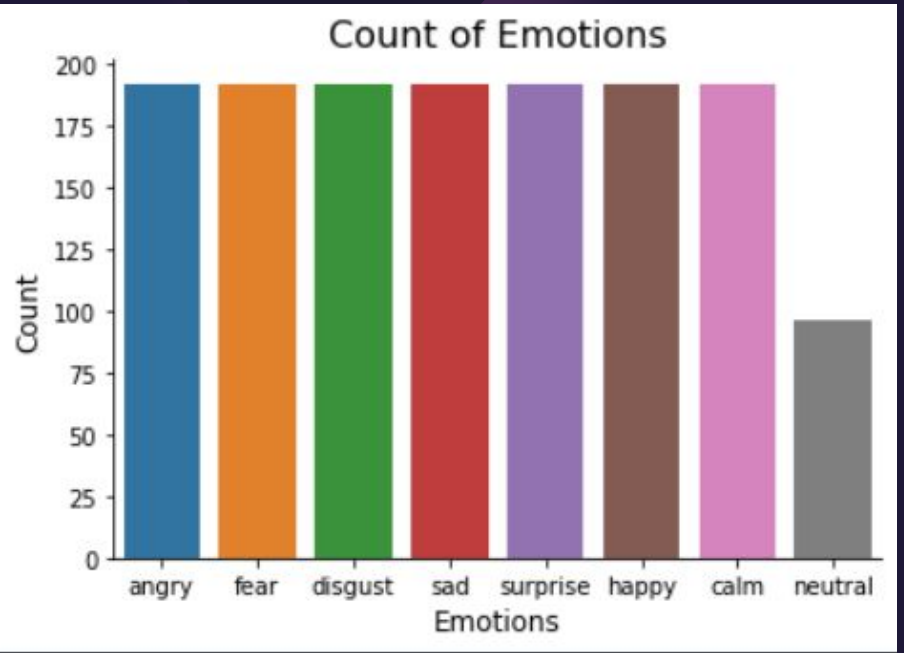
Emotion Detection Using Speech

By: Stephen William

The Problem

- Tools for Doctors and Therapists
- Classifying Emotions
- Assist in diagnostics
- Smart Watch App

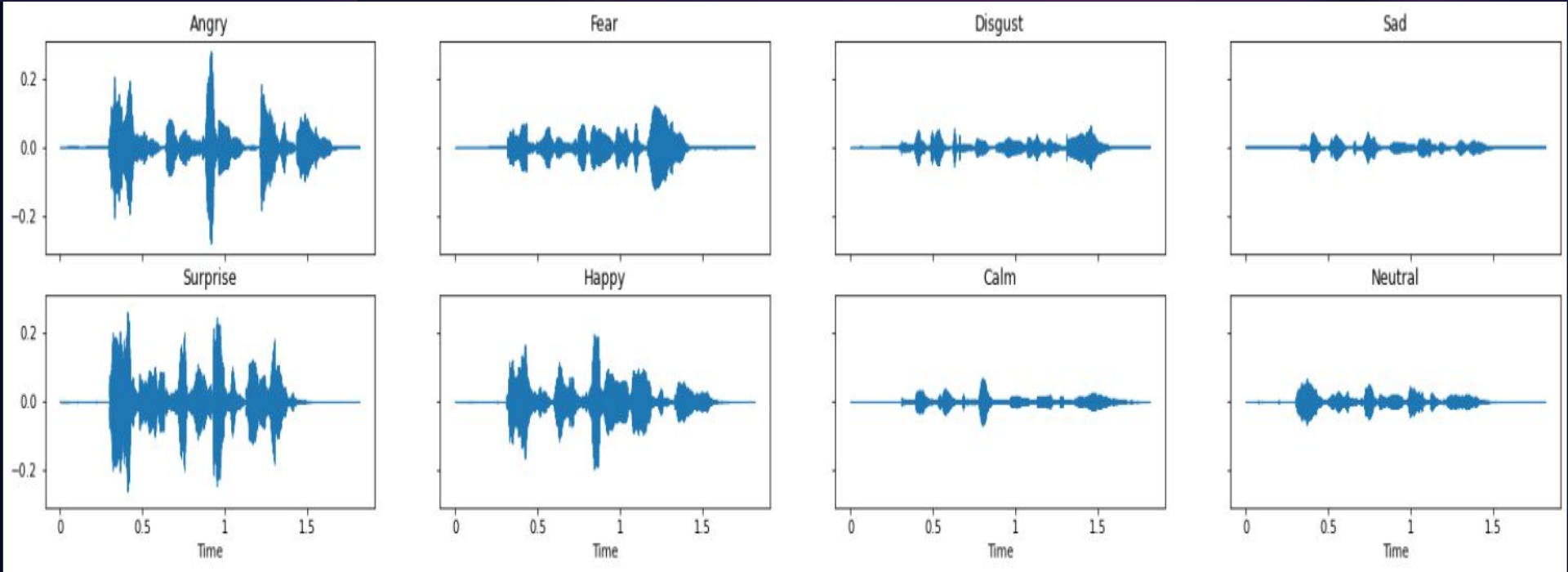
Data



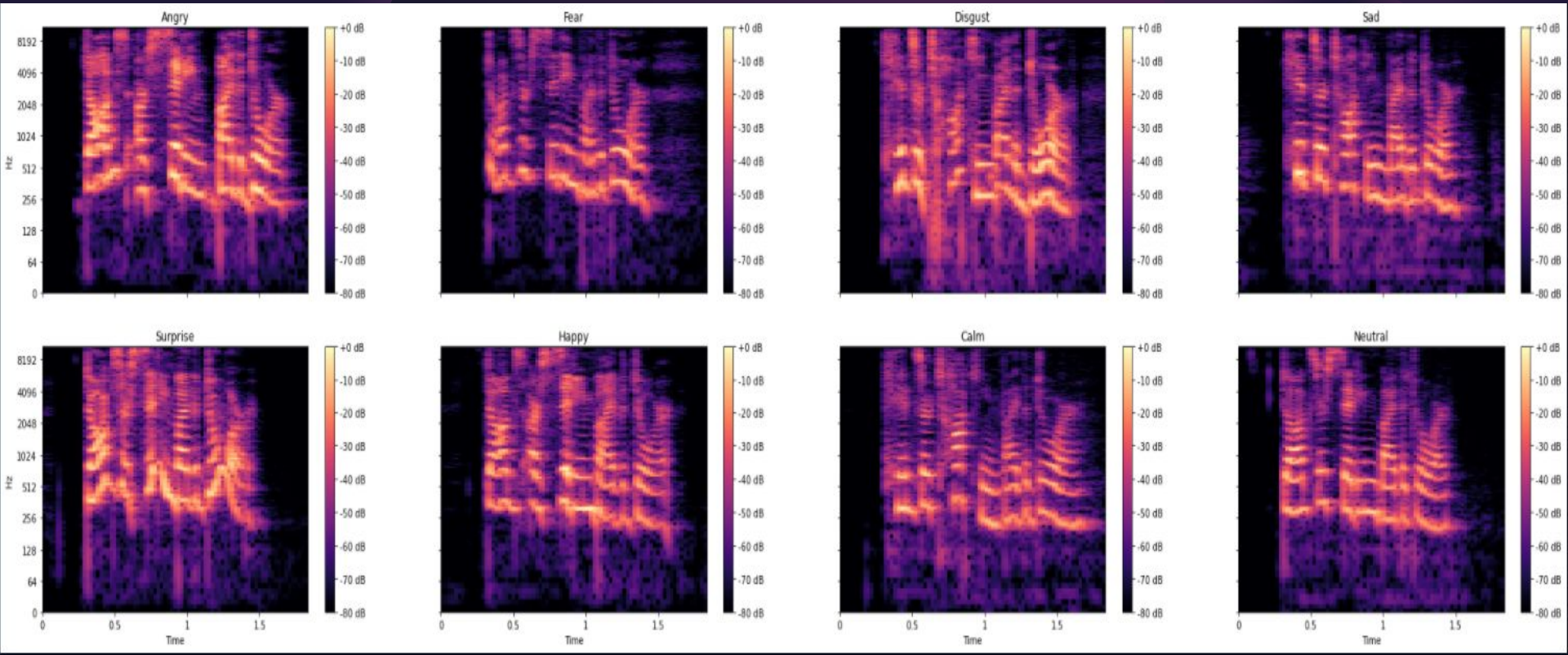
- RAVDESS
- 24 Actors/ Actresses evenly split
- 2 Statements Verbalized
- 2 levels of Intensity
- 8 Emotion Classes
- 1440 Observations
- 192 Per Class

Data Exploration

Emotions and Loudness

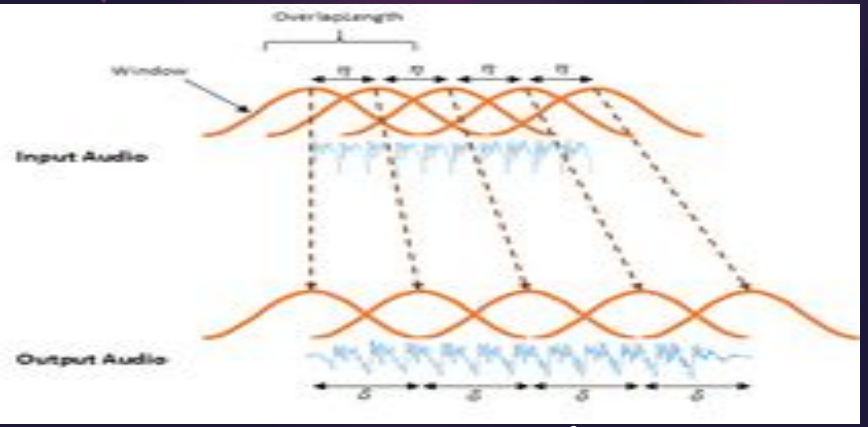


Melodic Content of Emotions

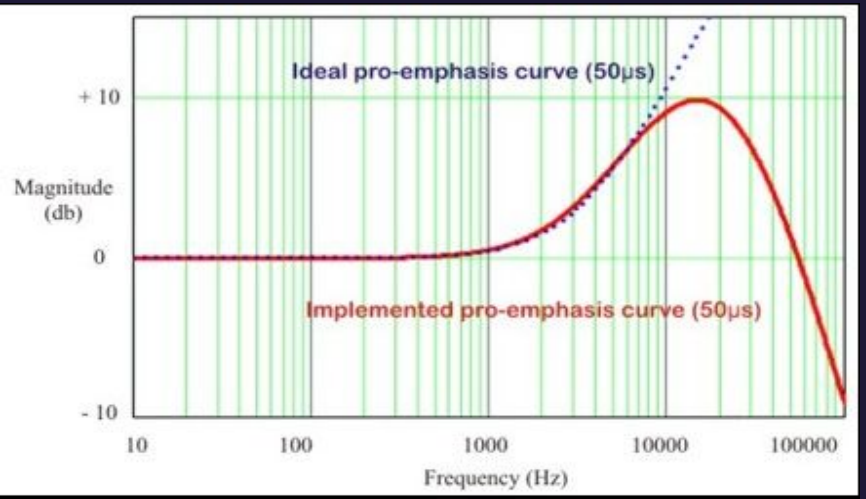


