# Emotional Crowd Sound

Predicting the emotional affect within a crowd

Inspired by: Emotional sounds of crowds: spectrogram-based analysis using deep learning

# Emotional Affect

Emotional affect is a way to determine arousal of emotions based on what emotion it is and the level or degree of **intensity** of that emotion.

It's common to identify emotions in **individuals,** but what about **crowds**?

Crowds tend to use '**mirroring**' and **synchronization**,
but can also have **multiple different emotions**

Could we use this in real-time to perceive crowds 'likeness' in music or events for marketing, or maybe aiding police with predicting if a crowd may turn into a mob?

# Process

Collect spectrogram images
and research previous models

Base CNN model

Data analysis

Final Model/App
Deployment

Tune
parameters and
rerun model

# Spectrograms EDA

## 01

## Normalized

Audio files 20-20k Hz range (which are audible to humans)

## 02

## Filtered

Filtered out silence blocks

## 03

## Spectrograms

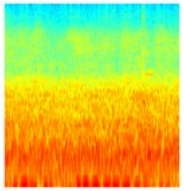
-spgrambw draw spectrogram function (MATLAB)
-png images using a 400 samples
- frame increment of 4.5 millisecond

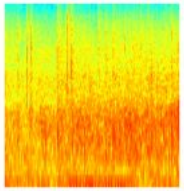
## 04

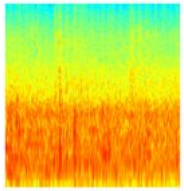
## Splitting

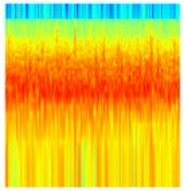
The data was heavily unbalanced with Neutral having over 5k images, Approval having around 3k images and Disapproval having just over 300

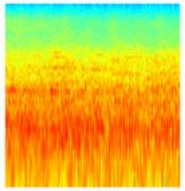
# What are Spectrograms?

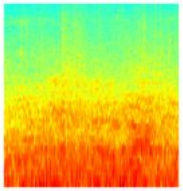

Neutral

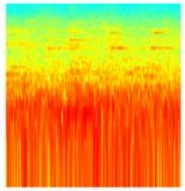

Approval

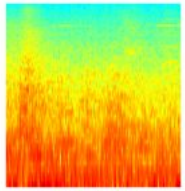

Approval


Approval


Disapproval


Disapproval


Disapproval


Disapproval

X axis = Time
Y axis = Frequency
Color Intensity = Amplitude

Amplitude can be interpreted as ‘loudness’ of the frequency

# Approaches to Models

## Their Approach

They trained an AlexNet with 4 epochs,L2 Regularization and were also getting a validation score at around 97% average over 4 networks

## Basic Sequential CNN

The model was overfit - added dropout layers
Control - only one dense layer at the end

## Stretched it

More nodes in the layers

## Compacted it

Compacting - using larger pool sizing to (then with less strides)

## Changing Strides

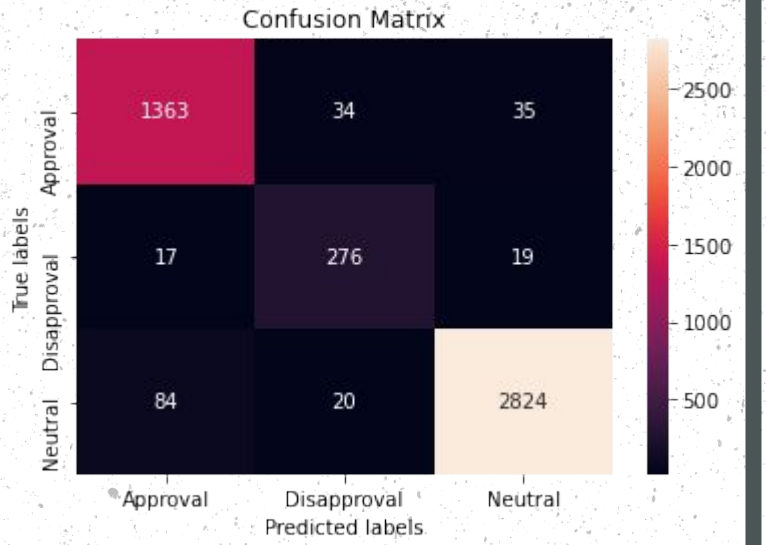
Overlapping the filters by taking less strides

