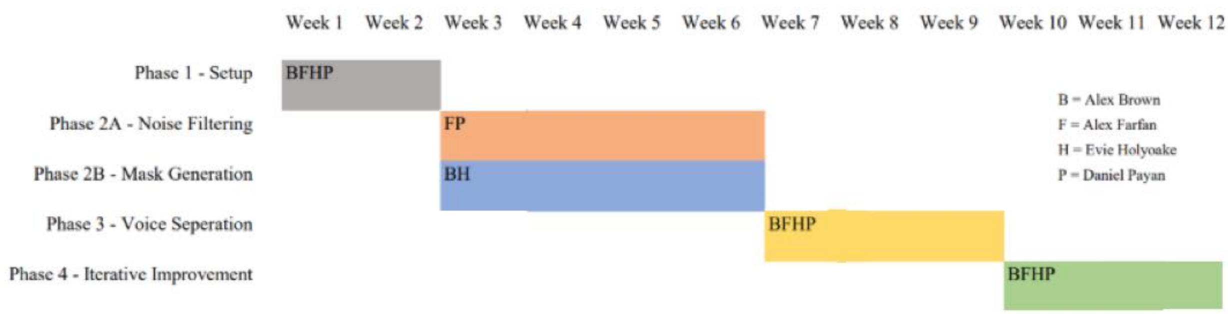
Alex Brown, Alex Farfan, Evie Holyoake, Daniel Payan

Team daea

Bi-Weekly Progress Meeting 3

3/22/2022

Gantt Chart:



Phase 1 - Setup:

- Set up workspace/environment
  - Version control using Git
  - Python3 development using Jupyter Notebook through in an Anaconda environment
  - Simulink and Matlab development using Matlab IDE
- Begin implementing initial filters
- Use Matlab to set up FIR/CIC filters to familiarize ourselves with noise reduction implementations
- Set up/Initialize Coral Board
  - Will need a similar development environment for testing of system implementation
  - Produce .wav files using on-board Digital PDM Microphones
- Research further into quantitative analysis system performance
  - As this is an investigative development, standards of performance may be subject to change
  - Performance may be relative across iterations, rather than compared to a known metric, as MFCC mask voice separation is a not well-explored domain

Phase 2A - Noise Filtering:

- Alex Farfan and Daniel Payan will lead this phase

- Attempt to implement separation of background noise from voices
    - Separate background noise from a single voice
    - Proceed to separate background noise from multiple voices
    - Proceed to separate background noise from stacked voices
- Implement MFCC feature vector production through Matlab
- Output MFCC feature vectors to file (for later use by Python3 code)
- Begin testing filter performance by comparing clean audio samples to artificial noise that was filtered by the program

Phase 2B - Mask Generation:

- Alex Brown and Evie Holyoake will lead this phase
- Develop infrastructure for interfacing with potential audio clip databases
- Implement MFCC feature vector production through Python3 (utilizing the Librosa library)
- Input MFCC feature vectors from file (previously output by Matlab code)
- Research into mask generation based on existing in generating masks in the Time-Frequency spectrogram domain
- Setup system to create an "enrollment profile" of audio clips for individuals from database sources
- Visualize MFCC time-domain data
- Implement a mask generation machine learning algorithm
    - Uses MFCC feature vectors produced from enrollment profiles
    - Produces a "mask" of multipliers/scalars that represent an individual's vocal presence in the MFCC-domain

Phase 3 - Voice Separation:

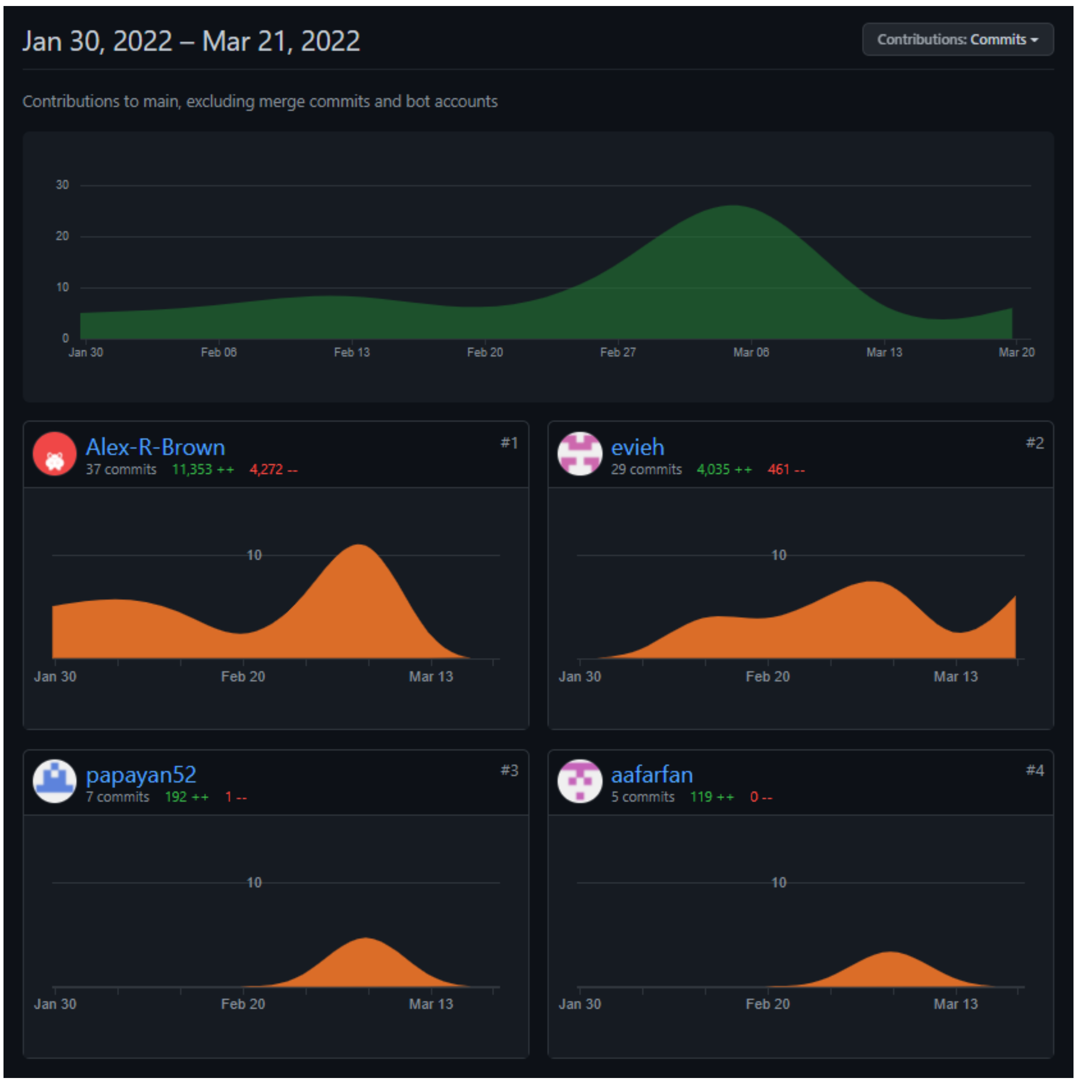
- Continue training mask generation model
- Use the MFCC feature masks to attempt to separate a known speaker from other speakers
    - Experimentally determine if MFCC feature mask multipliers are sufficient for attenuating non-target voices
    - Experimentally determine if a voice separation model is needed
- Reconstruct audio streams based on MFCC feature vectors with Python3 (utilizing the Librosa library)
- Implement the noise filtering algorithm on the Google Coral board, testing its performance

- Experimentally determine how to use both microphones concurrently whether this improves performance
- This must filter out noise recorded using the on-board microphone(s)
- Determine how to improve performance to meet latency requirements

Phase 4 - Iterative Improvement:

- Iteratively improve noise filtering implementation
- Iteratively improve mask generation model
- Iteratively improve voice separation implementation
- Iteratively improve testing mechanics

Git Repository:



1. Provide an update on your progress according to your Gantt chart in the proposal; provide over/under of your progress vs plan

We're in week 9 or Phase 3 of our Gantt chart.

After the mid-semester presentation, our team created a development plan (which has been uploaded with this document but was originally emailed to Professor Liu on 3/3). We have successfully implemented the wrapper that we proposed. This wrapper was presented to our sponsor, and our sponsor was understanding of our struggles to continue forward due to hardware limitations. This means we are in line with our development plan but have strayed away from our Gantt Chart.

The demo shown to our sponsor has been uploaded with this document.

2. Provide an update of teamwork: who is doing which tasks

- Alex Brown & Evie Holyoake
  - ✓ Interface with trained voice splitting model
  - ✓ Utilize functionality of noise reduce library
  - ✓ Normalize audio decibel levels of split audio streams
  - ✓ Creating custom data loader
  - o Training customer ConvTasNet architecture
- Alex Farfan & Daniel Payan
  - ✓ Interfaced with Coral Dev Board microphone and speaker output
  - ✓ Coral Dev Board environment set up
  - o Embedded environment for voice-splitting

3. If any new bottleneck/hurdle is discovered, please describe; also describe your solution to address the hurdle

- ☹ Hardware limitations
  - o We don't have the resources to train the unboxed ConvTasNet architecture
  - o We are working on setting up the environment on a more powerful PC but this option is still an uncertainty
- ☹ Google Coral Dev Board
  - o Issues with Google Coral Dev Board environment installations
    - ▪ Python 3.9
    - ▪ Librosa library
  - o We will be meeting with Professor Bell to see if he has any insight on how to work around the issues we are facing

4. If you are behind schedule please provide a plan to address

With the new development plan we are not behind schedule.

5. Describe your next step for the next 2 weeks

- Alex Brown & Evie Holyoake
    - Set up a new environment with the necessary resources to train with the ConvTasNet architecture
    - Adapt our randomized data preparation suite such that it creates metadata to interface with our custom PyTorch data loader object
    - Once this environment is set up, we will be working on iteratively improving our model
- Alex Farfan & Daniel Payan
    - Meet with Professor Bell to attempt to complete the Google Coral Dev Board's environment set up
    - Create a MATLAB CIC/FIR filter executable to run in front of the wrapper