Real-Time Sign Recognition for ASL to Text Conversion

Krishna Kumar, Nidhi (School: Olympia High School)

This project focused on the development of a model that maximizes accessibility by accurately and efficiently recognizing American Sign Language in real-time, converting it to text without any auxiliary devices. Although over 500,000 individuals in the US are ASL users, ASL interpreters are still scarce, expensive, and in high demand. ASL users typically employ a variety of technologies to communicate online, but these can be slow. Additionally, this technology is limited in day-to-day interactions, especially as populations are shifting to remote settings for work and other activities, which makes better interpreting methods increasingly needed. The initial prototype consisted of a single Convolutional Neural Network (CNN) model. However, a CNN only identifies 2D spatial features. It is only useful for image recognition, so while it can recognize still signs such as letters, it cannot recognize moving signs. An additional model is required for temporal recognition to sequence frames to identify moving signs. A Recurrent Neural Network (RNN) was included in the architecture for this purpose. The final model was a CNN-RNN structure that first sends the input through the Xception CNN model and then pushes the output of the CNN Model through an LSTM to obtain an RNN output. The CNN and RNN outputs were then merged through multiplication. The final model had an accuracy of 90.25%, validation accuracy of 73.06%, and a loss of 0.2177 using the categorical cross-entropy function.

Awards Won: