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Modern academics argue that grouping music into genres is inaccurate and outdated – categorizations are mostly based on arbitrary conventions.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 x 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
- Early iteration
- Overfitting
- Train Accuracy: 98%
- Dev Accuracy: 10%

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