Efficient Methods of ASL Video Classification Using 3D CNNs

Colden Bobowick[1], Anushka Narayanan[1], and Kaitlyn Williams[1]

[1]Brown University

Abstract

This project addresses the challenges involved in creating an efficient model for classifying American Sign Language (ASL) videos. It demonstrates the limitations of the current dataset, including lack of diversity in signers and insufficient samples per class. The research contrasts the 3D convolution, 2+1D Convolution, and Convolutional LSTM models in terms of the efficiency-accuracy tradeoff. The findings suggest that, to create a functional model, a significantly more diverse and comprehensive dataset is required. Additionally, access to computational and memory resources are necessary to train over multiple epochs. The project emphasizes the need for greater representation within ASL signers and sets a foundation for future research in efficient ASL classification models.