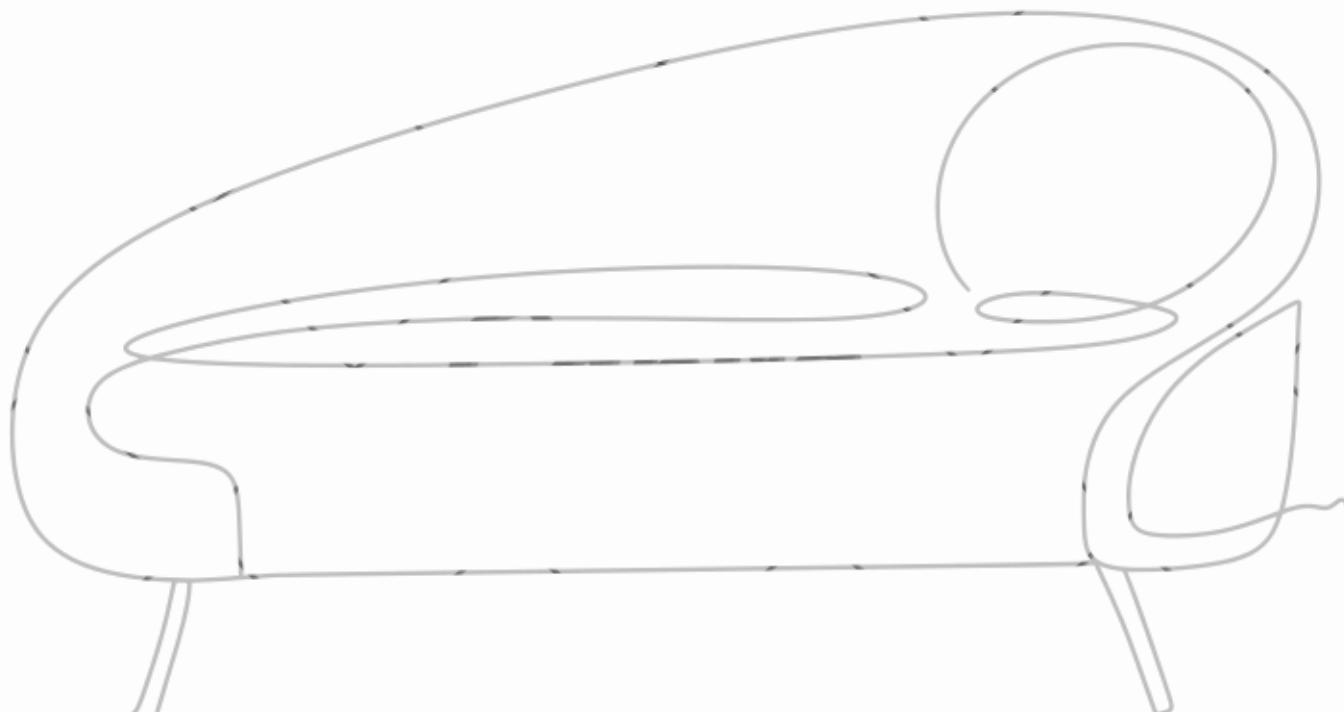
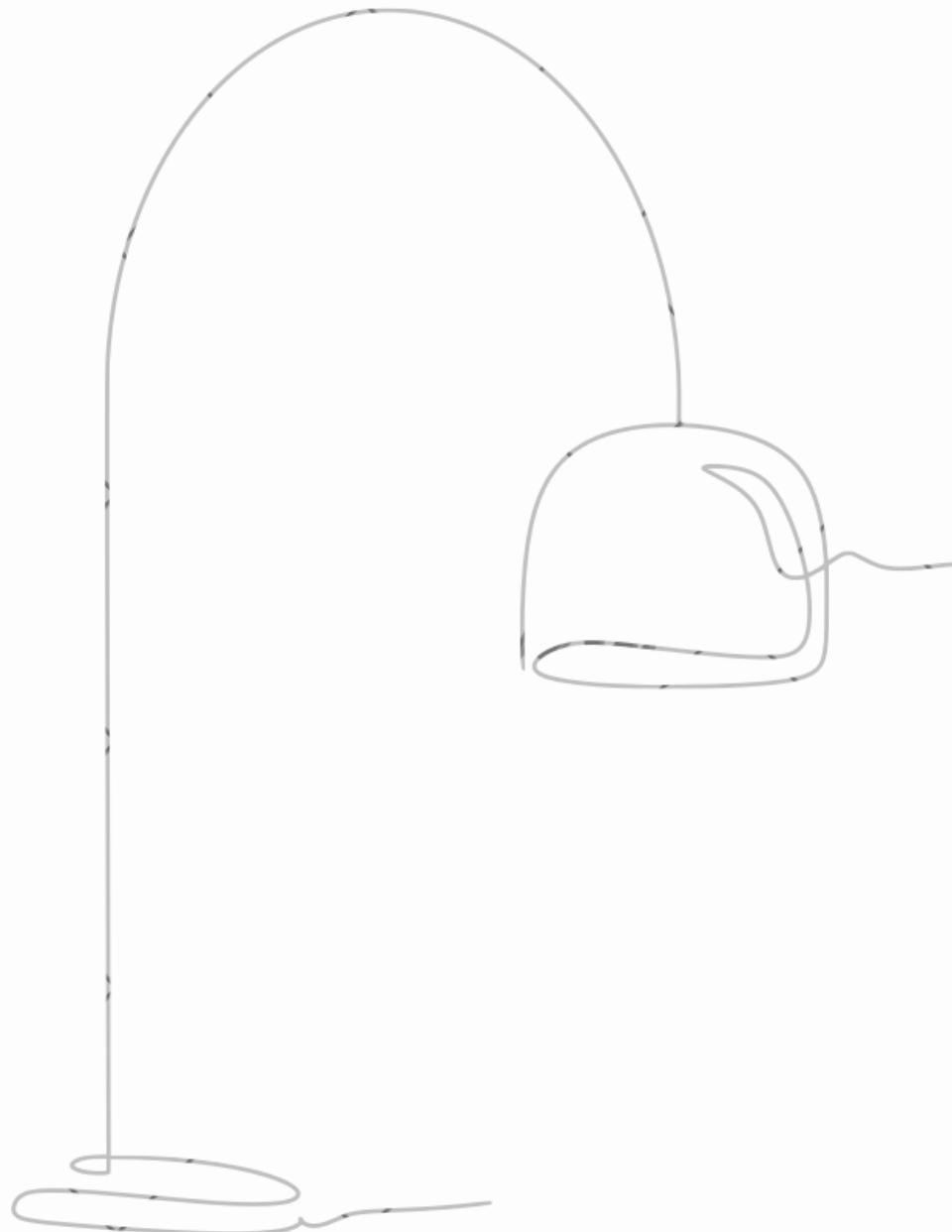


