Progress Report for CSS 586: Modeling Latent Patterns in Music

Alex Kyllo
akyllo@uw.edu
University of Washington
Bothell, WA, USA

ABSTRACT

This report explores recent research in modeling music with deep learning and provides a progress report on a project to train a generative model of classical music on the MusicNet dataset.

KEYWORDS

deep learning, neural networks, music

1 INTRODUCTION

Machine learning models of music have interesting applications in music information retrieval and creative tools for musical artists and educators. Music is complex and challenging to model because it exhibits a hierarchy of recurring patterns.

Depending on the task, machine learning models of music may be trained on the audio signal itself, either in a time domain or a frequency domain representation, or they may be trained on a digital symbolic representation of music, the most common of which is MIDI (Musical Instrument Digital Interface) notation. MIDI is an encoding of music as streams of bytes in one or more tracks or channels, each representing a sequence of 128 possible pitch values, along with timing, pressure and instrument values. A music transcription model may convert an audio signal into MIDI (which can then be transcribed into sheet music), while a synthesizer model can convert MIDI representations into audio signals.

2 RELATED WORK

Google's Magenta is an umbrella project for music deep learning research and development of software tools to expose these models for use by creative artists and students.

MusicVAE is a variational LSTM autoencoder for MIDI that incorporates a novel hierarchical structure using a "composer" recurrent layer in its encoder model to better capture structure at multiple levels [2].

Music Transformer is a generative model that borrows its approach from the Natural Language Processing (NLP) domain, using an attention network to model MIDI music as a sequence of discrete tokens with relative positional dependencies [1].

3 PLANNED METHODS

4 PROGRESS

5 FUTURE WORK

REFERENCES

[1] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Ian Simon, Curtis Hawthorne, Andrew M. Dai, Matthew D. Hoffman, Monica Dinculescu, and Douglas Eck. 2018. Music Transformer. (2018). arXiv:1809.04281 http://arxiv.org/abs/1809.04281

[2] Adam Roberts, Jesse Engel, Colin Raffel, Curtis Hawthorne, and Douglas Eck. [n.d.]. A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music. In *Proceedings of the35th International Conference on Machine Learning* (2018). 10.