Music Genre Classification with Deep Learning

Rocco Clemente Bonjour
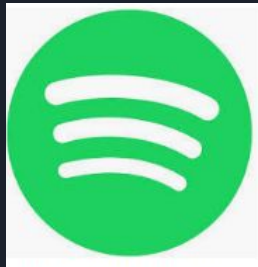Josep Torné Chertó

Overview

Objective

- Design a Music Genre Classification using Deep Neural Networks with an accuracy of the 90%.

Motivation

- Appearance of audio streaming platforms (Spotify)
    - Music recommendations based on listened genre
    - Automatic list creator
- Song recognition from recorded sample (Shazam)

Previous work

- First work by Tzanetakis et al. [1]:
  - Extraction of feature vectors:
    - Spectral centroid
    - Spectral Roll-off
    - Spectral flux
    - Time-domain zero crossings
    - Mel-frequency Cepstral Coefficients

- Appearance of deep learning: extract patterns and trends from large pool of data.
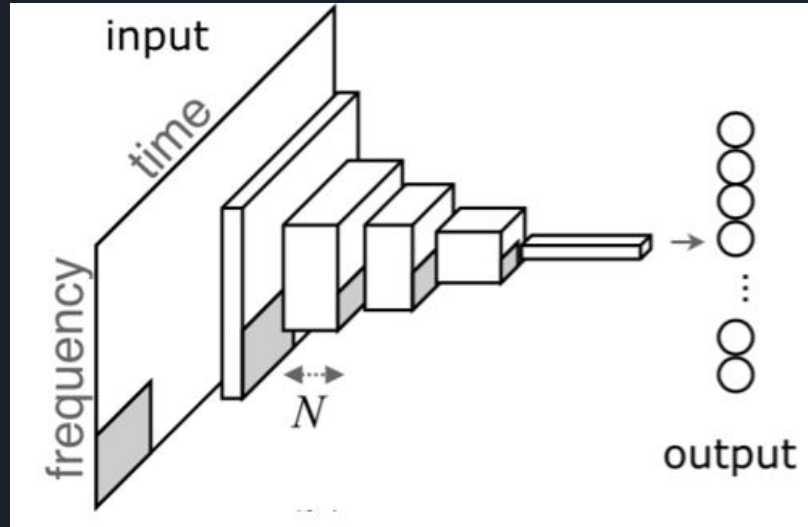 - Extracted feature: Mel-Spectrogram

Framework

- Audio libraries
  - Librosa
- Deep Learning framework
  - Keras over Theano
- Database
 - GZTAN: 1000 songs of 30 seconds, 100 songs per genre

Deep Learning algorithm - Definition

- Usage of a 5-layer 2D CNN with the Mel-Spectrogram of the song as input (Based on Choi et al [2]):
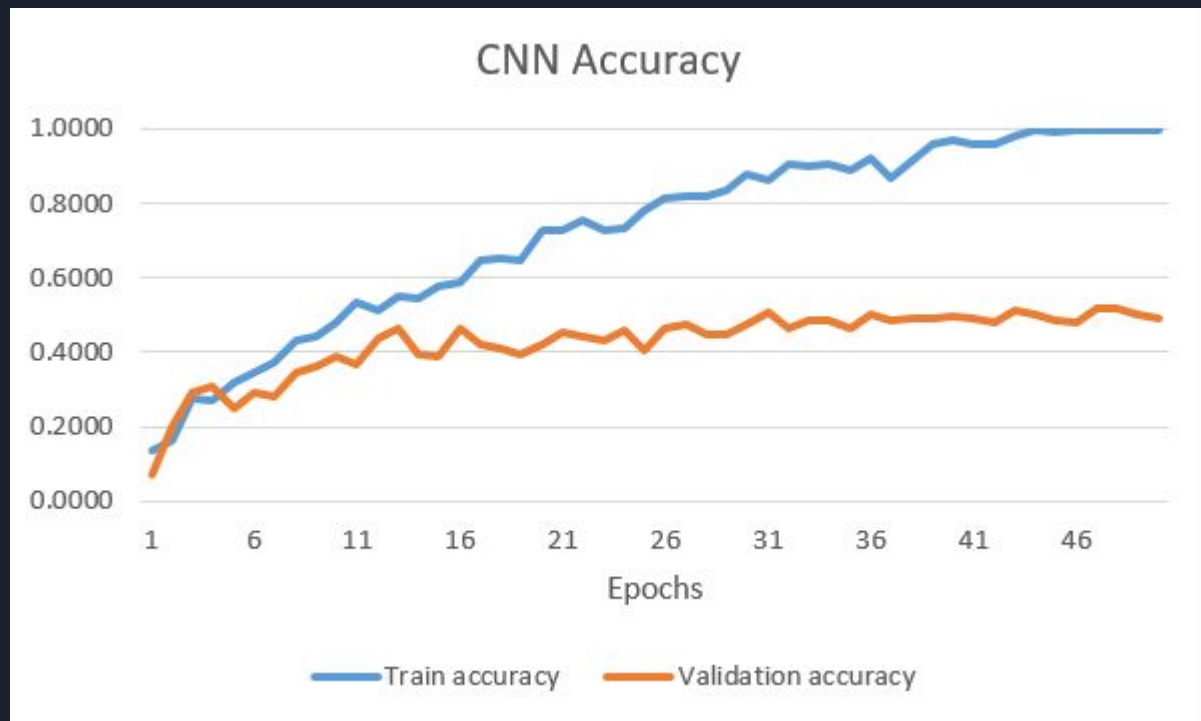
Deep Learning Neural Network - Model Plot

