USING SPECTROGRAMS & CNNs FOR MUSIC GENRE CLASSIFICATION

Gene Tanaka
gtanaka@stanford.edu

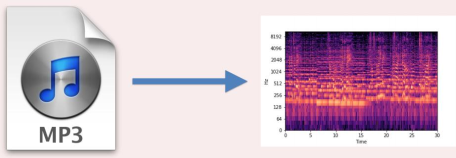
Jaime Hurtado-Lopez
jaimeh37@stanford.edu

Introduction

Modern academics argue that grouping music into genres is inaccurate and outdated – categorizations are mostly based on arbitrary conventions. We investigated how accurately convolutional networks can identify these "patterns" in music of shared genres.

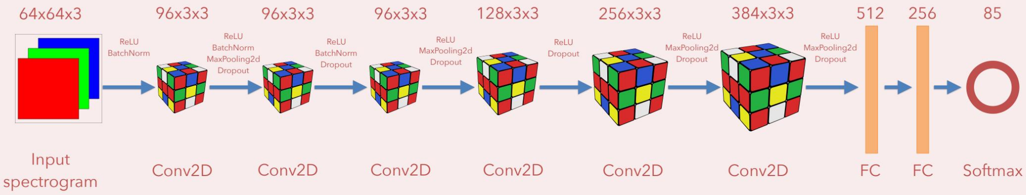
Dataset & Preprocessing

- Free Music Archive
- 8000 .mp3 files, accompanied by .csv file containing genres
- 85 different genres
- LibROSA Python package
- Use library to extract time-series and frequencies from audio files
- Converts .mp3 files into spectrograms
- 432 × 288 RGB images (.png files)
- Crop spectrogram to 64x64



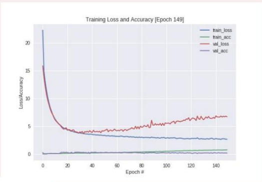
Model

- 6 convolutional layers with increasing filter density to extract features of images
- Pooling and dropout layers to reduce overfitting
- ReLU activation
- 2 fully connected layers at end, softmax output



Results

FIGURE A
- Early iteration
- Overfitting
- Train Accuracy: 98%
- Dev Accuracy: 10%



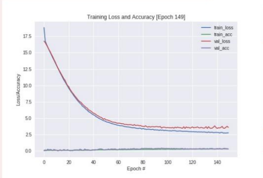


FIGURE B
- Reduced overfitting with batch-norm, L2 regularization, dropout, more data
- Train Accuracy: 72%
- Dev Accuracy: 38%

Discussion

- Smaller filters led to greater accuracy
- Our accuracy was not as high as we would have liked it to be for a number of reasons:
  - Small dataset (8000 training examples)
  - Too many genres (85)
  - Some songs had no genres or multiple genres
- Genres are not truly based on audio patterns

References

[1] Choi, et al. "Transfer Learning for Music Classification and Regression Tasks." ArXiv.org, 13 Sept. 2017, arxiv.org/abs/1703.09179. [2] Krizhevskyet al., 2012. ImageNet classification with deep convolutional neural networks [3] Brian McFee, Matt McVicar, Stefan Balke, Vincent Lostanlen, Carl Thomé, Colin Raffel, ... CJ Carr. (2019, February 13). librosa/librosa: 0.6.3 (Version 0.6.3). Zenodo. http://doi.org/10.5281/zenodo.2564164 [4] Xu, Adrian Yijie, and Adrian Yijie Xu. "Urban Sound Classification Using Convolutional Neural Networks with Keras: Theory and..." Medium, GradientCrescent, 27 Feb. 2019, medium.com/gradientcrescent/urban-sound-classification-using-convolutional-neural-networks-with-keras-theory-and-486e92785df4. [5] Varshney, Kapil. "How to Plot the Model Training in Keras - Using Custom Callback Function and Using TensorBoard." Medium, Medium, 6 Aug. 2018, medium.com/@kapilvarshney/how-to-plot-the-model-training-in-keras-using-custom-callback-function-and-using-tensorboard-41e4ce3cb401.