Film Music Analysis

My goal with this project is to analyze the different ways that film use sound/music to tell stories and capture emotion. I decided to focus on modeling the colors, emotions, and instruments from films + their music. These are both classification problems that use two different data sets.

Emotion Classification

Dataset

Emotion Labeled Data that assigned different emotion labels to a collection of audio features. The possible labels are:

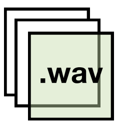
amazed-suprised sad-lonely quiet-still

angry-aggresive relaxing-calm happy-pleased

Model

Random Forest Classification

Predictions / Creating Data



audio files

centroid	mfcc_1	mfcc_0

dataframe of extracted audio features

*for this step I matched/normalized the same features used in the training data (for full breakdown see the notebook **audioFileToData.ipynb**)*

randomForest Classifier

trained model

relaxing-calm	amazed-suprised	angry-aggresive	quiet-still	sad-lonely	happy-pleased
1	0	0	1	0	0
0	1	0	0	0	1
0	0	0	0	1	0

predictions (1 if its predicted to have that emotion, 0 otherwise)

Instrument Classification

Dataset

MusicNet Dataset from Kaggle. Contains pre-split train/test data where each data point is an audio.wav file which has a corresponding labeled .csv file which has the timestamps+instruments+notes.

Model

OneVs.All + RandomForest Classifier

- multi class classification using the instruments as the target variable

yamnet model

- Loads an existing yamnet model which extracts audio features from each .wav form and classifies different audio events. These are used as features in the randomforest / one v all classifier.

note: *I found this notebook online, and used it as a guide for classifying my audio files.*

Predictions

The model takes in a .wav file and outputs an instrument label (these are encoded by numbers as seen below):

1: Piano
7: Harpsichord
41: Violin, fiddle
43: Cello
61: French horn
72: Clarinet
74: Flute
42: Violin, fiddle
44: Double bass
69: Wind instrument, woodwind instrument
71: Wind instrument, woodwind instrument

Testing

The model performed well on the test files from the MusicNet dataset. In general, it was good at predicting "piano", but sometimes struggled picking up woodwind instruments. I wonder if there is one loud instrument, it can overpower the other instruments making it more difficult to detect them.

```
Hamming Loss: 0.04
Classification Report:
                                     precision    recall  f1-score   support

                              Piano       0.92      0.94      0.93       179
                        Harpsichord       0.00      0.00      0.00         0
                     Violin, fiddle       0.79      0.58      0.67       118
                              Cello       0.00      0.00      0.00        22
                        French horn       0.96      0.68      0.80        73
                           Clarinet       0.00      0.00      0.00         0
                              Flute       0.00      0.00      0.00        40
                     Violin, fiddle       0.00      0.00      0.00         0
                        Double bass       1.00      0.04      0.08        23
Wind instrument, woodwind instrument       0.00      0.00      0.00        44
Wind instrument, woodwind instrument       0.00      0.00      0.00         0

                          micro avg       0.89      0.58      0.70       499
                          macro avg       0.33      0.20      0.23       499
                       weighted avg       0.71      0.58      0.61       499
                        samples avg       0.58      0.58      0.58       499
```

We see that the highest precision is from the piano, violin, french horn, and bass. Woodwind instruments, cellos, clarinets, all got more lost, but its possible that the training/testing data has less of these instruments

Color Analysis

For this aspect of the project, i wanted to explore different methods from extracting an images color pallette. I wanted unique colors not just the most common colors. For this, I explored using clustering and a color threshold and compared the results.

Clustering: I used Gaussian Clustering on 8 clusters and then selected the 5 most saturated cluster colors

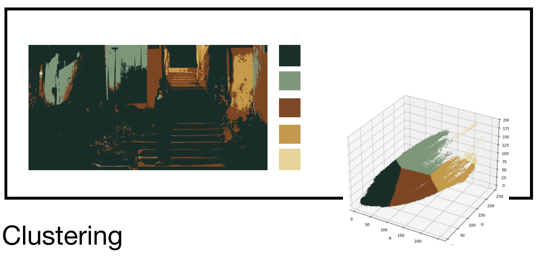
Color Distance: I set a threshold of 100 and iterated through the pixels grouping them if they are within a euclidean distance of 100 to each other (based on their rgb values)

Observations

After visualizing this process on several images, the most saturated colors come from the color distance method. This makes sense if my goal is to get the mos unique colors of an image. The clustering method gave more muted colors which makes sense as I was getting the color from the cluster's center which is more "averaged". Additionally the clustering algorithm tries to balance the clusters which would explain why some less prominent but bright colors aren't picked up on (although they might be with a larger cluster number)

Input Image

Color Distance Threshold



Clustering