a few studies. The Google Text-to-Speech (GTTS) API is widely recognized for its ability to generate natural-sounding speech, making it a popular choice for converting recognized text into spoken language.

**6. Limitations in Existing Systems**

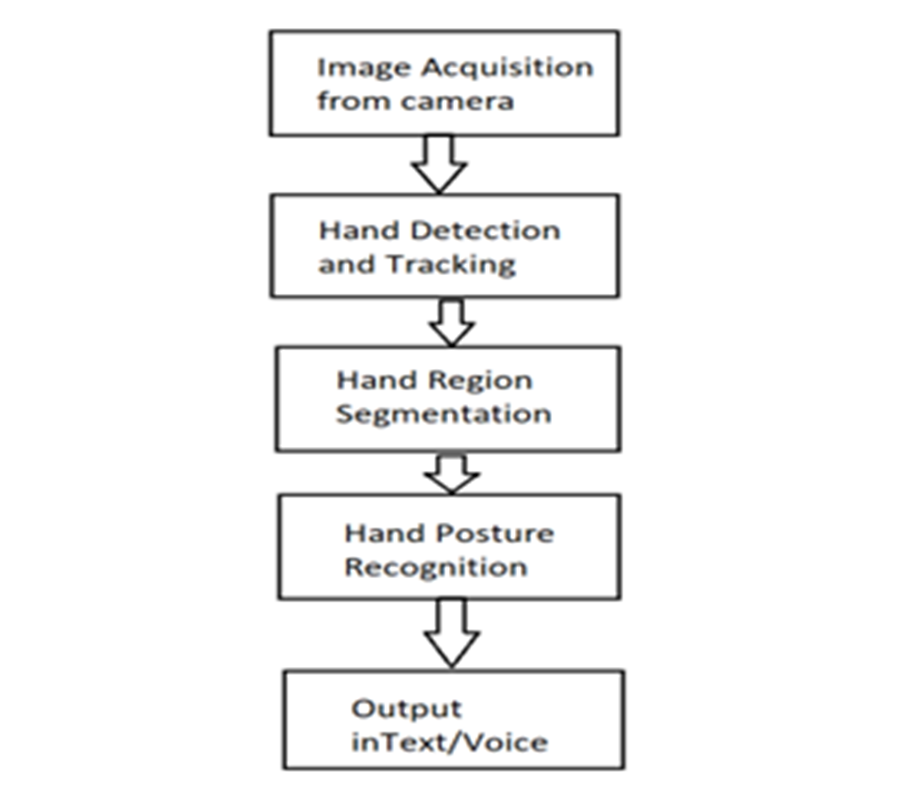
Despite advancements, many existing systems face challenges such as:

* Limited scalability to different sign languages.
* Difficulty in recognizing complex gestures or gestures performed at varying speeds.
* High computational requirements, making them unsuitable for portable or real-time use.

**7. Research Gap**

Current solutions either require expensive hardware, lack adaptability to multiple environments, or fail to provide real-time results. This project addresses these gaps by using a vision-based approach combined with deep learning and speech synthesis to create a cost-effective, portable, and efficient system.

**Proposed System (Block Diagram):**

****

**Conclusion:**

As we conclude this project, it is evident that the path forward inSign Language To Speech Recognition is filled with exciting opportunities. Our work, alongside the body of research we've surveyed, underscores the importance of continued innovation in this domain. The development of a real-time sign language to speech recognition system has the potential to significantly reduce communication barriers between sign language users and non-signers. By leveraging computer vision, deep learning, and gesture classification techniques, combined with Google Text-to-Speech (GTTS) for speech synthesis, the system provides a cost-effective, accurate, and portable solution for translating sign language gestures into spoken words or phrases.

This project demonstrates that advancements in machine learning and vision-based technologies can be effectively utilized to address real-world accessibility challenges. The system not only empowers individuals with hearing or speech impairments but also promotes inclusivity and independence by enabling seamless communication in diverse settings.

**References:**

1. A. Halder and A. Tayade (2021) "Real-time vernacular sign language recognition using media pipe and machine learning", Journal homepage, pp 7421, 2021, [online] Available: www.ijrpr.com

2. S. Goyal, I. Sharma and S. Sharma (2013) "Sign language recognition system for deaf and dumb people", International Journal of Engineering Research Technology, vol. 2, no. 4.

3. A. S. Nikam and A. G. Ambekar (2016) "Sign language recognition using image based hand gesture recognition techniques", 2016 Online International Conference on Green Engineering and Technologies (IC-GET), pp. 1-5.

4. Y. H. Chao, C. Y. Chen and L Chen, (2018) "American Sign Language Recognition Based on 3D Convolutional Neural Networks and Convolutional Recurrent Neural Networks", Journal of Healthcare Engineering 2018.

5. J. Yang, X. Zhang and C Xu, "RGB-D sign language recognition based on CNN and LSTM" in , IEEE, 2019.

**Annexure:**

**Annexure I: Form A-Title Approval (for offline mode)**

**Annexure II: Form B-Market and financial feasibility (verify from guide)**

**Annexure III: Literature survey paper or links**