Music Genres Classification



Son Tung - 25%
Mai Luong - 30%
Thanh Hue - 25%
The Phong - 20%

Appendix



I. Introduction

II. Model

1. Convert audio to image
2. Convert audio to features

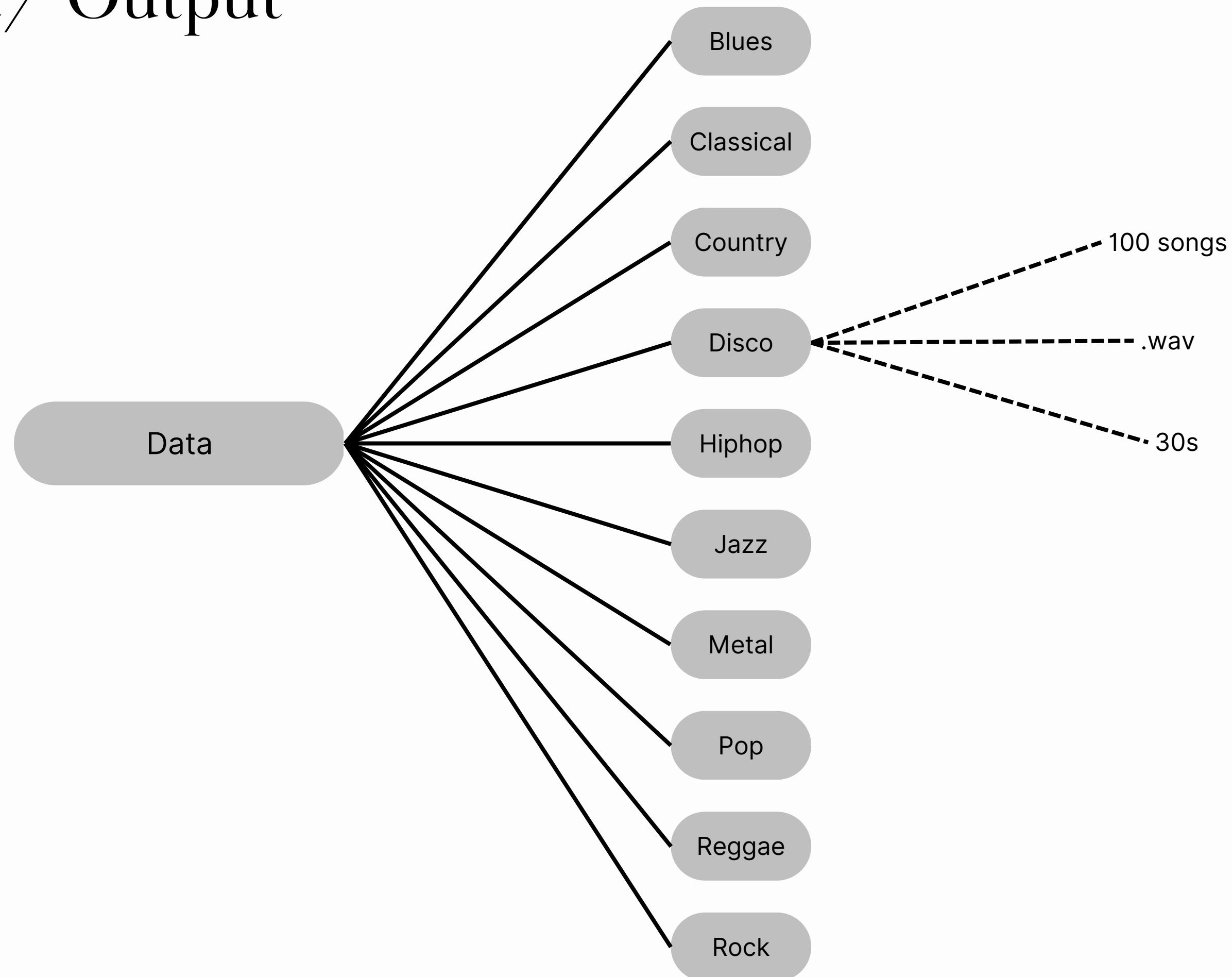
III. Comparision

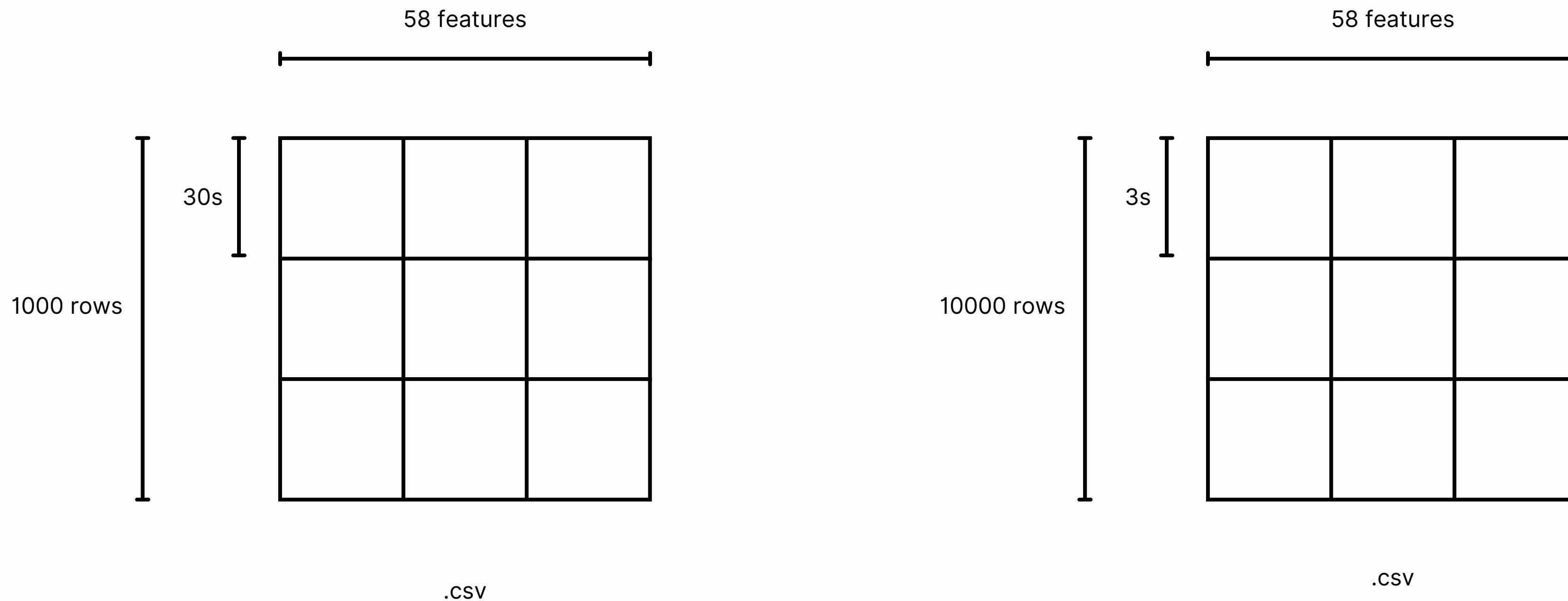
IV. Conclusion

I. Introduction



Input / Output





II. EDA



Waveform

Spectrogram

Spectral rolloff

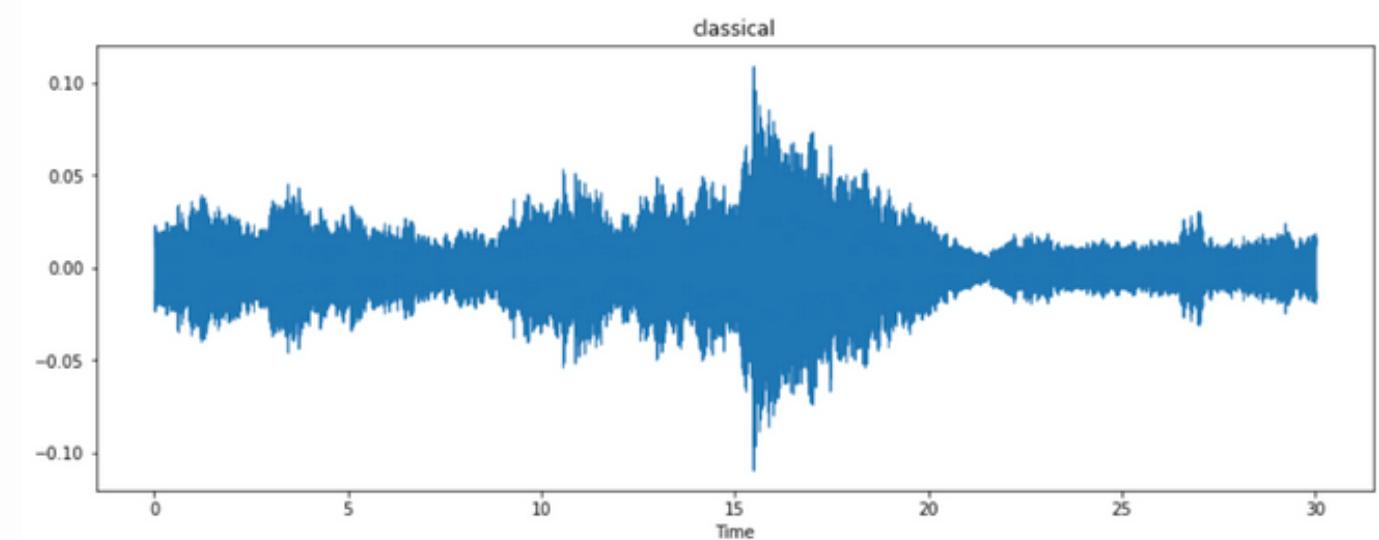
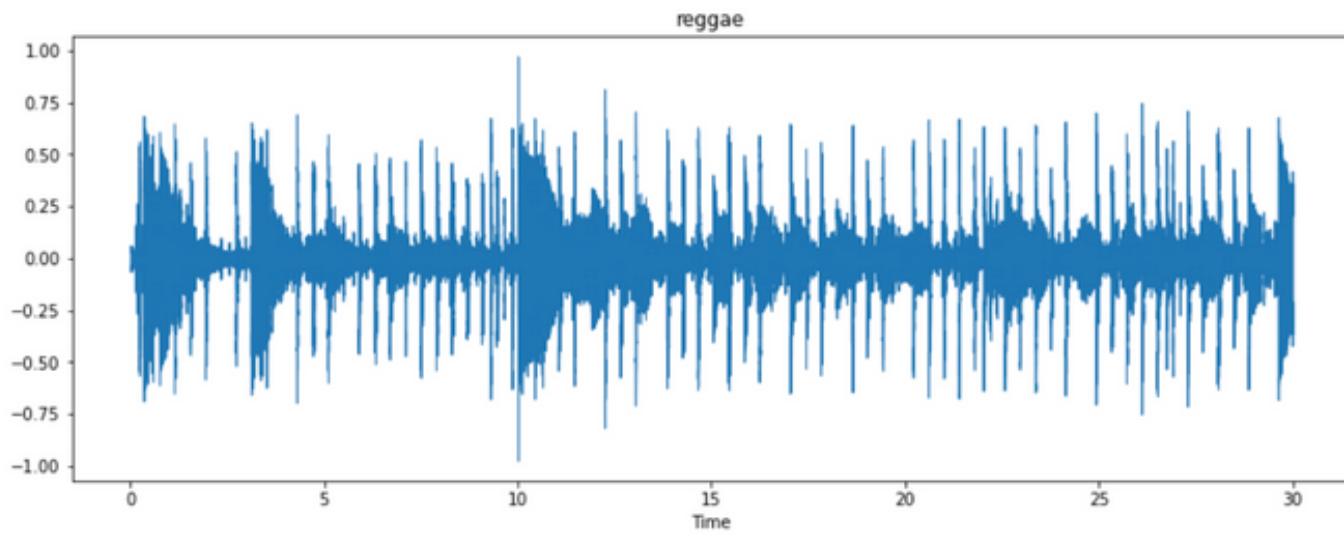
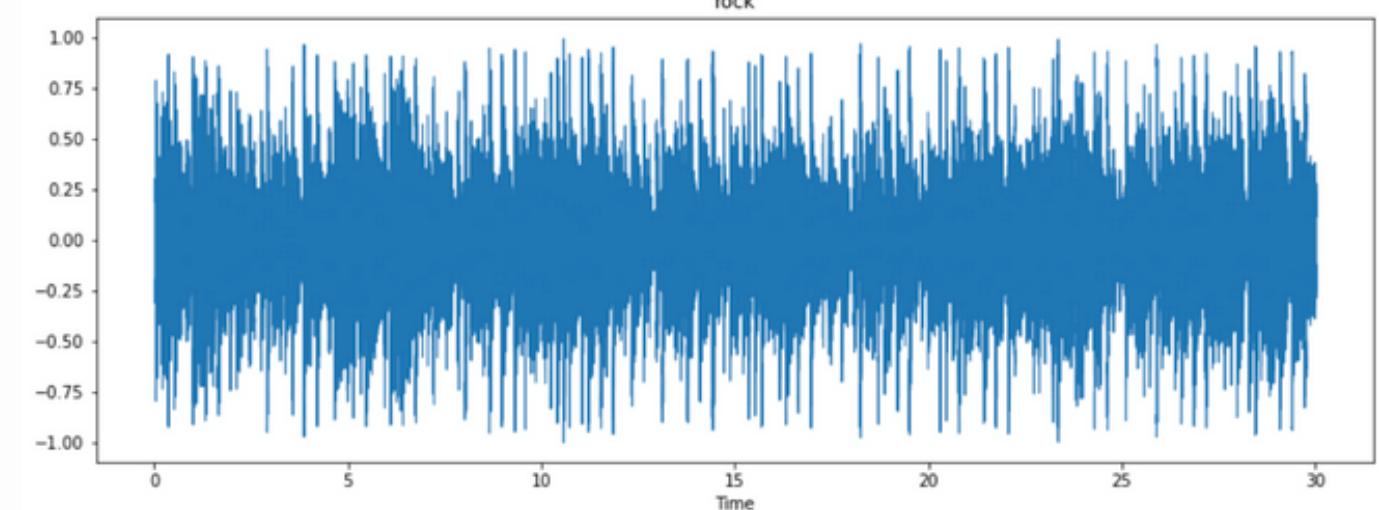
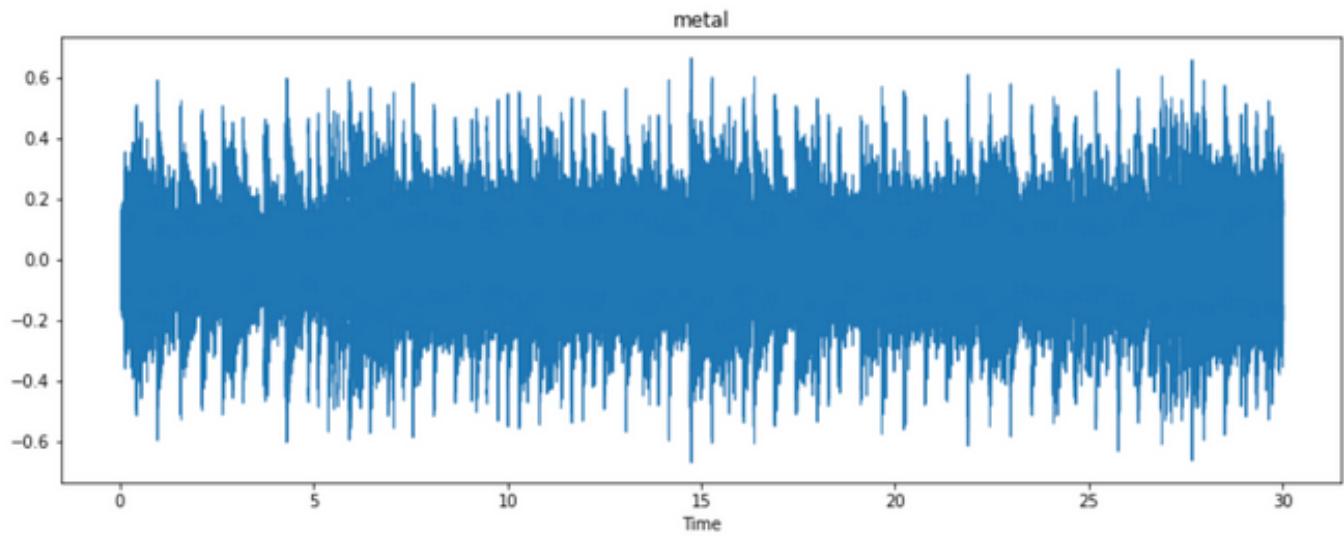
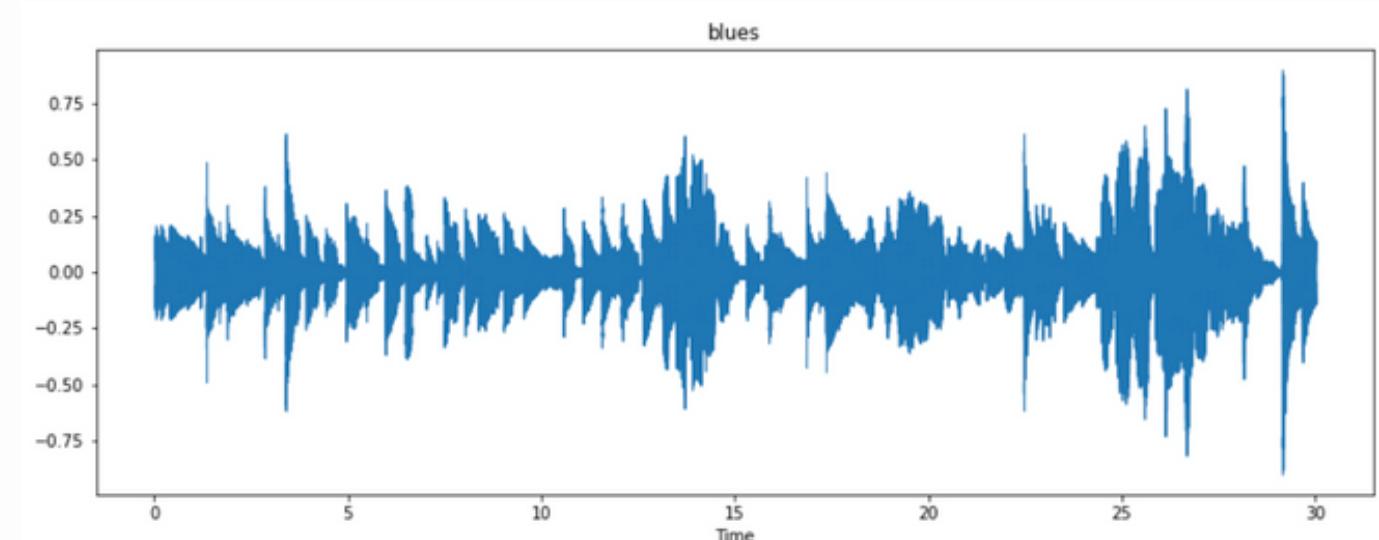
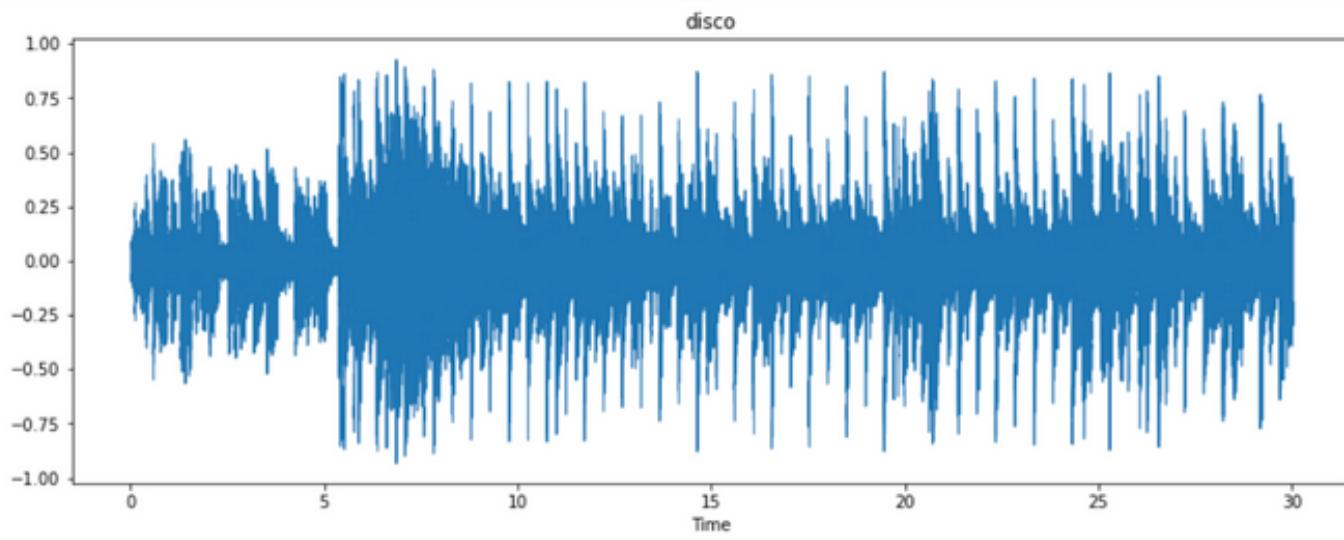
Chroma feature

