Music Genre Classifier Specification

Zach Neveu

January 19, 2020

Inputs

- The only input will be a multi-channel audio input. This will be merged down to a stereo representation, as this is what the dataset includes.

Outputs

- The Genre Classifier will output a one-hot vector which will select one of 16 possible music genres.
- Possible genres (limited by the training dataset) are: Rock, Electronic, Experimental, Hip-Hop, Folk, Instrumental, Pop, International, Classical, Old-Time/Historic, Jazz, Country, Soul-RnB, Spoken, Blues, and Easy Listening
- Fewer genres can be used if this is convenient, however more genres cannot be added

Architecture

- The genre classifier will use a simple convolutional architecture based on Sander Dielman's paper "End to End Learning for Music Audio"
- Future iterations could use a fork off of the audio2midi model to increase efficiency

Dataset and Training

- Dataset Used: Free Music Archive (medium partition): https://github.com/mdeff/fma