**Methodology**

(little intro)

**Data Collection**

The dataset used in model training, testing and validation is LibriCSS (cite), a corpus specifically designed to evaluate continuous speech separation systems. It consists of synthetic mixtures generated from the LibriSpeech corpus (cite). LibriCSS contains 682 distinct speakers, whose clips are concatenated at various levels of overlap.

(a table showing the levels of overlap and what they mean)

**Data Pre-processing**

Trimming the audio, fetching labels

**Model Architecture**

Creating the CNN – why I chose 3 vs 4 conv layers

**HP Tuning with Optuna**

Why I used Optuna, applying a variation of SH over 100 tests instead of (the max amount), due to time restrictions.

**Pushing to production/real-time app that listens for audio and predicts speakers**