Estimating Synthesizer Parameters with Neural Networks
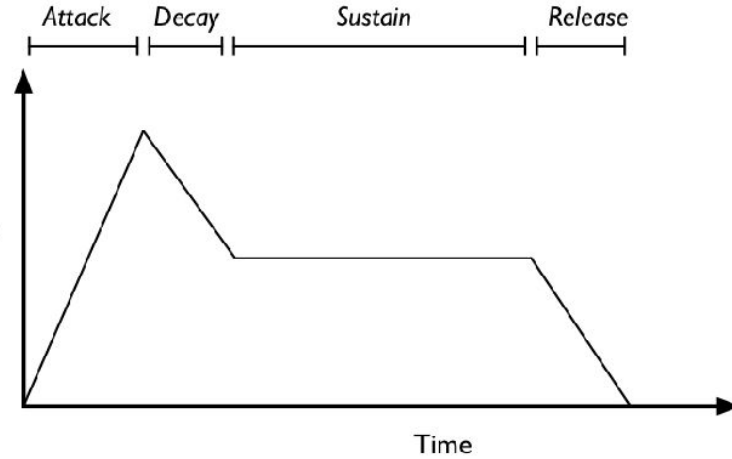
•••

Kaleb Tsegaye

What is a synthesizer?

- Musical instrument, became popular throughout mid-to-late 20th century
- Makes sound by generating electronic signals
- Nowadays, there are a plethora of software synths

Attack | Decay | Sustain | Release

Time

Altering a synthesizer's sound

- Oscillators - generates the sound you hear
  - Common waveforms: Sine, square, sawtooth, triangle
- Envelopes
  - Attack, decay, sustain, release (ADSR)
  - Can be applied to amplitude, pitch, etc.
- Many other features/effects
  - Low frequency oscillators (LFO), filters, panning, etc.

Problem Statement

- How accurately can convolutional neural networks classify oscillator waveforms and estimate amplitude envelopes from audio samples?