- Convolution 2D layers:
  - Size 3x3 kernels.
- Max pooling layers:
  - Between convolutional networks and with different sizes.
- Dense layer (softmax):
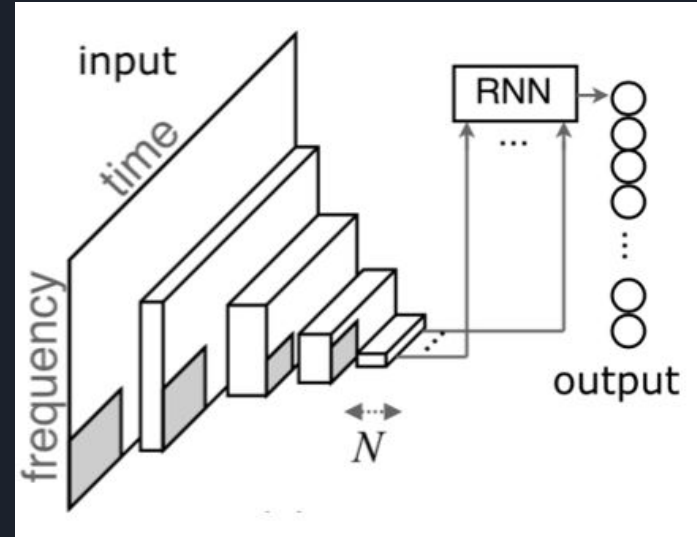 - Flatten layer before.
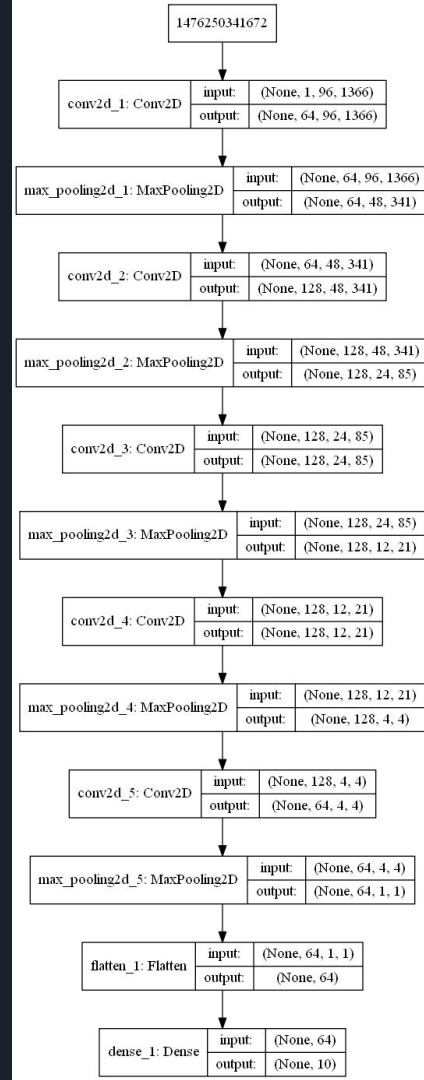
Results

Results - Comparison with previous work

Tzanetakis	61.00%
Holzapfel	74.00%
Benetos	75.00%
Lidy	76.80%
Bergstra	82.50%
Our model	52.00%

Future work

- Parameter optimization
- Implement a CNN + RNN

References

[1]: A. Tzanetakis, G. and Cook, P. "Musical genre classification ofaudio signal", IEEE Transactions on Speech and Audio Processing,Vol. 10, No. 3, pp. 293-302, July 2002.

[2]: Keunwoo Choi, George Fazekas, Mark Sandler, P. "Automatic Tagging using Deep Convolutional Neural Networks", 17th International Society for Music Information Retrieval Conference, New York, USA, 2016

Questions