Zero crossing rate

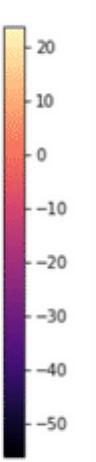
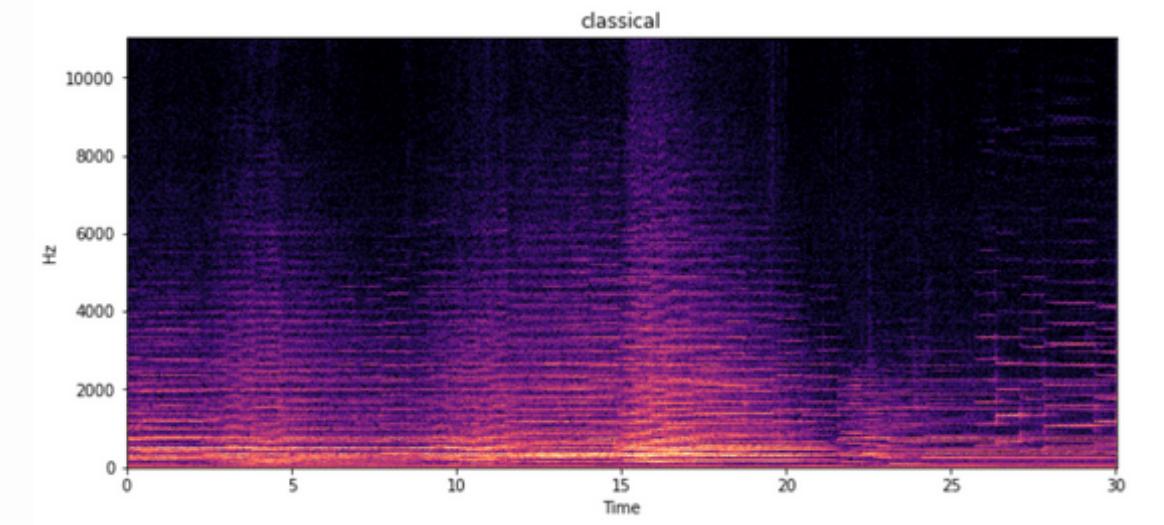
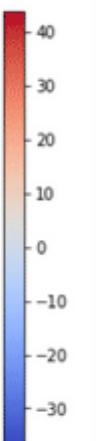
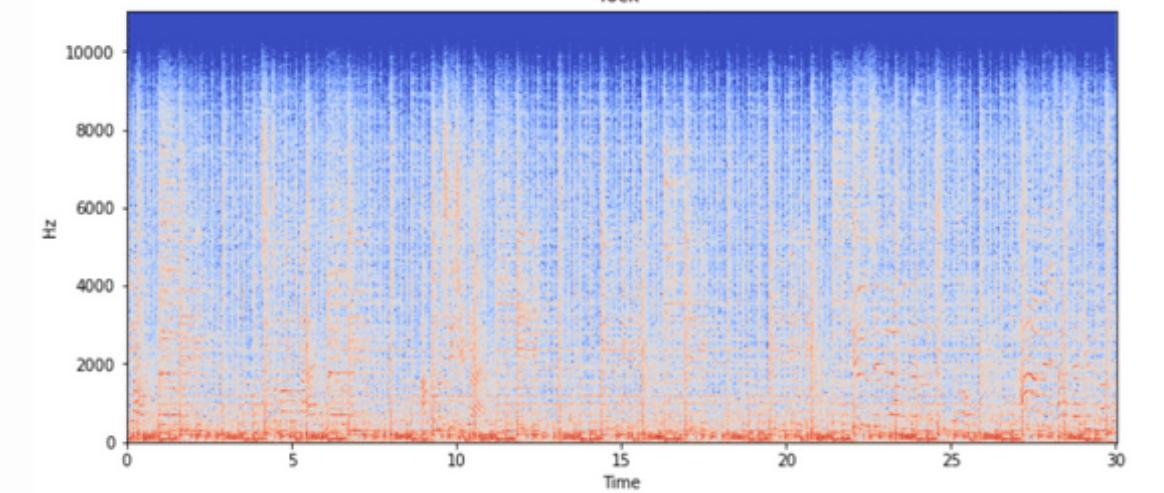
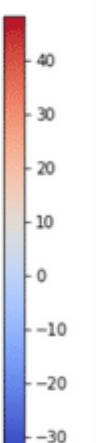
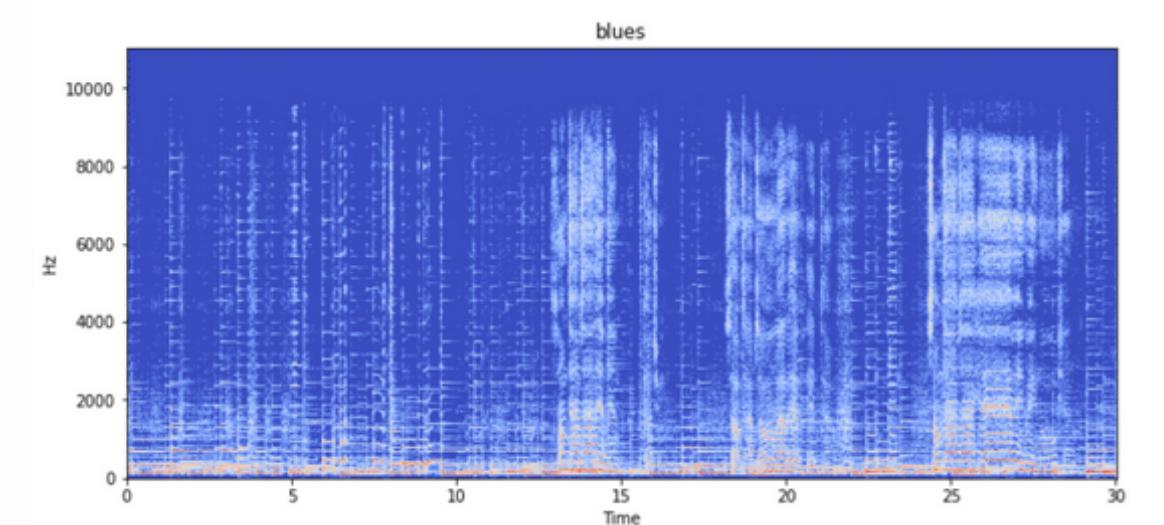
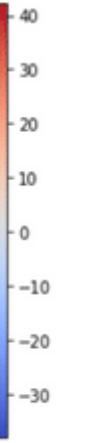
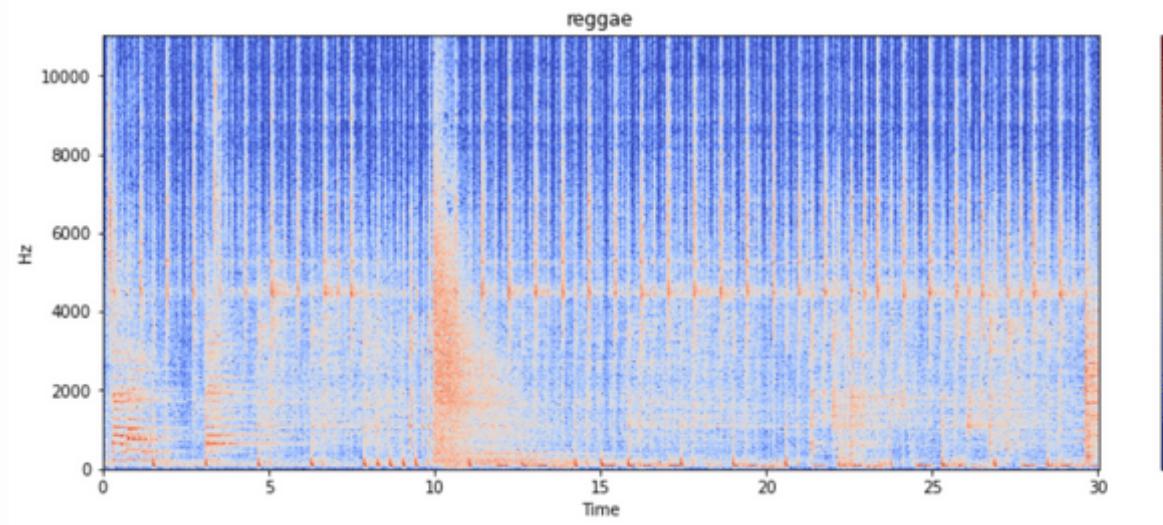
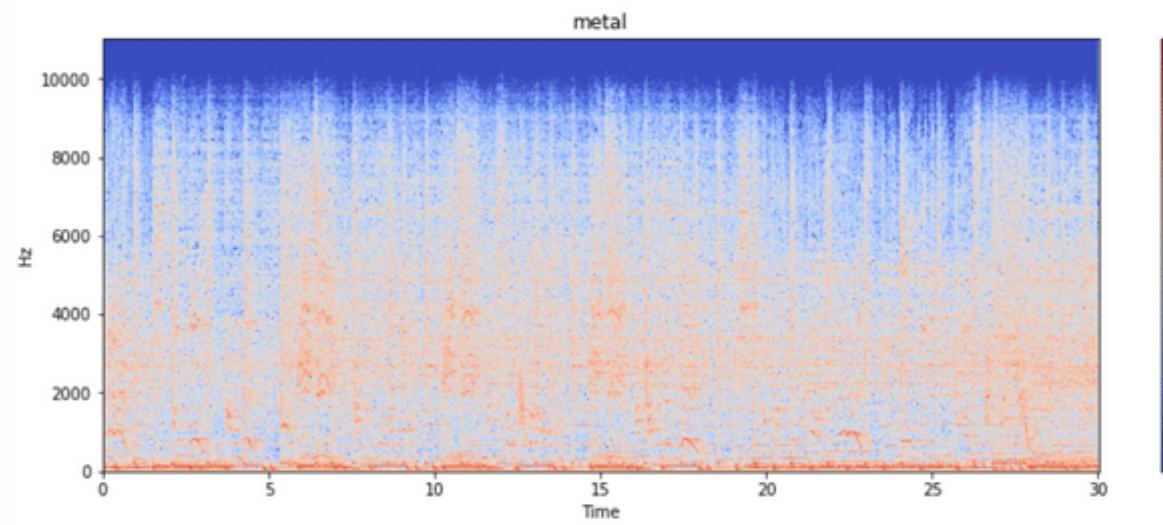
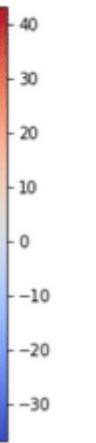
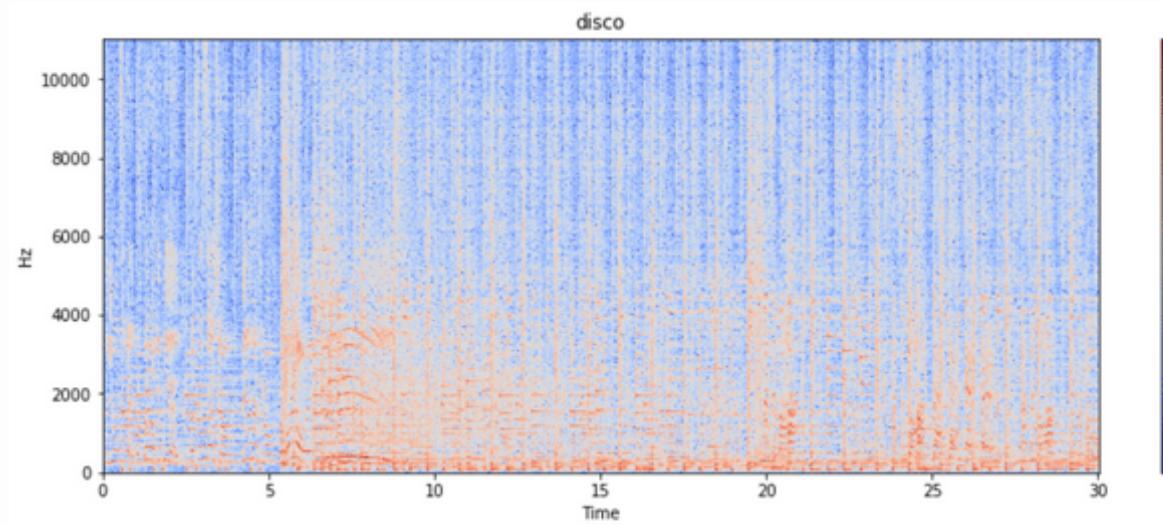
MFCC

