Realtime ASL Translation Using ML and Computer Vision

Tiptur, Enzo (School: Academies of Loudoun)

There are over 1 million ASL speakers in North America and it is claimed to be the third most studied language in the United States. The language barrier between people with impaired or no hearing and people who can hear is massive and makes communication significantly harder. What if there was a "google translate" esque app for ASL translation? Using Computer Vision and Machine Learning, these barriers can be broken and a bridge can be built allowing for communication. While there are other solutions for this problem like a physical translator or a glove-like device to translate, these solutions are both expensive and inconvenient. Using Google's Media Pipe framework, a real-time translation can be provided for simple ASL words/phrases as well as alphanumeric characters using just a smartphone camera. To do this, datasets for both words/phrases and alphanumeric characters had to be obtained. The alphanumeric characters are static and therefore can use images to train the neural network, but for the word/phrases, videos need to be used due to the dynamic nature of the signs. Using the SVM neural network an average accuracy of ~96% can be obtained allowing for comfortable ASL translation. Word/phrase accuracy is lower at ~78%, but shows a lot of promise with a revised dataset. A platform like this one allowing for quick and easy ASL translation could work wonders for people struggling to communicate with a close friend or family member or even help an ASL novice learn the language better.