

Music Genre Classification

George Herbert
Department of Computer Science
University of Bristol
Bristol, United Kingdom
cj19328@bristol.ac.uk

Abstract—

Index Terms—music information retrieval, convolutional neural networks

I. INTRODUCTION

Music information retrieval (MIR) is the interdisciplinary science of retrieving information from music. Genre classification—classifying a sample of music into one or more genres—is a fundamental problem in MIR.

Schindler et al. [1] investigated performance differences of different convolutional neural network (CNN) architectures on the task of genre classification.

II. RELATED WORK

III. DATASET

The GTZAN dataset [2] contains 1000 WAV audio tracks, each 30 seconds in length. There are 100 tracks for each of the 10 genres in the dataset: blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae and rock.

The CNNs were not trained on the raw tracks. Each audio track was divided into chunks, with log-mel spectrograms produced for randomly selected chunks.

To produce a training and validation set, a stratified split was deemed suitable to prevent imbalance. Spectrograms for 75 of the 100 WAV audio tracks for each genre were randomly selected to make up the training set, with the spectrograms for the other 25 audio tracks for each genre making up the validation set.

IV. CNN ARCHITECTURE

V. IMPLEMENTATION DETAILS

VI. REPLICATING QUANTITATIVE RESULTS

VII. TRAINING CURVES

VIII. QUALITATIVE RESULTS

IX. IMPROVEMENTS

X. CONCLUSION AND FUTURE WORK

REFERENCES

- [1] A. Schindler, T. Lidy, and A. Rauber, “Comparing shallow versus deep neural network architectures for automatic music genre classification,” Nov. 2016.
- [2] G. Tzanetakis and P. Cook, “Musical genre classification of audio signals,” *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.