Spectral centroid

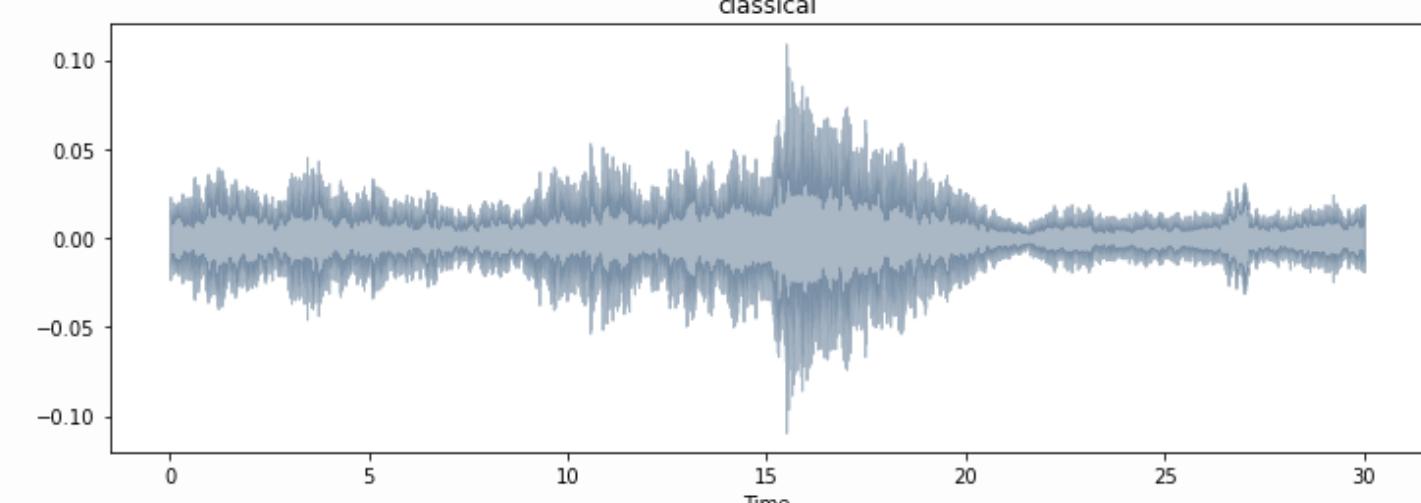
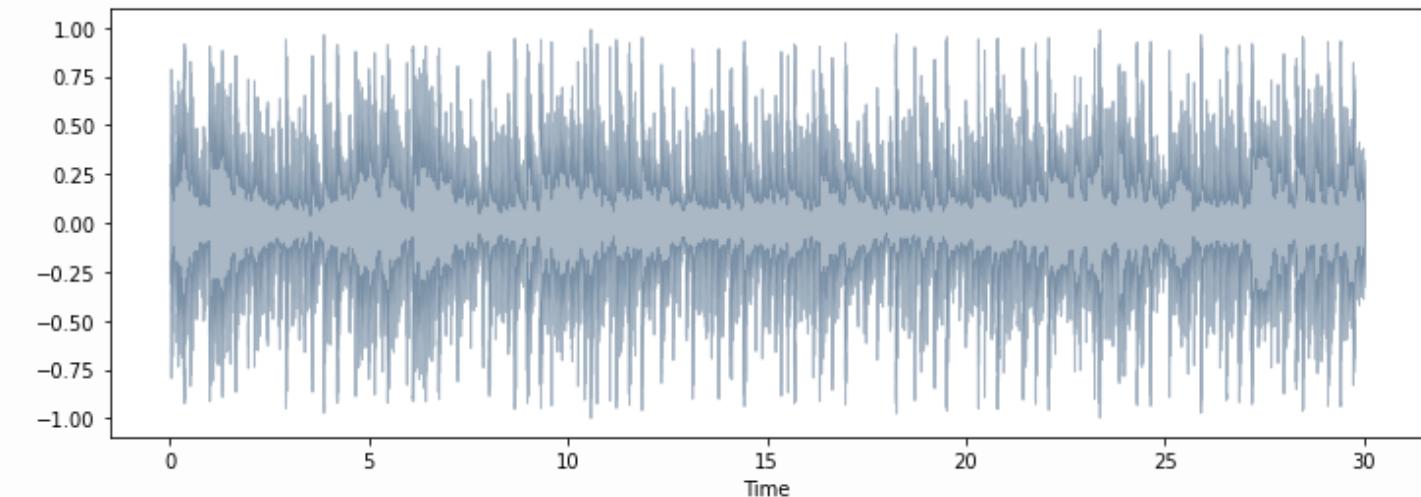
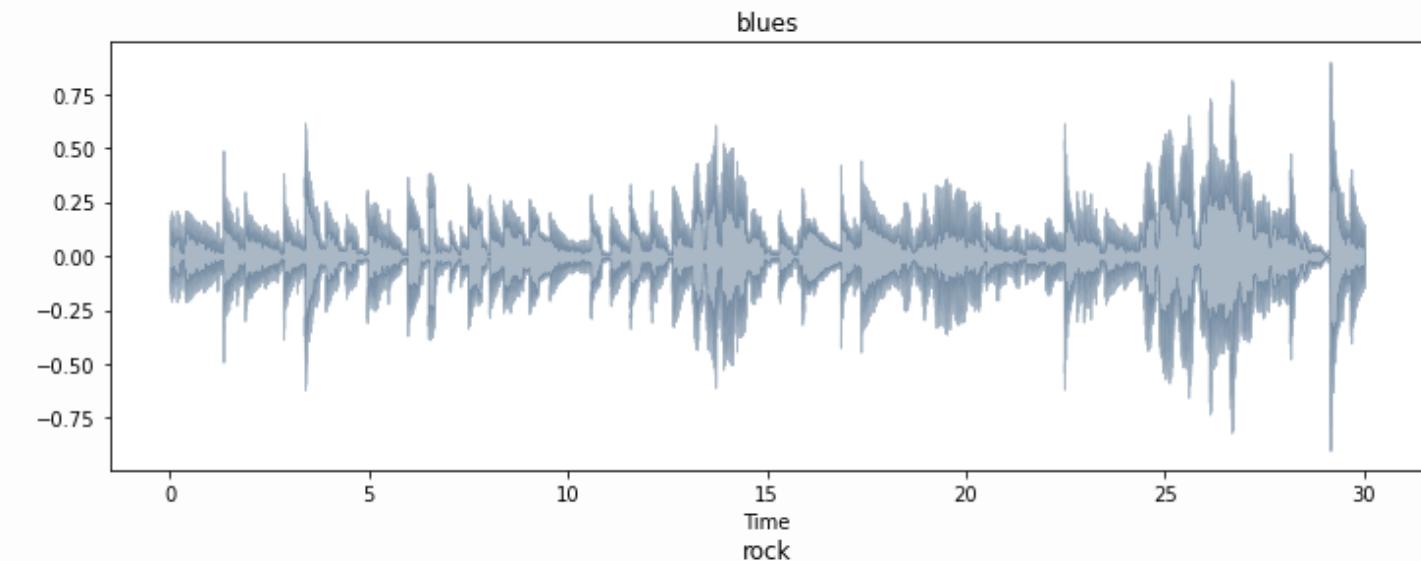
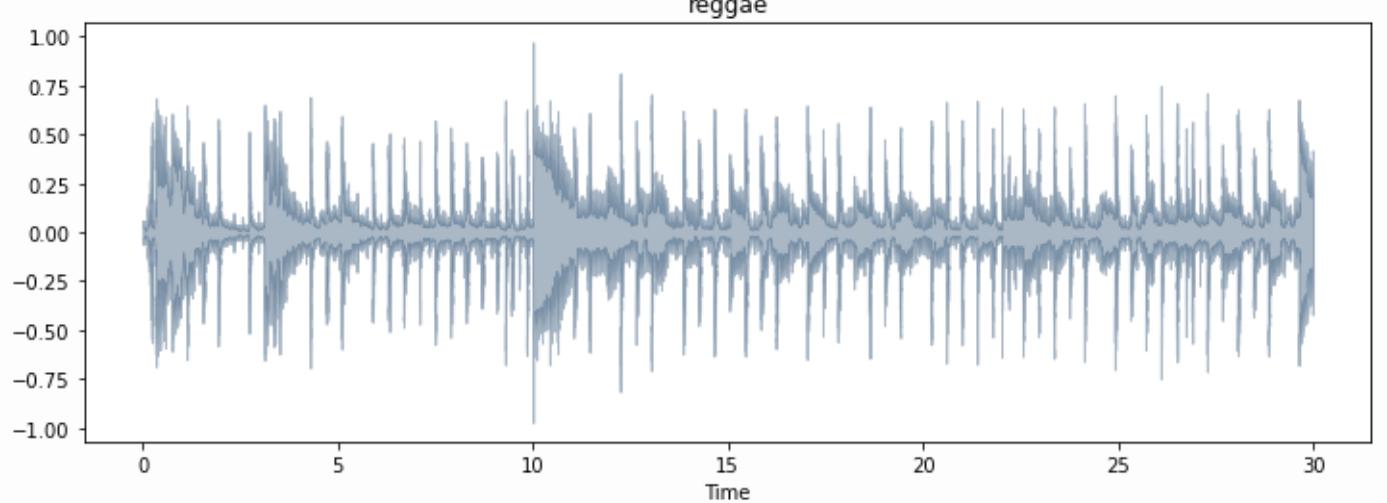
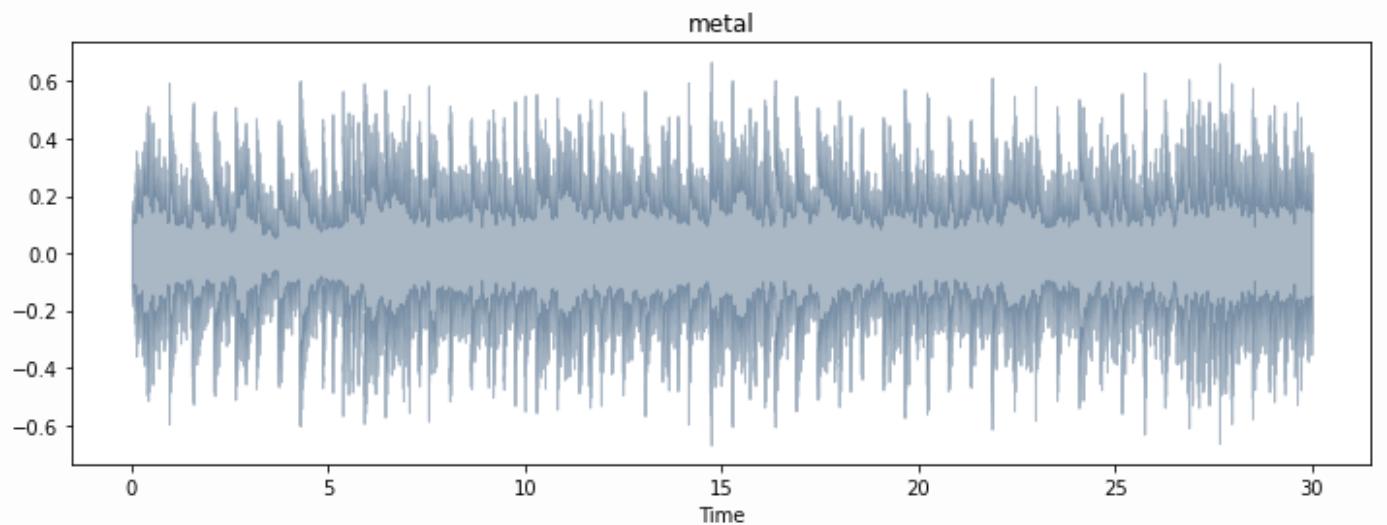
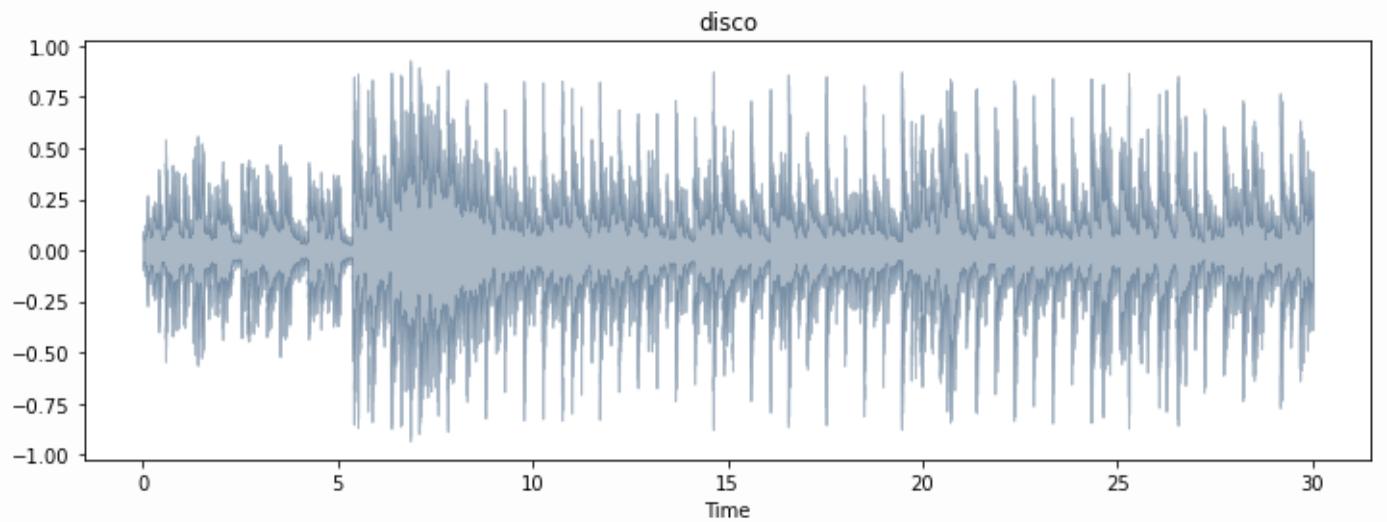
Waveform