Preprocessing

Time/ Pitch Stretching

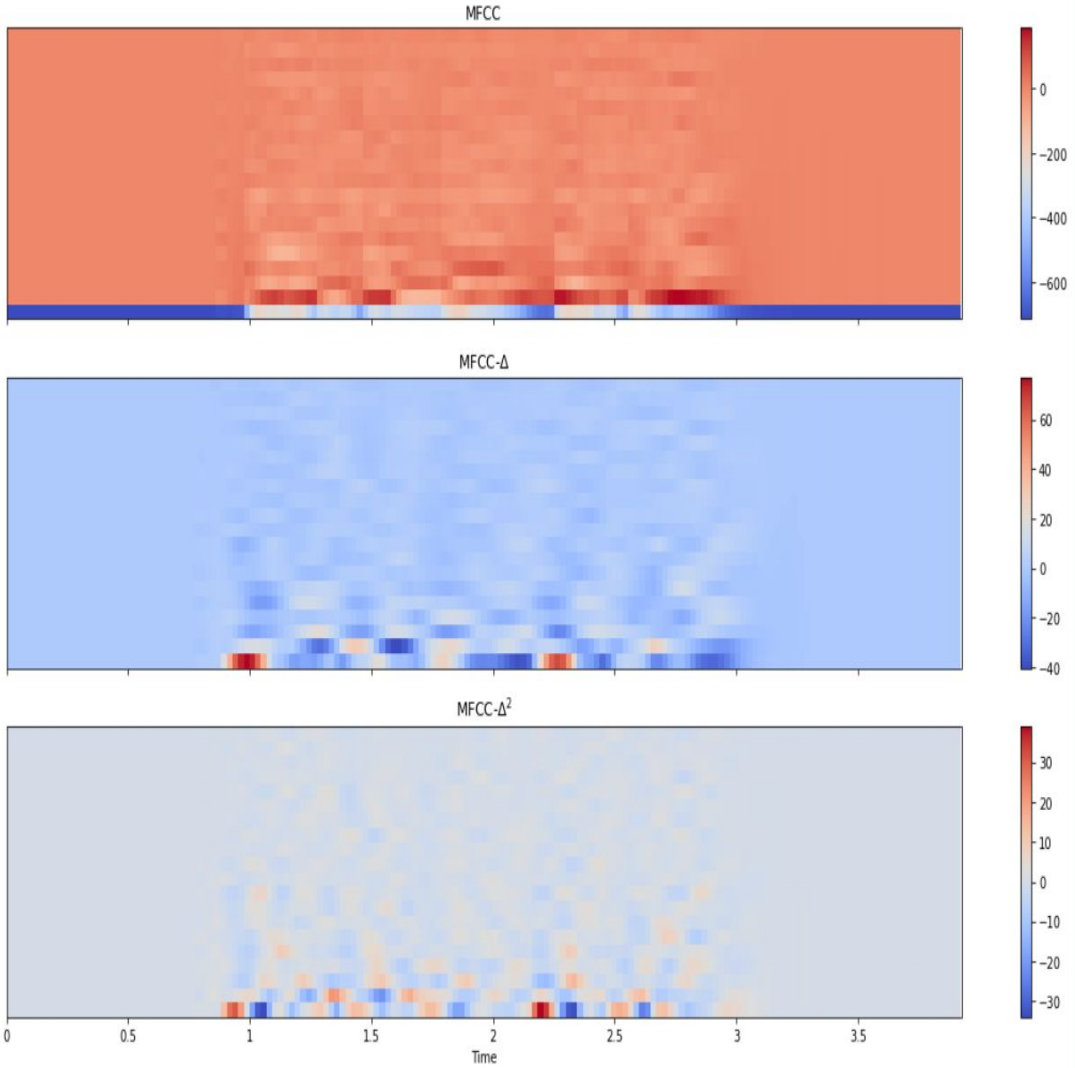


Pre Emphasis



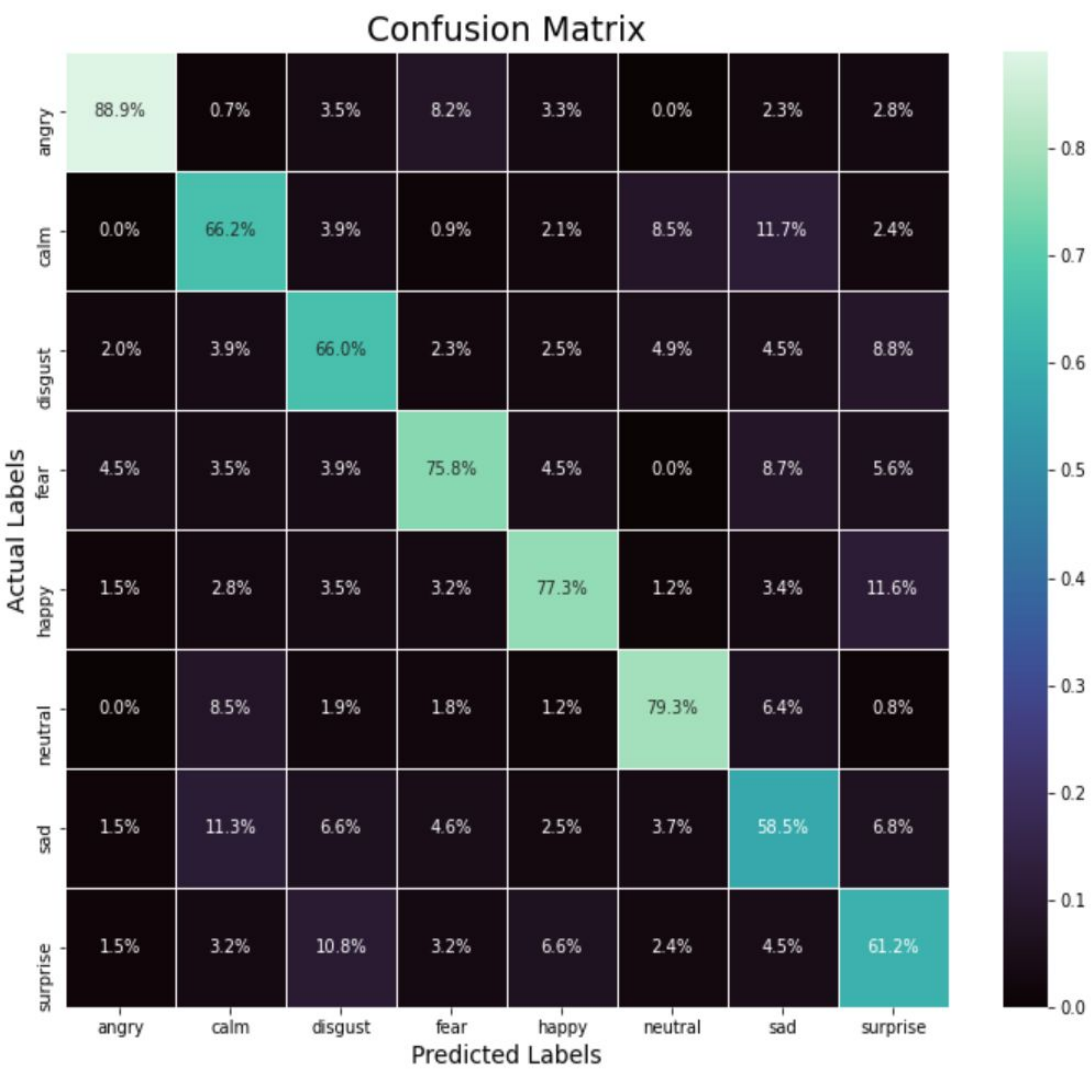
Synthetic data:

