**CS236781 – Deep Learning on Computational Accelerators**

Final Project

(Winter 2019-2020)

**DDSP**

Yakir Helets and Gal Fleissig

**Abstract**

**Intro**

The paper relevant to out project is [*DDSP: Differentiable Digital Signal Processing*](https://arxiv.org/abs/2001.04643), by Jesse Engel, Lamtharn Hantrakul, Chenjie Gu, & Adam Roberts of the Brain Team at Google Research, published as a conference paper at ICLR 2020.

The domain of this paper is the *Digital Signal Processing* (DSP, without the first “D”), which is a domain that is integral to modern society in many aspects such as telecommunications, transportation, audio and medical technologies. This paper brings a new approach to the table, by combining deep learning techniques, and by that improves upon existing approaches.

There are three existing approaches prior to this paper, all of which are part of the family of audio that belongs to generative models. The first generates samples in the domain of time. The second generates samples in the domain of frequency. These two approaches can express any signal, but they are inefficient because they do not take advantage of existing knowledge on sound properties. The third approach which is vocoders or synthesizers does utilize existing knowledge on sound properties, however since it has limited expressivity and it is difficult to adapt it to modern auto differentiation machine learning methods – it has been less actively researched.

The contribution of the relevant paper is by introducing a new method called *Differentiable Digital Signal Processing* (or DDSP), along with its library. The new method uses a neural network to convert input into complex DSP controls that can produce more realistic signals. This method unlike the previous methods enables for direct integration of classic signal processing elements using deep learning methods. In other words, the DDSP units are differentiable (hence the name) which allows to train the neural network to learn a new dataset using the known process of backpropagation, and by that to achieve better results than existing approaches.

The paper focuses on audio synthesis and by that achieves high quality audio generation without the need for large models or adversarial losses. Additionally, it shows that combining different modules allows for manipulation of each of the model components individually. And in general, the DDSP approach towards the domain is interpretable and modular while still using the benefits of deep learning techniques.

Some of the interesting applications of this approach include independent control of pitch and loudness, transfer of the acoustics of one room to another and transfer of audio signals from one domain to another (e.g. singing to a violin).

We are planning to …….. (maybe not necessary)

**Methods**

**Implementation and Experiments**

**Results**