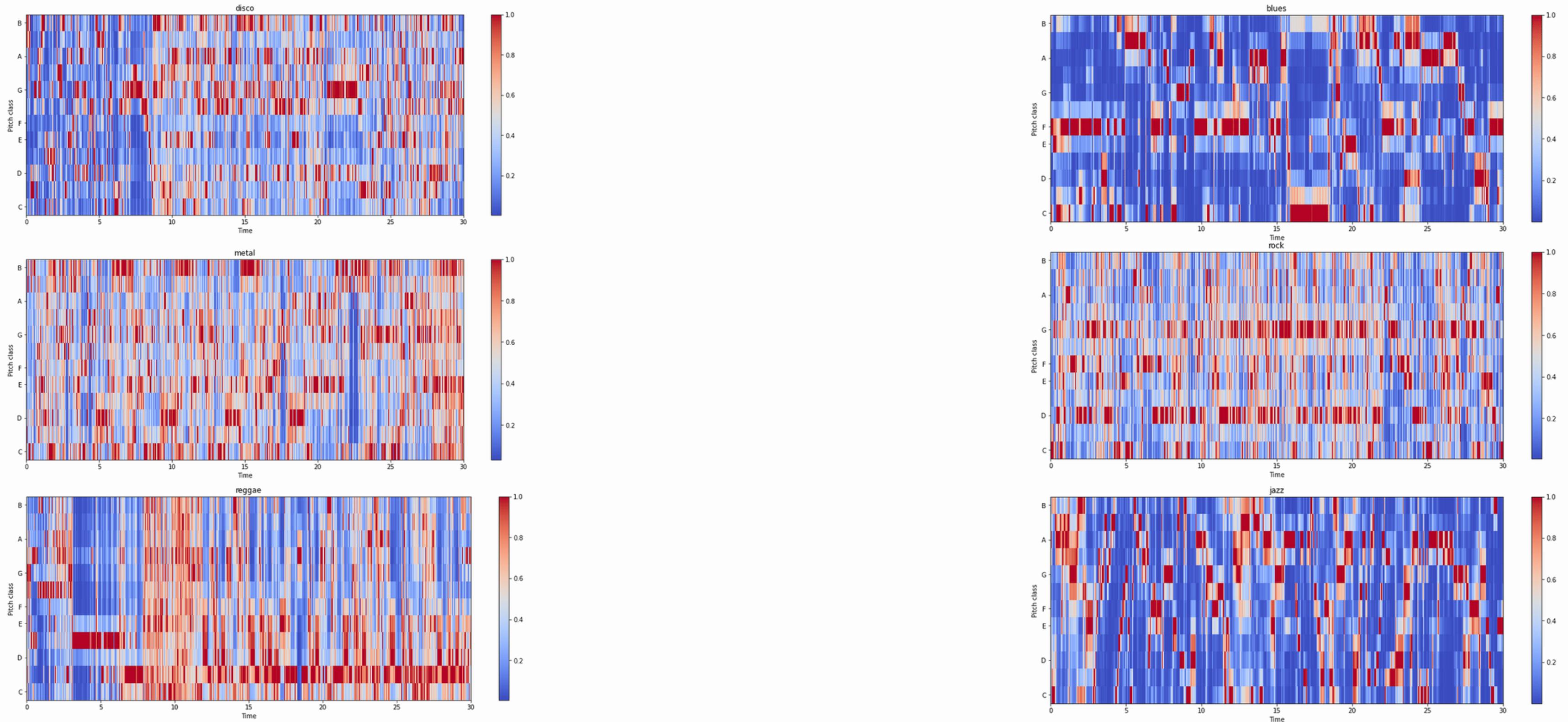
Spectrogram



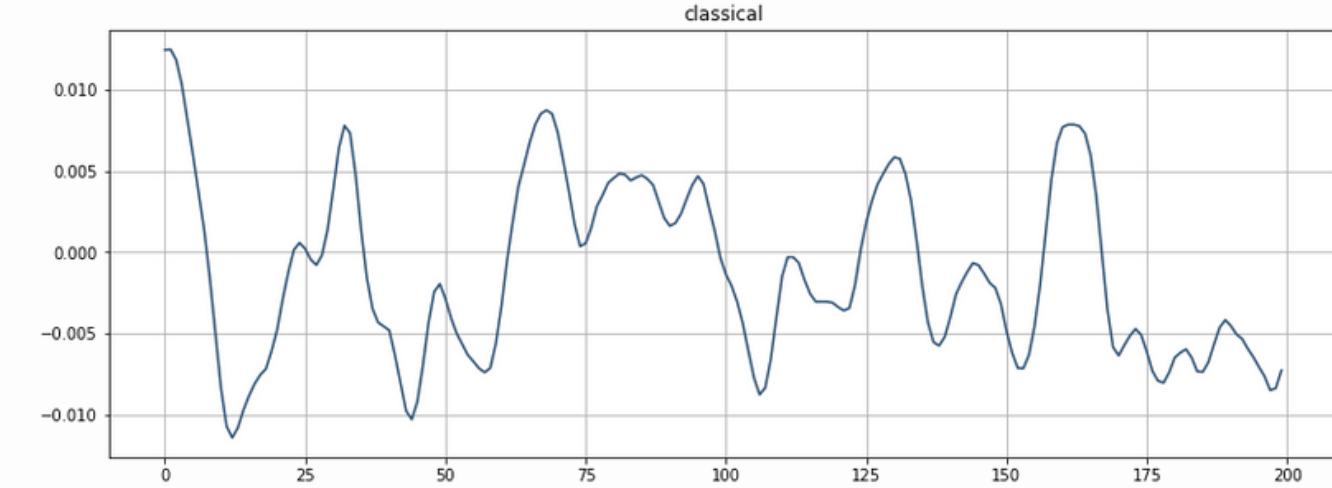
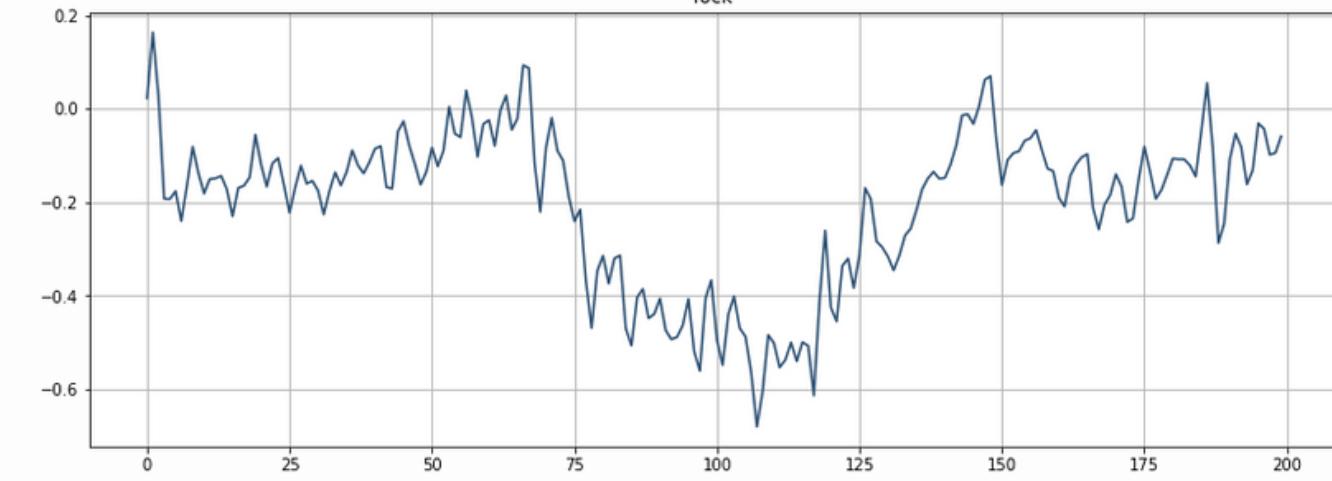
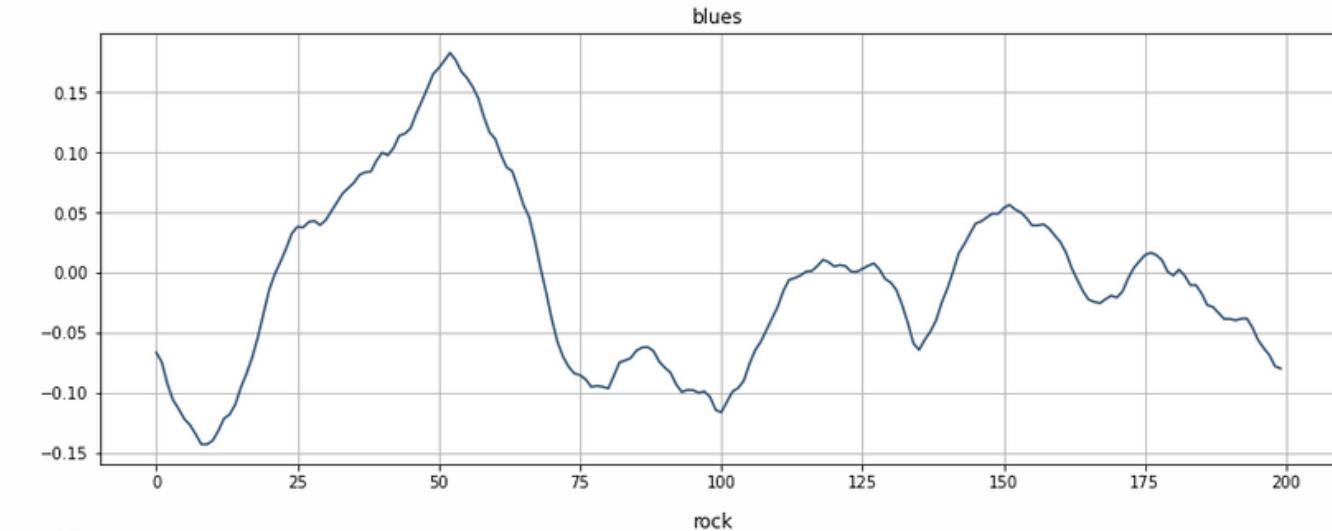
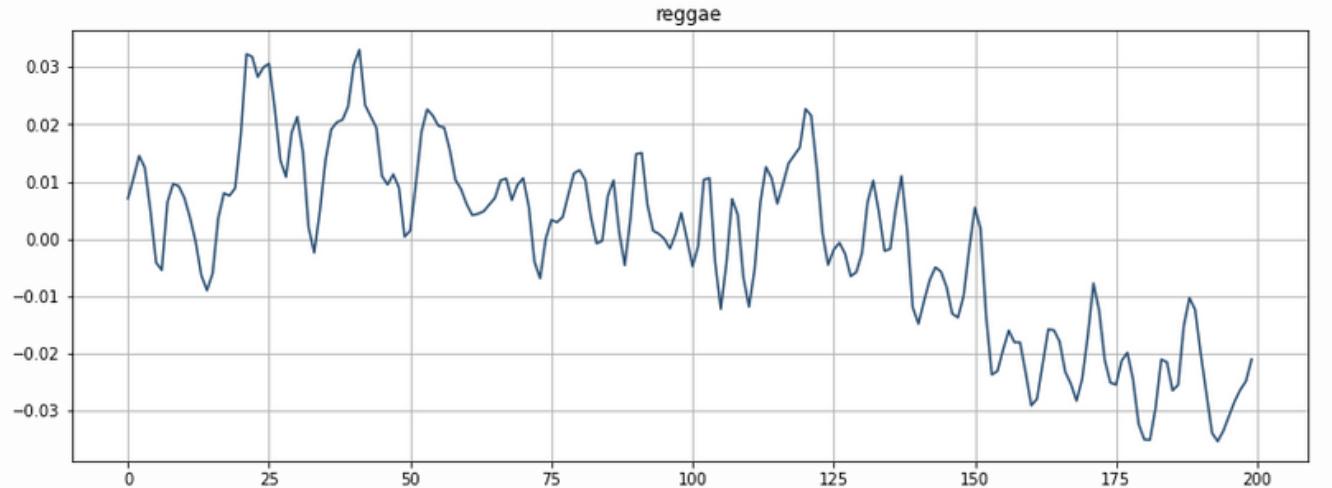
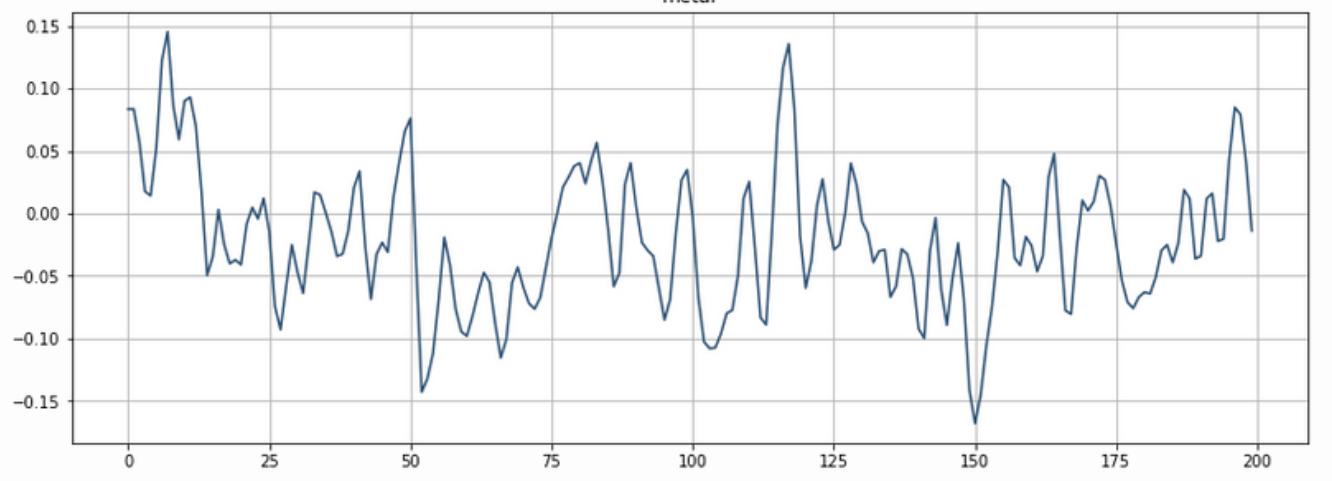
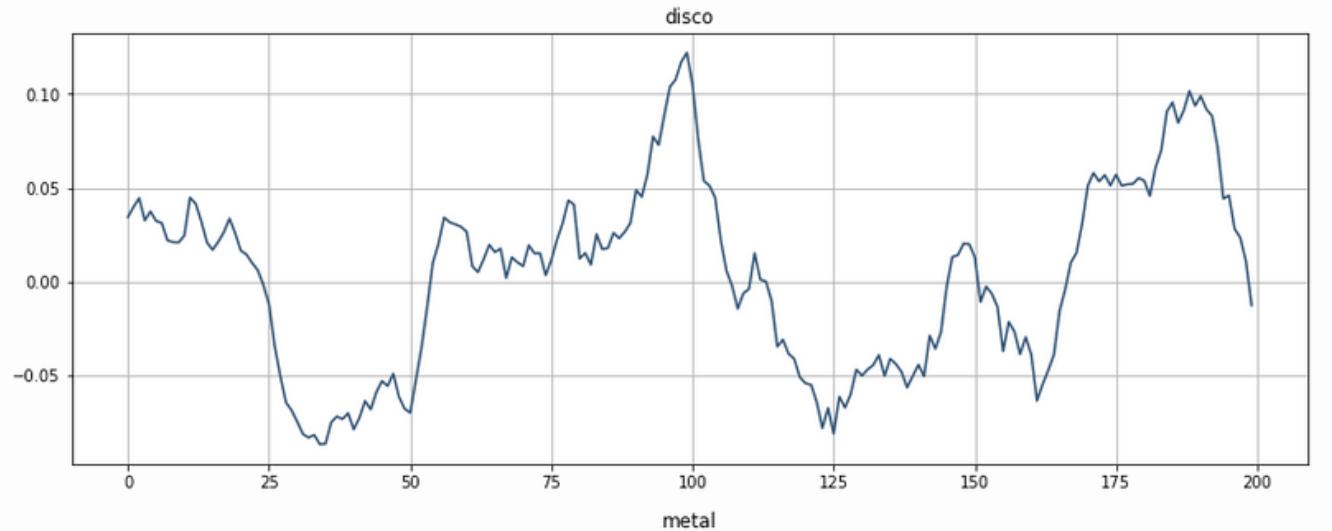
Spectral rolloff



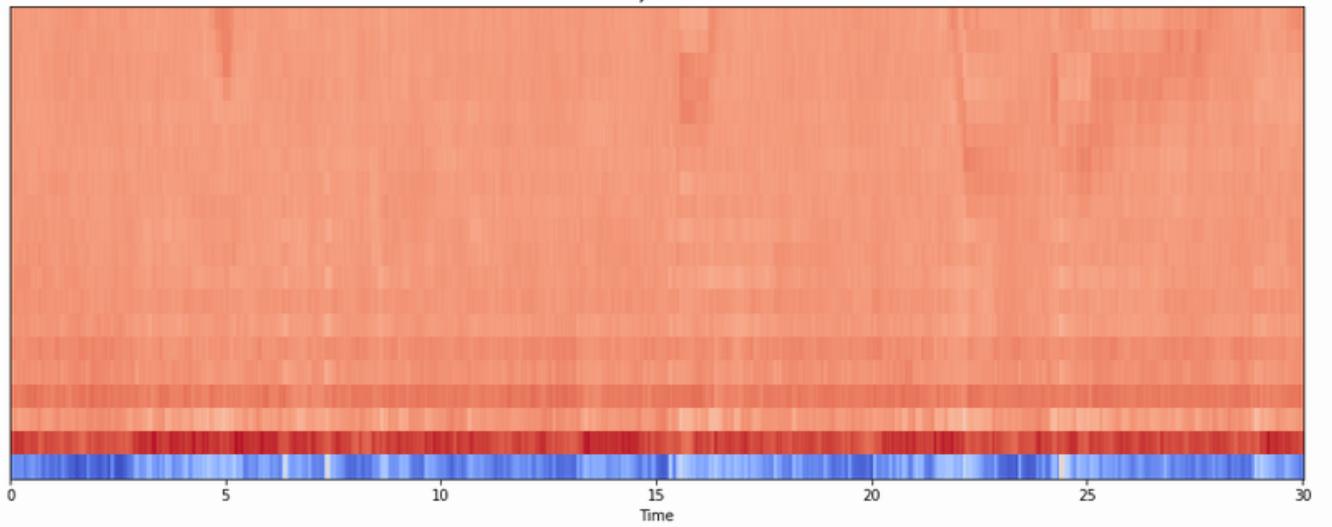
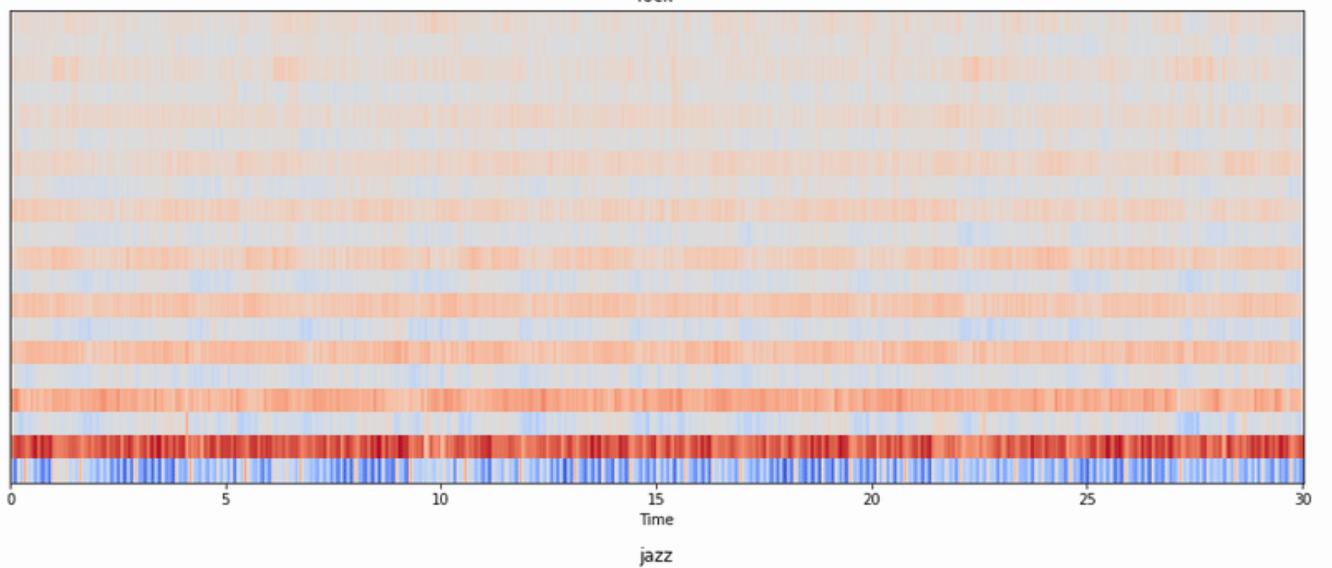
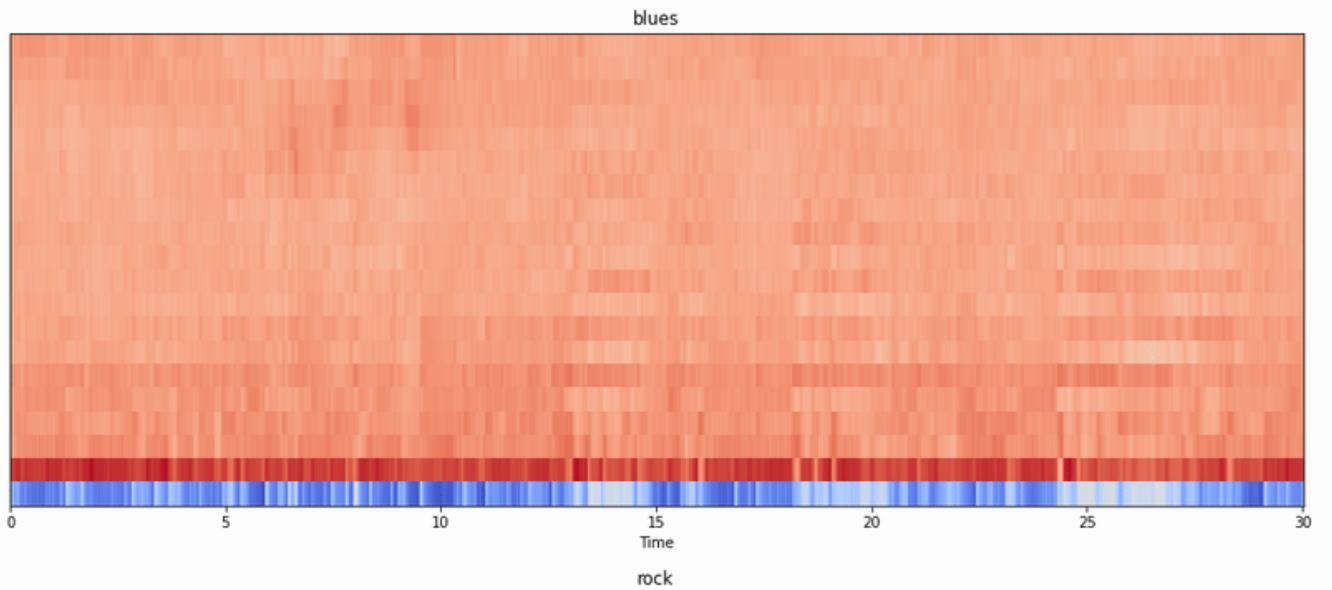
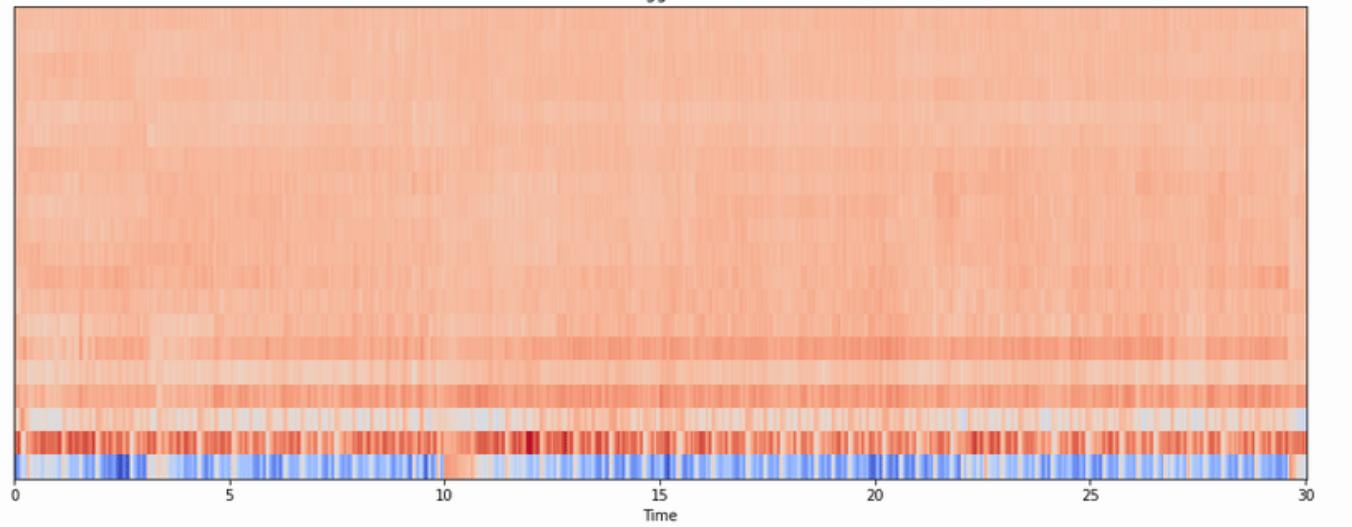
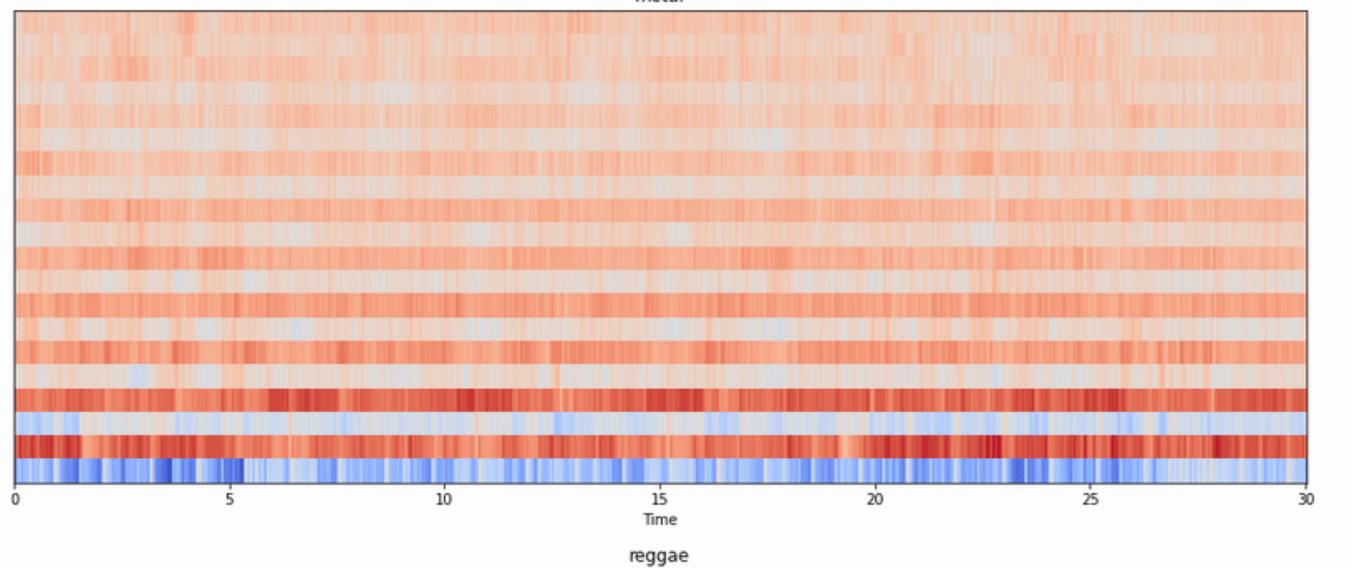
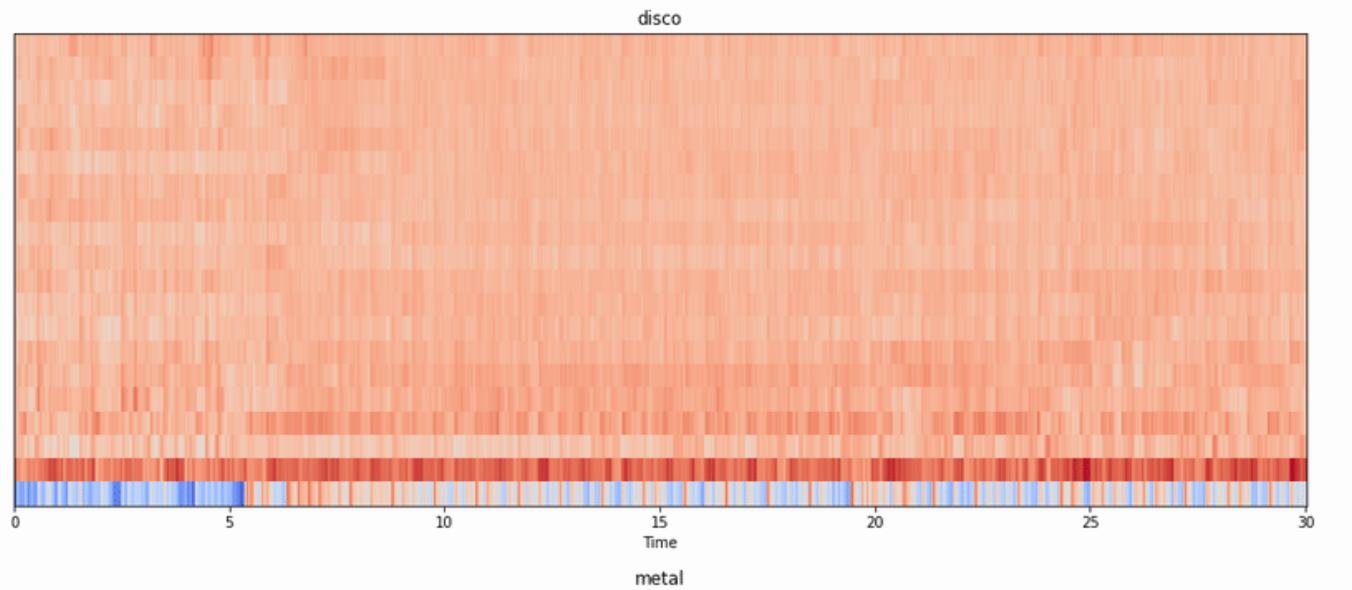
Chroma feature



