Unobstructed Voice for Signers: A Real-Time Centered Sign Language Recognition System

Espinola, May (School: Obra D. Tompkins High School)

Sign language serves as the main means of communication for the Deaf and Hard-of-Hearing community, yet barriers persist for non-signers who either lack the practical exposure to master it or a cultural complacency towards learning, due to little applicability in their lives. This creates a challenge for the deaf community, forcing them to rely on tools and make sacrifices to accommodate the general population. With remote working on the rise and artificial intelligence becoming more integrated into education and various fields, we can use advancements, specifically in computer vision, to remove communication barriers for the deaf. For practical adoption, the system must function in real-time without significant delays. This project aims to build a real-time, scalable system for American Sign Language (ASL) recognition, capable of large-vocabulary recognition and translation capabilities at over 30 FPS for both word- and sentence-level signing. By integrating MediaPipe for hand keypoint tracking and a hybrid static-dynamic fusion approach with a lightweight transformer architecture like MobileViT, the system calculates static spatial features alongside temporal dynamics with the use of multi-scale attention mechanisms to ensure high accuracy and low latency. Key functionalities of the system include sign-to-text/speech conversion, multimodal fusion of manual (hand gestures) and non-manual (facial expressions, body posture) markers, and realtime feedback via text display or speech synthesis. These features bridge the gap between signers and non-signers, not only enhancing accessibility and inclusion for the Deaf and Hard-of-Hearing community but also setting a foundation for the future integration of assistive technologies in everyday communication.

Awards Won: