

Group 13 Project Proposal

Jahya Burke
ECE Department
UC San Diego
San Diego, USA
jlburke@ucsd.edu

Yisheng Ji
ECE Department
UC San Diego
San Diego, USA
y3ji@eng.ucsd.edu

Shiwei Rong
ECE Department
UC San Diego
San Diego, USA
srong@ucsd.edu

Shuangcheng Yang
ECE Department
UC San Diego
San Diego, USA
shy003@ucsd.edu

Alan Yessenbayev
ECE Department
UC San Diego
San Diego, USA
ayessenb@ucsd.edu

***Index Terms*—component, formatting, style, styling, insert**

I. PROBLEM

For our project, we will address the task of creating a neural composer. For thousands of years, cultures around the world composed music for recreational, religious, and artistic reasons. Up until now, music has been composed by humans but with the rise of machine learning there are opportunities for music to be composed using learning algorithms. Aside from the cultural benefits of music, there is now a large commercial industry surrounding the music composition for entertainment venues, movies, and video games. If a successful music composing application were to be created, it could be used for artistic as well as commercial benefits. Current endeavors like NSynth by Google Magenta team are used for creating new sounds [1].

II. DATASET

In order to train our neural network, we will require large datasets of music. The datasets will need to be restricted to a specific genre to avoid confusing the network with wide variation in compositional style. Based on examples that we have seen [references], we expect that different types of networks will perform well for different genres of music. In order to allow for flexibility in the type of network we construct, we will consider composing 3 different genres of music; jazz music, which tends to be sporadic and free form, classical music, which tends to be dynamic and structured, and video game music, which tends to be formulaic and looping.

III. SOLUTION

We will generate an auto-encoder to achieve our goal. For the encoder network part, a 96 notes with a time step of 96 ticks structure will be generated as the training dataset. Additionally, a third dimension for the measure itself may also be established due to measure often repeats. The number of third dimension is calculated based on the number of measure songs. PCA or LDA method will be used to compress the high-dimensional training data to a reasonable dimension for us to train our model. A supervised learning model of deep neural network called neural ordinary differential equations solver[1] is used. Instead of specifying a discrete sequence

of hidden layers, this model will parameterize the derivative of the hidden state using a neural network and the adjoint sensitivity method is used to calculate the gradients. The output of the network is computed using a black-box ordinary differential equation solver. As a result, we will generate a composing AI interface. User can directly change a certain number of top principal components to change the style of music at any time. Also, notes in the piano roll format will be shown to the user. There may have other useful controls inside the interface to compose various of music.

IV. EXPERIMENTS

REFERENCES

- [1] (2017, Apr.) Neural audio synthesis of musical notes with wavenet autoencoders. [Online]. Available: <https://arxiv.org/abs/1704.01279>