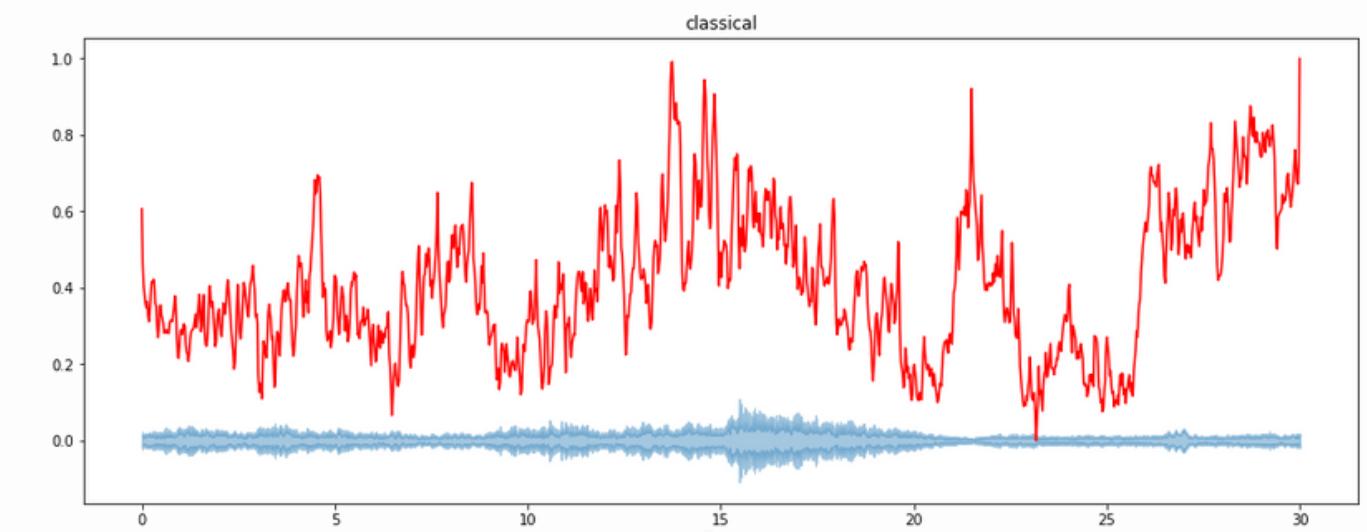
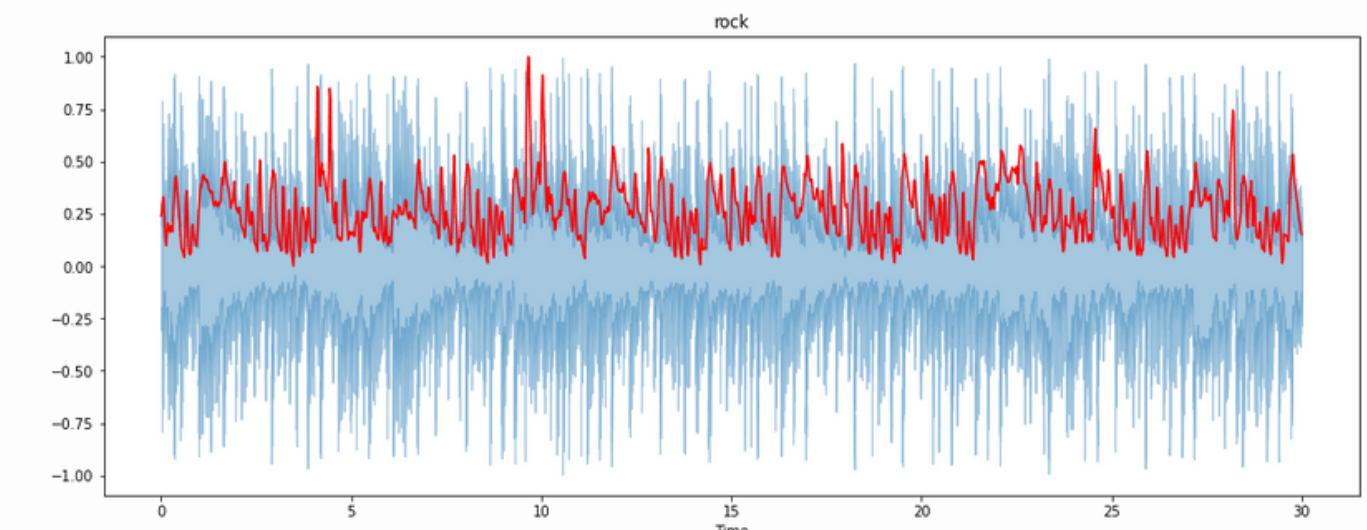
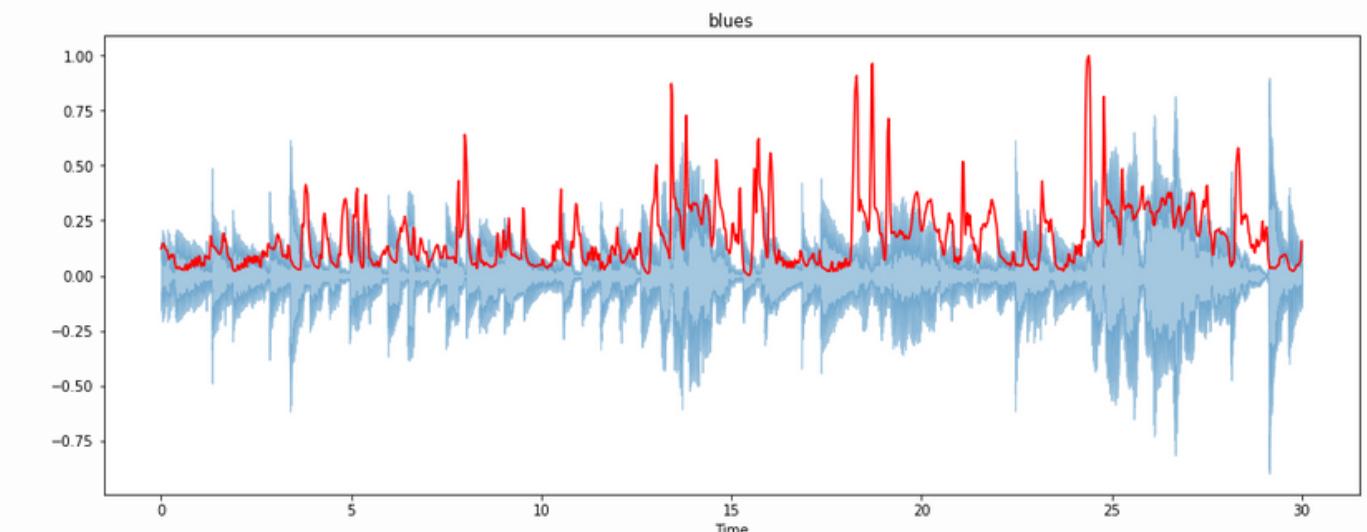
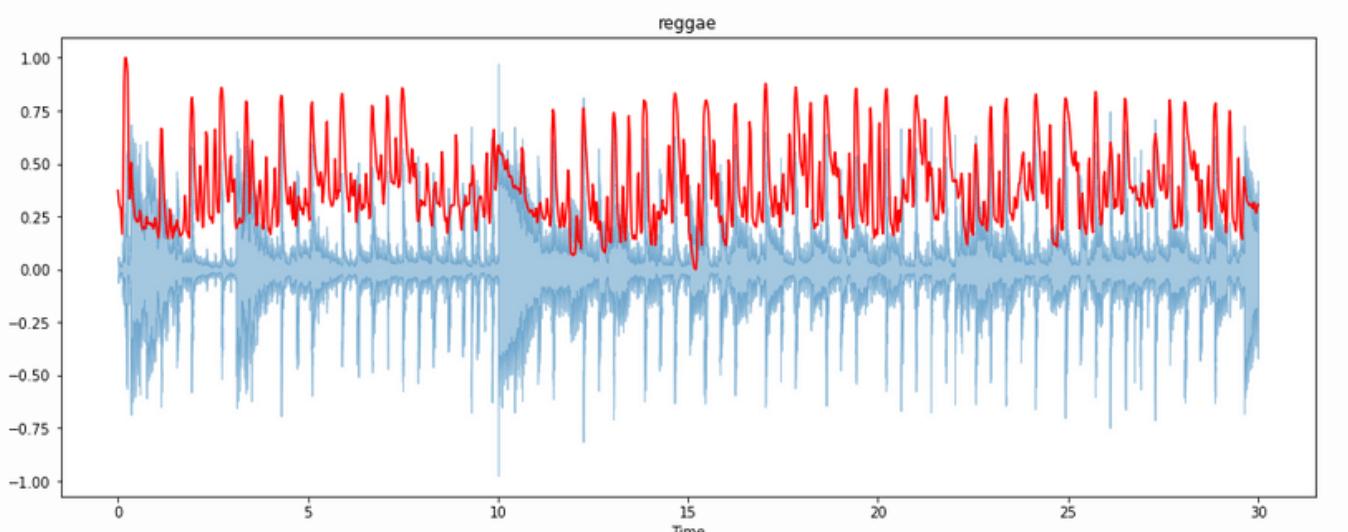
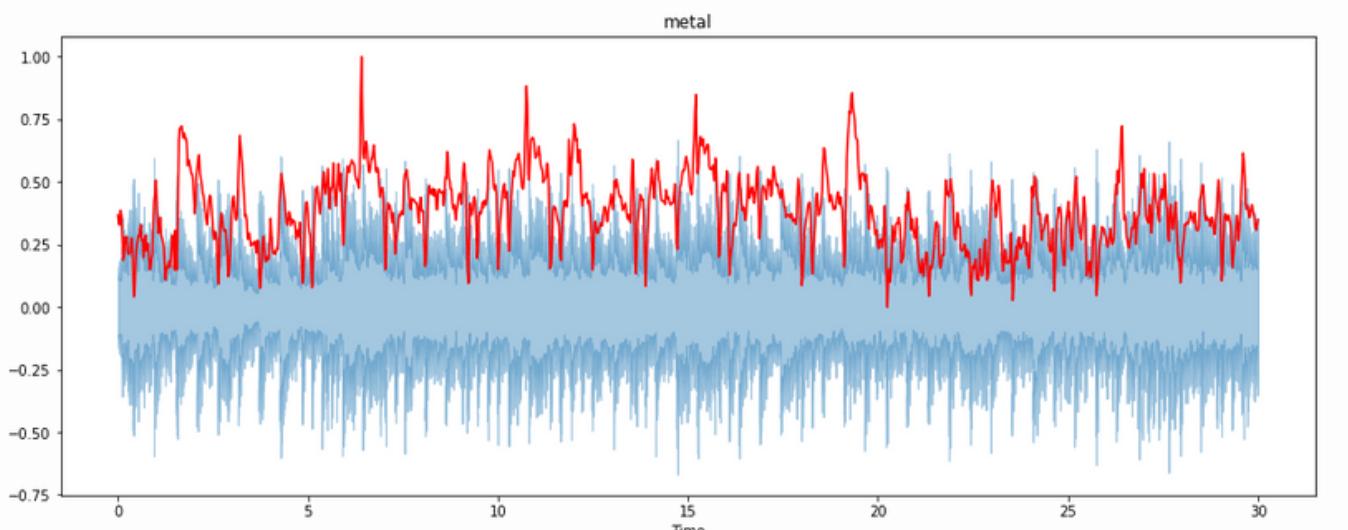
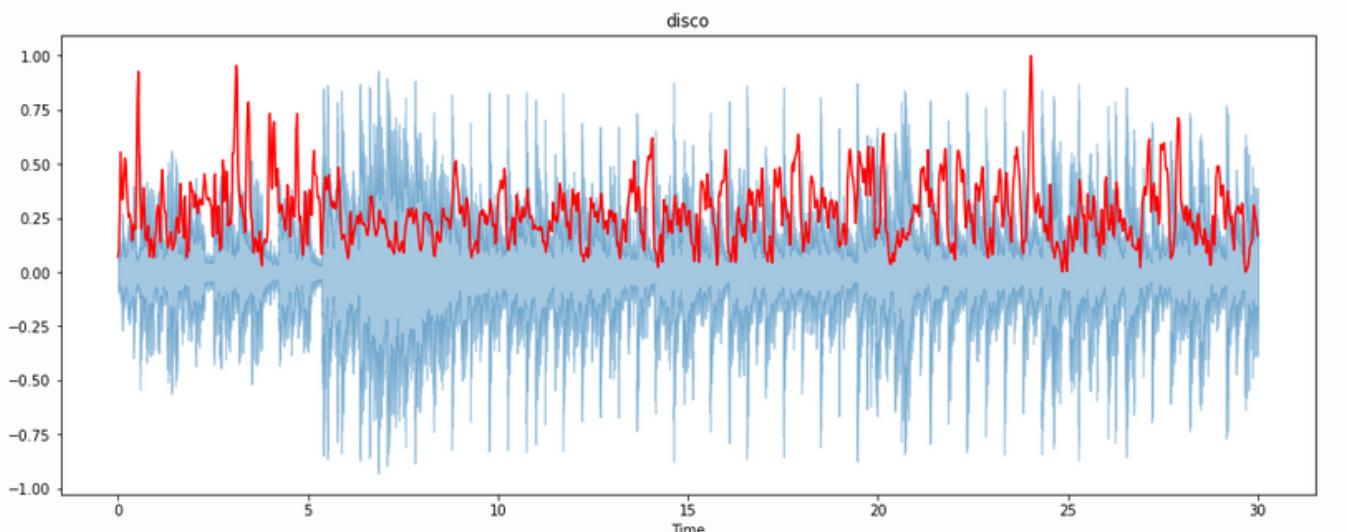
Zero crossing rate



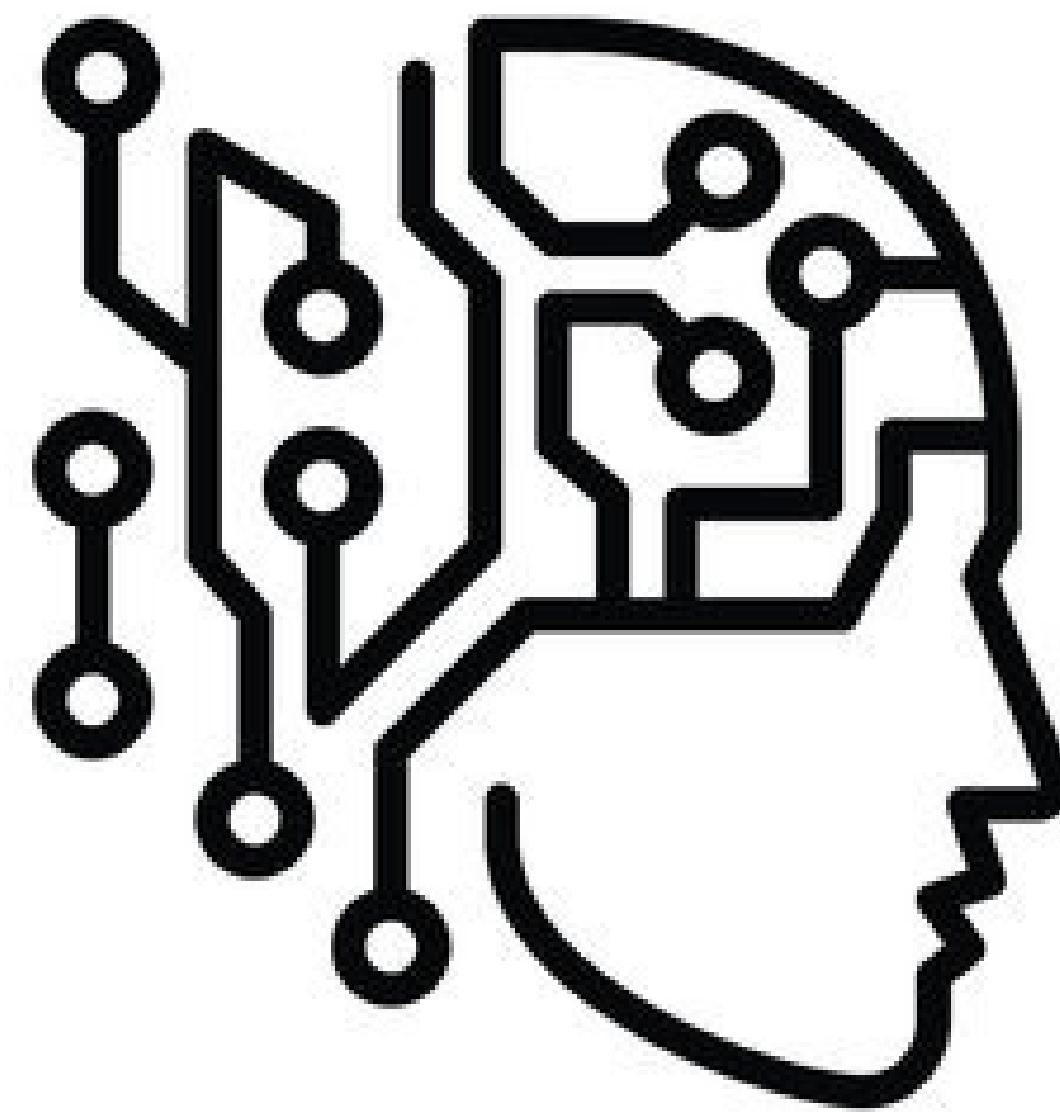
MFCC



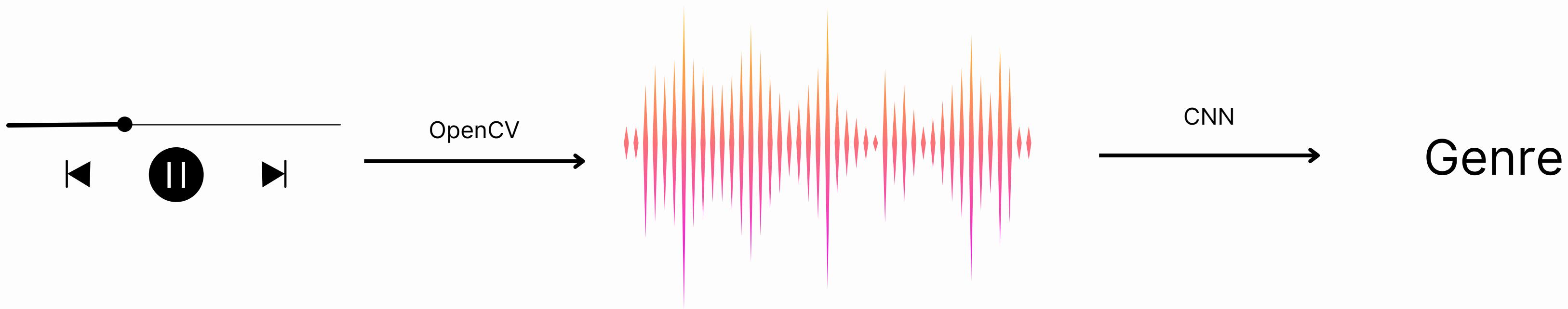
Spectral centroid



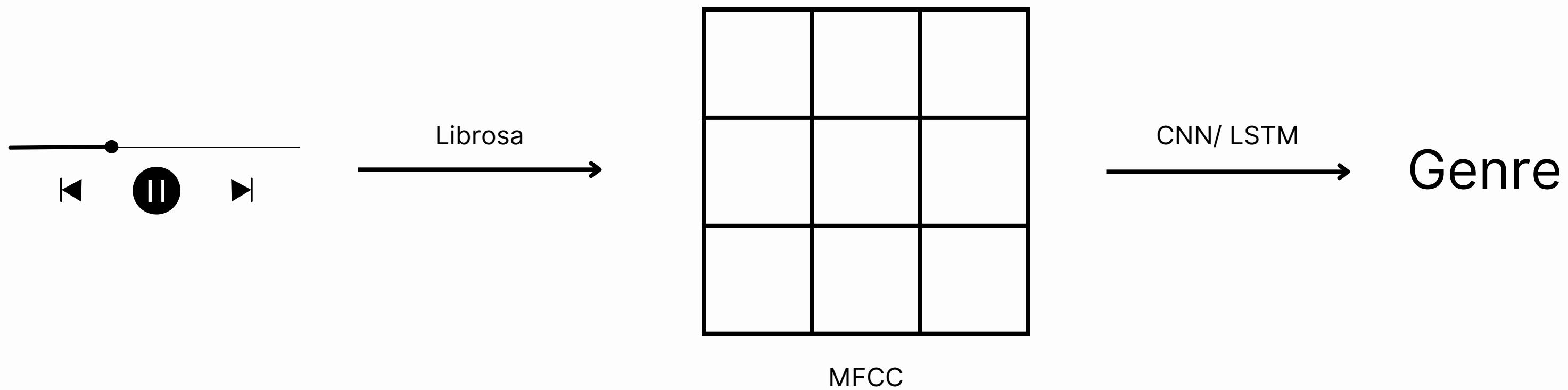
III. Model



1. Convert audio to image



2. Convert audio to features



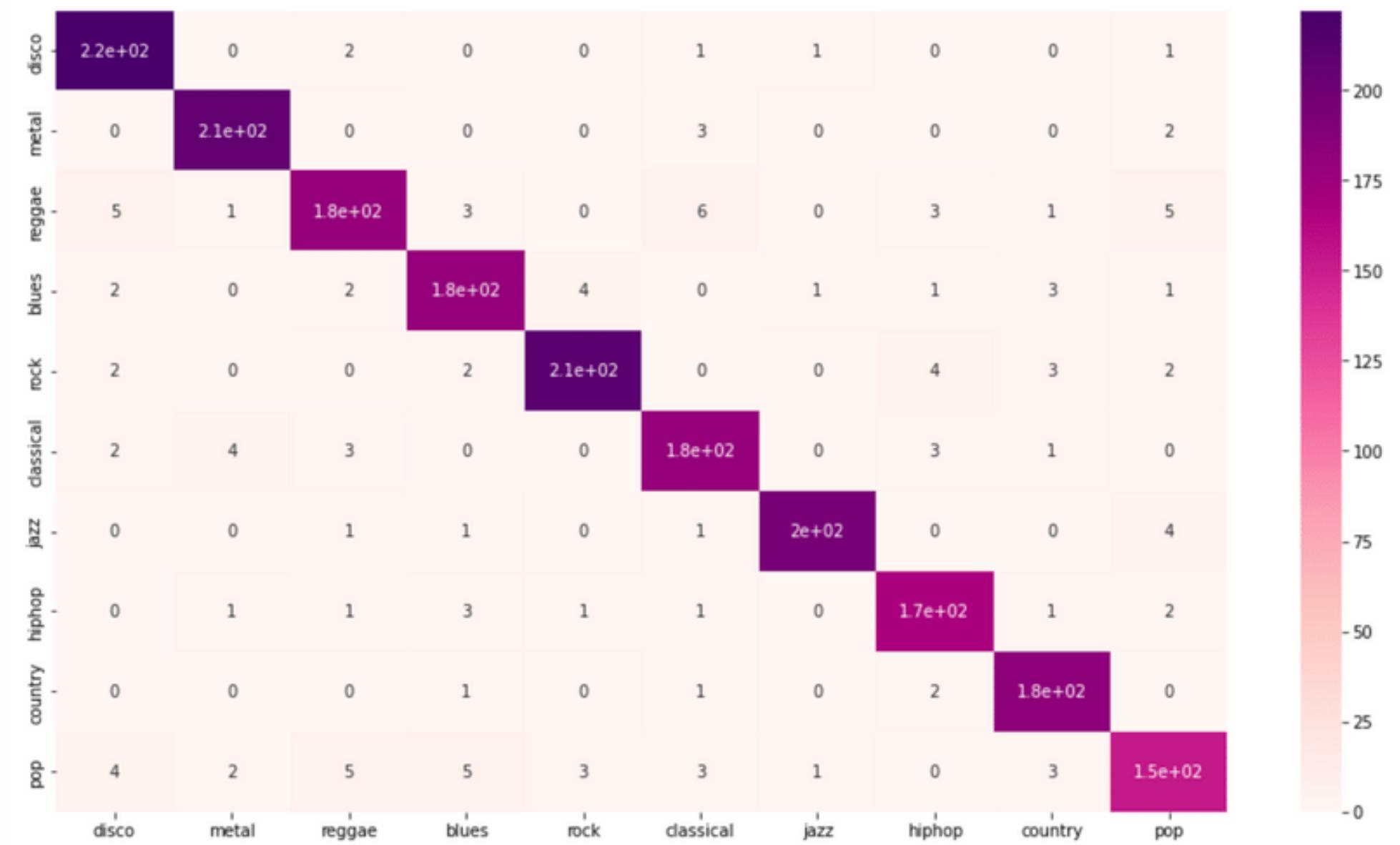
3. Train with .csv





	Audio to image	Audio to features	.csv
CNN	0.42	0.89	0.94
LSTM		0.83	

	precision	recall	f1-score	support
disco	0.94	0.98	0.96	227
metal	0.96	0.98	0.97	212
reggae	0.93	0.88	0.90	203
blues	0.92	0.93	0.93	193
rock	0.96	0.94	0.95	224
classical	0.92	0.93	0.93	192
jazz	0.98	0.97	0.97	202
hiphop	0.93	0.94	0.94	178
country	0.94	0.98	0.96	187
pop	0.90	0.86	0.88	180
accuracy			0.94	1998
macro avg	0.94	0.94	0.94	1998
weighted avg	0.94	0.94	0.94	1998



The end