# Best Results

Classification Report

| | precision | recall | support |
|---|---|---|---|
| Approval | 93% | 95% | 1432 |
| Disapproval | 84% | 88% | 312 |
| Neutral | 98% | 96% | 2928 |

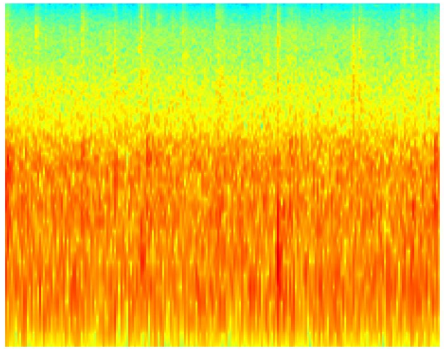
Average accuracy 96%




‘Compacted Sequential Model’

- larger pool size (4, 4) in one stretched out layer
- smaller strides (3)

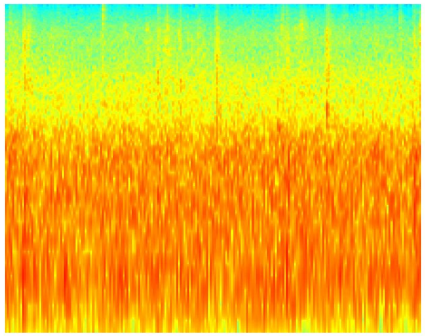
# Differences In Predicting Affect (0-3.5 kHz)



Approval

Predicted Neutral

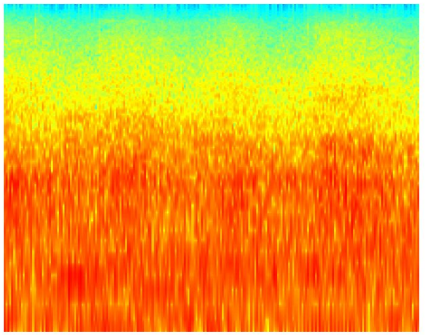
44% Approval
52% Neutral



Approval

Predicted Approval

49% Approval
47% Neutral

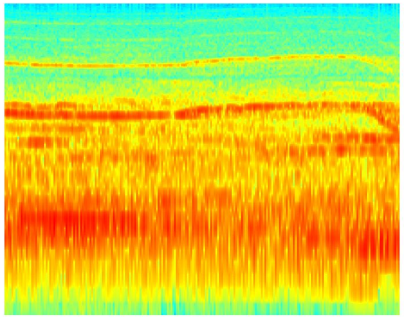


Neutral

Predicted Neutral

47% Approval
51% Neutral

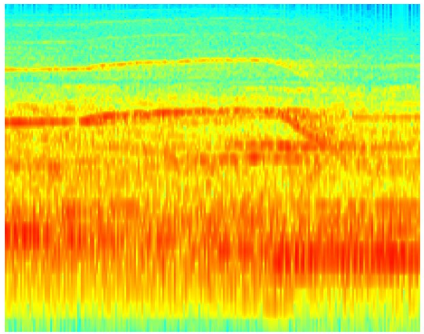
# Differences In Predicting Affect (0-3.5 kHz)



