CS371L - Artificial Intelligence

Real Time Sign Language Translation

Session: 2022 – 2026

Submitted by:

Muhammad Tayyab 2022-CS-133

Submitted To:

Mr. Numan Shafee

Department of Computer Science

University of Engineering and Technology Lahore Pakistan

1. Objective

The objective of this project is to build an AI-based system that can translate sign language into spoken language in real time using video input. This system aims to facilitate communication between deaf and hearing individuals by recognizing hand gestures and mapping them to words or phrases using machine learning models.

2. Background and Motivation

- Motivation: Communication barriers for individuals who rely on sign language create challenges in many settings, including education, healthcare, and everyday interactions.
- Proposed Solution: Leveraging recent advances in deep learning and computer vision, we aim to enable real-time translation of sign language into spoken language.
- Technological Foundation:
  - CNNs for gesture recognition to detect spatial patterns.
  - RNNs (LSTM/GRU) for modeling temporal dependencies in the sequence of gestures.
- Social Impact: This project aims to enhance accessibility and inclusivity, providing a solution for communication between deaf and hearing individuals.

3. Proposed Methodology

We will approach the project in several structured steps to ensure a smooth development process.

Data Collection

- Source: Publicly available sign language datasets (e.g., American Sign Language dataset).
- Custom Dataset: If needed, we will generate our own dataset using video recordings.
- Annotation: Each gesture will be labeled with its corresponding sign or phrase for model training.
- Data Augmentation: Techniques like image rotation, flipping, and scaling will be applied to increase dataset variability.

Algorithms/Models

- CNN (Convolutional Neural Networks):
  - CNNs will be used for feature extraction from individual video frames.
  - Pre-trained CNNs (e.g., InceptionV3, ResNet) can be fine-tuned for gesture classification.
- RNN (Recurrent Neural Networks):
  - RNNs (LSTM or GRU) will capture the sequential nature of gestures, enabling accurate interpretation of sign language phrases.
  - The RNN will handle temporal dependencies between gestures to form complete sentences.

System Workflow

- Video Input: Capture real-time video using a webcam.
- Hand Gesture Detection: Detect hand gestures using OpenCV and MediaPipe.
- Gesture Classification: Use CNN to classify individual frames.
- Temporal Modeling: Use RNN to analyze the sequence of gestures for full phrase translation.
- Speech Output: Convert the recognized sign language into spoken words using a text-to-speech system.

Flow Diagram

Below is a placeholder for a flow diagram that illustrates the system architecture:

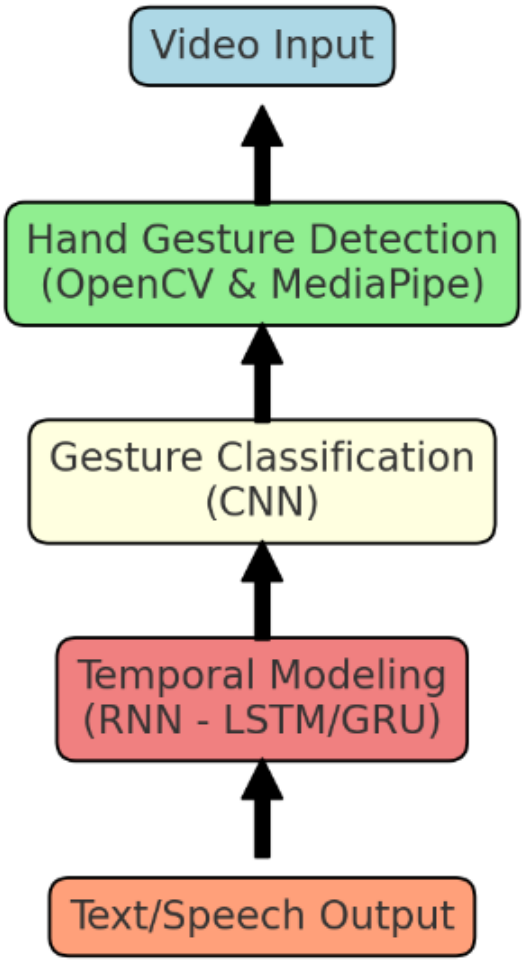

Figure 1: System Architecture Flow Diagram

Tools and Technologies

- Python: The primary programming language for implementation.
- Libraries:
  - `TensorFlow` and `Keras` for model development.
  - `OpenCV` for real-time video processing.
  - `MediaPipe` for hand tracking and gesture detection.
- Hardware: A standard webcam for video input.

Evaluation Metrics

- Accuracy: Overall performance in correctly translating gestures into words.
- Precision and Recall: Measures of how well the model identifies correct gestures and avoids false positives.

- F1-Score: The harmonic mean of precision and recall, providing a balance between both.
- Latency: Time taken to translate gestures in real time.
- Word Error Rate (WER): Measures how often the recognized words deviate from the correct translation.

4. Expected Outcomes

- Goal: A working prototype capable of real-time sign language translation with high accuracy.
- Challenges:
  - Variability in signing speed and gestures.
  - Regional variations in sign languages.
- Potential Applications:
  - Healthcare communication between doctors and deaf patients.
  - Educational tools for sign language learners.
  - Public service accessibility improvements.

5. References

- Intelligent Predictive Model for Hepatitis C, 2023 [**?**].
- Cleft Prediction Before Birth Using Deep Neural Networks, 2020 [**?**].
- Deep Learning for Sign Language Recognition Using CNN and RNN, 2022 [**?**].
- MediaPipe Documentation on Hand Detection, 2021 [**?**].
- Real-Time Gesture Recognition with OpenCV, 2020 [**?**].

6. Conclusion

By the end of this project, we aim to demonstrate a fully functional prototype capable of real-time sign language translation. This solution will contribute to improving communication for individuals who use sign language, fostering greater inclusion in society.