Deep neural network for classifying metal subgenres

ELEC5305 Something Awesome Project Proposal

James Walsh

School of Electrical and Information Engineering
University of Sydney
Sydney, Australia
jwal7309@uni.sydney.edu.au

Abstract—**Music genre classification has traditionally been performed by inputting handcrafted audio features (e.g. MFCCs) into machine learning classifiers (e.g. SVM, k-NN) to make broad classifications. However, subgenres are harder to classify as audio features are less distinguishable within genre groups (see my initial research in Figure 3). I propose the use newer deep learning techniques that use the visual representation of an audio signal (spectrograms) as input to discover the features for best distinguishing between two extreme metal subgenres: black metal and death metal.**

Index Terms—**deep learning, music classification, audio analysis**

Fig. 1. A death metal album (Scream Bloody Gore by Death - 1987).

Fig. 2. A black metal album (Transylvanian Hunger by Darkthrone - 1994).

I. Background

Music classification (labelling songs with their corresponding genre) is useful for music streaming services and recommender systems. However, music is complex, and there are usually smarter ways than music classification to recommend new music to listeners (such as through network analysis of listener habits—i.e. 'other listeners liked...'—or online-activity monitoring). Thus, music classification is of most use to databases that lack existing genre labels and require an automatic labelling tool to get their music service started (before any listener habits have been learned). Here, two subgenres of extreme metal will be targeted for a binary classification problem using recent deep learning-based music classifiers.

- **Death Metal (1st-wave death metal)**—pioneered in the 80s in USA by bands such as Death, Obituary, Morbid Angel and Possessed, death metal is characterised by its heavy distortion, low tuned guitars, use of palm muting, blast beats and deep growling vocals (Figure 1).
- **Black Metal (2nd-wave black metal)**—pioneered in the 90s in Norway by bands such as Mayhem, Darkthrone, Emperor and Gorgoroth, black metal is characterised by its heavy distortion, high-pitched guitar tone, raw recording quality, dissonance and shrieking vocals (Figure 2).

An initial investigation has shown that audio features such as spectral roll-off do not offer a clear-enough delineation between black and death metal (Figure 3). Spectral centroid (not shown here) showed slightly better delineation between the two genres, but not how one would exptect: the death metal

songs actually produced higher spectral centroids. Nevertheless, this project will ignore these audio features in favour of a hidden deep learning process—as performed in [1]—using a convolutional neural network fed by 15-second long snippet song spectrograms.
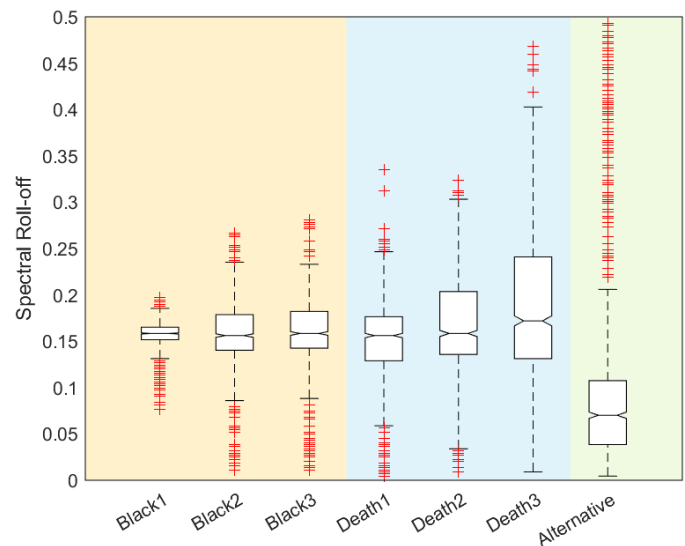

Fig. 3. Spectral roll-off (the frequency at which 90% of the signal energy is below) for various 20 second segments of black metal songs (Black1, Black2, Black3), death metal songs (Death1, Death2, Death3) and a radiohead song as a control (Alternative).

II. Credit Requirements

To achieve a credit for this project I will:

- Investigate a variety of deep learning techniques for music classification [2], [3].
- Prepare data for training the neural network.
- Choose and design an appropriate neural network architecture.
- Train the neural network.
- Test the neural network and visualise results using 2×2 confusion matrix.

III. Extension Marks

These additional points will be addressed if time permits:

- Investigate methods to normalise the data (i.e. removing high frequency components, normalising volume levels, etc.) in order to guide the neural network towards identifying features that human's use to classify the genres.
- Attempt to identify and characterise what audio features are seen as *important* by the neural network.
- Provide an example of how this network could be extended to multiple and nested genre classification tasks.

IV. Data Management

This project will be undertaken on MATLAB and stored locally and on GitHub. Reports will be produced on Overleaf and video presentations uploaded to Youtube. Music files will be downloaded from purchased albums on Bandcamp (where possible) and stored locally. Online open-source databases for music genres are woefully small (e.g. ISMIR 2004 and GTZAN Genre Collection) and obviously did not have existing subgenre labels.

References

[1] S. Oramas, F. Barbieri, O. Nieto, and X. Serra, "Multimodal Deep Learning for Music Genre Classification," *Transactions of the International Society for Music Information Retrieval*, vol. 1, no. 1, pp. 4–21, 2018, ISSN: 2514-3298. DOI: 10.5334/tismir.10.

[2] A. Elbir and N. Aydin, "Music genre classification and music recommendation by using deep learning," *Electronics Letters*, vol. 56, no. 12, pp. 627–629, 2020, ISSN: 0013-5194. DOI: 10.1049/el.2019.4202.

[3] L. Nanni, Y. M. G. Costa, R. L. Aguiar, C. N. Silla, and S. Brahnam, "Ensemble of deep learning, visual and acoustic features for music genre classification," *Journal of New Music Research*, vol. 47, no. 4, pp. 383–397, 2018, ISSN: 0929-8215. DOI: 10.1080/09298215.2018.1438476.