Disapproval

Predicted Disapproval

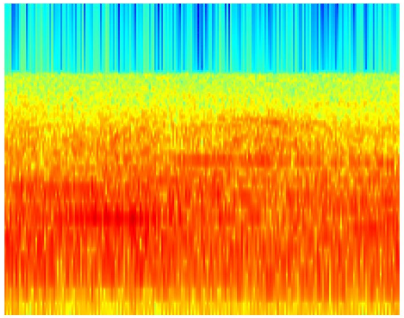
31% Approval
46% Disapproval



Disapproval

Predicted Approval

45% Approval
31% Disapproval

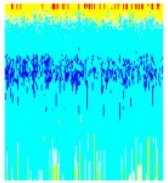


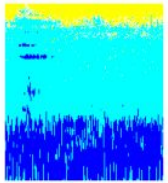
Disapproval

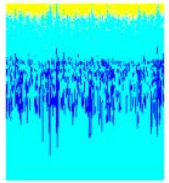
Predicted Neutral

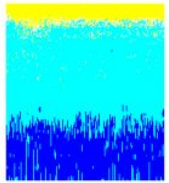
14% Disapproval
63% Neutral

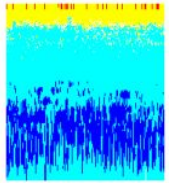
# Random Sampling (0-3.5 kHz)

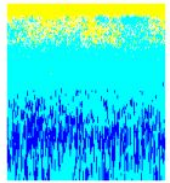

Approval

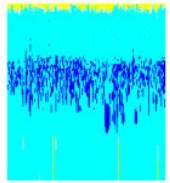

Neutral

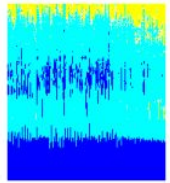

Approval


Disapproval


Disapproval


Disapproval

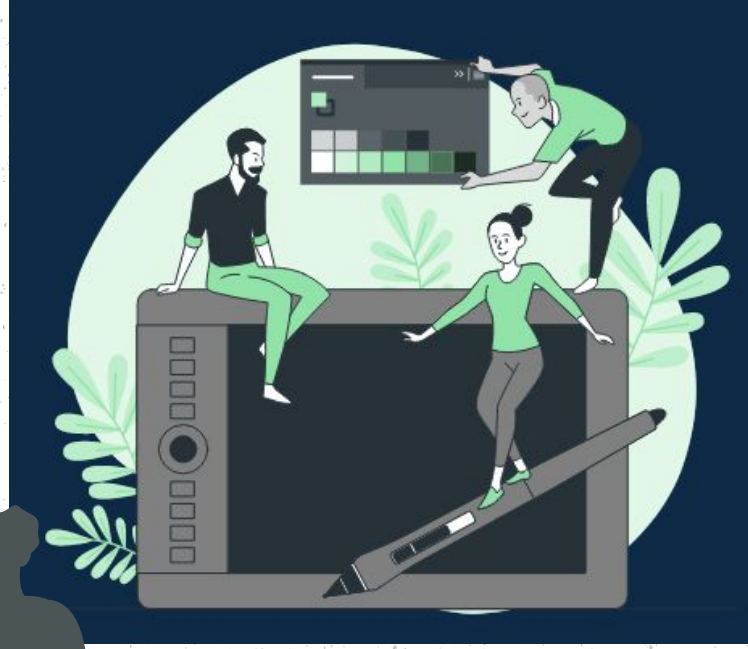

Approval


Approval

RGB (Red, Green, Blue Channel)
to
BGR (Blue, Green Red Channel)
^ i.e. Blue Dominate color

# Further Research

- Even though my process was to get to transfer learning, I did not however, my next steps would be to use Global Average Pooling Layers after transfer learning model
- If this research is going to be used in real time, I think it would be important to include and to train the model to know what silence is and not take them out.
- Also, as stated before, I would like to use Independent Component Analysis to aid in learning to distinguish and predict emotional affect within crowds where there are multiple different emotions.



# Thank You for Your Time

Tara Blackburn
941-960-5021
tara.blackburn3@gmail.com

GitHub

LinkedIn

Original dataset:
https://ieee-dataport.org/open-access/emotional-crowd-sound
Paper:
https://link.springer.com/article/10.1007/s11042-020-09428-x