- Random Noise
- Time and pitch shift
- Pre Emphasis



Data Features

- MFCC
- MFCC DELTA’S
- DELTA-DELTA’S

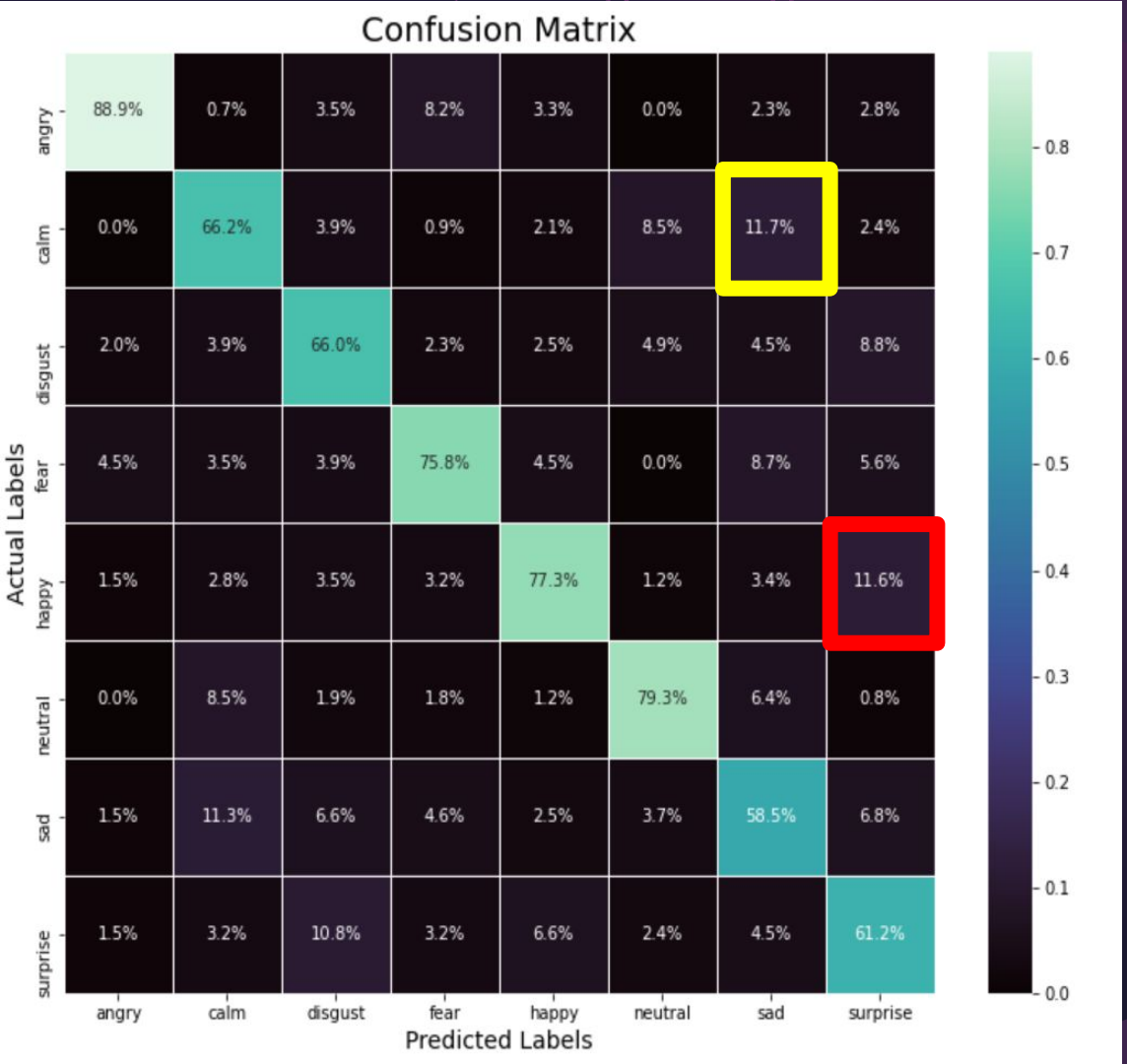


Results

70% Accuracy

Most successful detection:
Anger @ 88.9%

Least successful detection:
Sad @ 58.5%



Results

70% Accuracy

Sad for Calm: 11.7%
Surprised for Happy: 11.6%

Least successful detection: Sad @ 58.5%

Moving Forward

- More data
- Dropping the “Neutral” class
- Male Vs. Female voices
- More Data Augmentation

Thank You

Stephen William
ASR Data Science/ Audio Engineer

Swilliam.productions@gmail.com

https://github.com/Jyve00/