Team 9 – daea

3/3/2022

Development Plan

After hitting multiple dead ends with our own implementation of a voice separation model, we have decided to leverage an open-source pre-built model that employs the ConvTasNet architecture for voice separation. The pre-trained models that we could find are limited, and only separate two unknown voices from a stream, but these models have the framework available for us to build off of and fit to the objectives that our sponsor set for us.

This framework will allow us to build a wrapper environment around the model. This wrapper will filter background noise from human speech, feed the human audio into the model, and then take the separated audio from the model and classify it by speaker. Once we have the wrapper environment set up, we can move forward to improving the implementation of the voice separation model by training a model that follows the ConvTasNet architecture with a customized audio data loader. Creating a custom data loader made to fit the data set we have been processing would be very time consuming and is unclear if it would work, but once the wrapper environment is set up we would have the freedom to test it.

To properly implement this voice separation model, we will be utilizing the on-board microphone, passing audio files through the MATLAB filter as well as our wrapper environment, and then playing back the separated audio sources all from the Google Coral Dev Board. Alongside creating the implementation of the wrapper environment, we will be making sure the board has the proper audio pipeline set up to demonstrate our model.