Real-Time Sign Language Recognition System

– Amrutha Girishkumar

Introduction

Communication is a fundamental aspect of human interaction, and for individuals with hearing or speech impairments, sign language serves as a vital medium. However, not everyone understands sign language, leading to a communication gap. To bridge this gap, we present a real-time sign language recognition system that translates hand gestures into text and speech. This application is designed to assist in inclusive communication using computer vision and machine learning techniques.

Abstract

This project presents a real-time American Sign Language (ASL) recognition system that converts hand gestures into text and speech. Utilizing webcam input and MediaPipe's hand tracking, the system predicts letters accurately using a trained machine learning model. A user-friendly GUI developed with Tkinter allows real-time prediction, sentence formation, gesture guide, and text-to-speech output. The model achieves high accuracy and responsiveness, making it suitable for communication assistance tools.

Tools Used

- Python – Core programming language
- OpenCV – For capturing and processing webcam input
- MediaPipe – For real-time hand landmark detection
- Scikit-learn – For training and deploying the MLP classifier
- Tkinter – For developing the graphical user interface
- pyttsx3 – For text-to-speech functionality
- Pillow – For handling and resizing the gesture guide image
- NumPy / Pandas – For data processing

Steps Involved in Building the Project

1. Data Generation:
    - Captured 63 hand landmark features per frame using MediaPipe.
    - Stored labeled gesture data (A–Z) for training the model.
2. Model Training**:**
    - Preprocessed the landmark data using scaling.
    - Trained an MLPClassifier on the features to recognize letters.
    - Achieved high accuracy after optimizing model and feature filtering.
    - The model was trained using the ASL Alphabet dataset, which contains labeled images (upto 2000) for each alphabet (A–Z). Dataset Link**:** https://www.kaggle.com/datasets/grassknoted/asl-alphabet
3. Prediction Script:
    - Used a live webcam feed to extract real-time hand landmarks.

- Predicted letters using the trained model with confidence filtering.
- Applied prediction smoothing to reduce flickering using a buffer.
- Enabled key controls:
  - Enter to add letter
  - Space for space
  - Backspace to delete
  - C to clear
  - S to speak sentence
  - G to toggle gesture guide
  - Q to quit

4. GUI Development**:**
   - Built using Tkinter with a user-friendly and visually appealing interface.
   - Webcam and gesture guide displayed side-by-side with equal sizing.
   - Prediction and sentence labels styled with distinct fonts:
     - Comic Sans MS for Prediction
     - Times New Roman for Sentence
   - All controls accessible through buttons and keyboard shortcuts.
   - Included a toggle for dark/light theme and suppressed all unnecessary warnings.

Conclusion

The ASL recognition system successfully detects and translates static hand gestures into English letters and forms complete sentences in real time. Its responsive GUI, accurate model, and intuitive controls make it an effective assistive tool for communication and learning. This project can be expanded in the future to include dynamic gestures, word-level recognition, and multilingual output, improving accessibility and inclusivity.