Methodology

- **[18alantom 'synth' project repository](#)**
  - A synthesizer written in Python
  - Slightly altered for purposes of my project
- ~2000 samples created
  - Four waveforms: sine, square, sawtooth, triangle
  - Between 1 and 4 present per sample
  - 10 seconds long, ADSR randomized
- Intended results:
  - What waveforms are present?
  - What ADSR settings are being used?
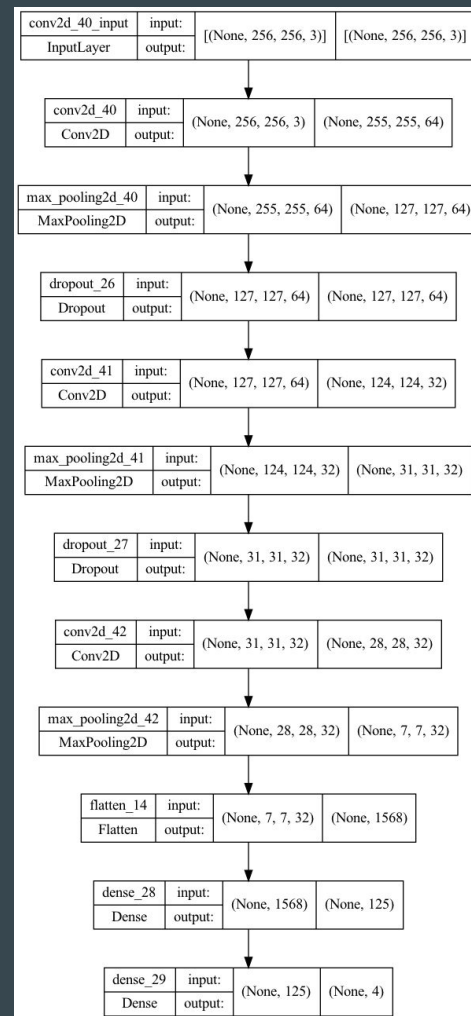   - Durations
    - Sustain volume
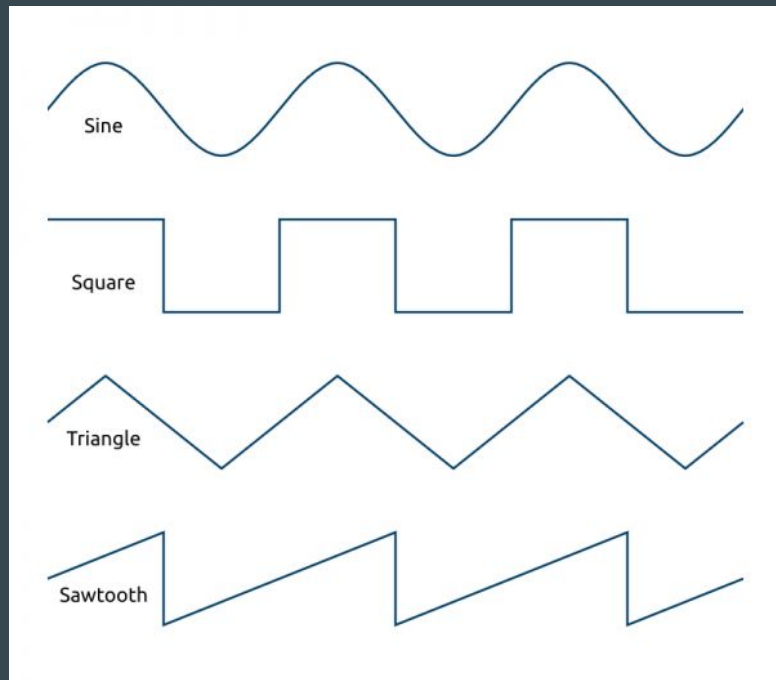
Representing Audio as Images - Spectrograms/Amplitude Plots
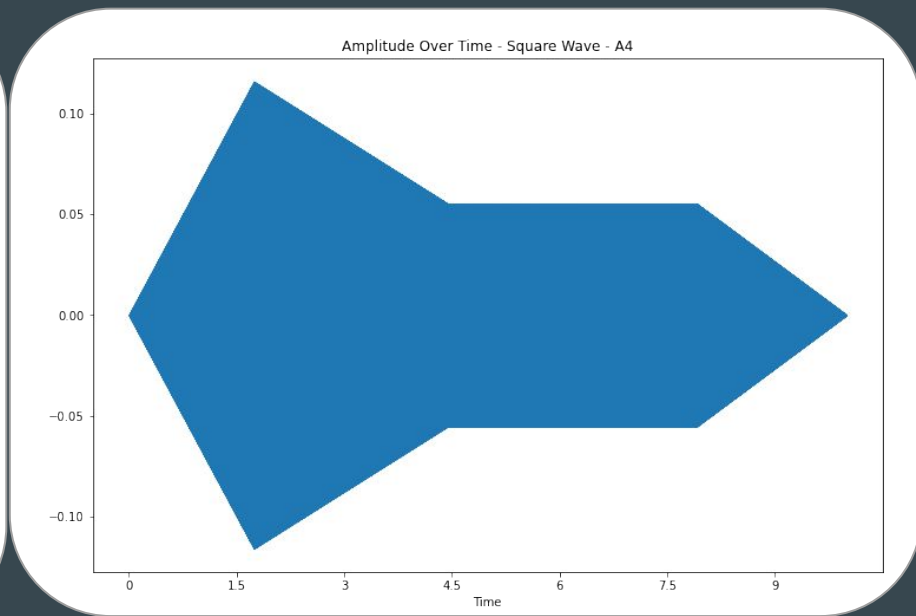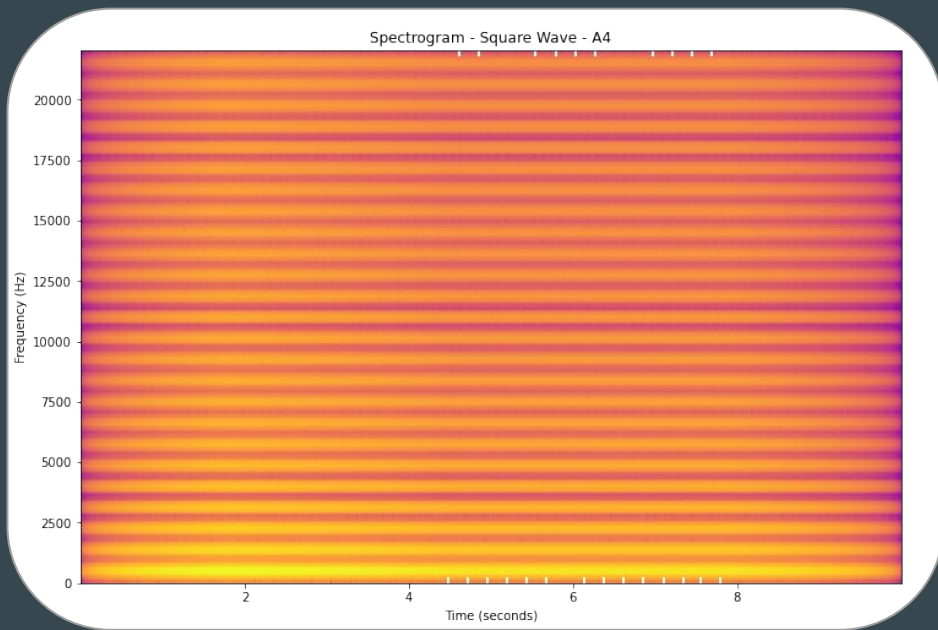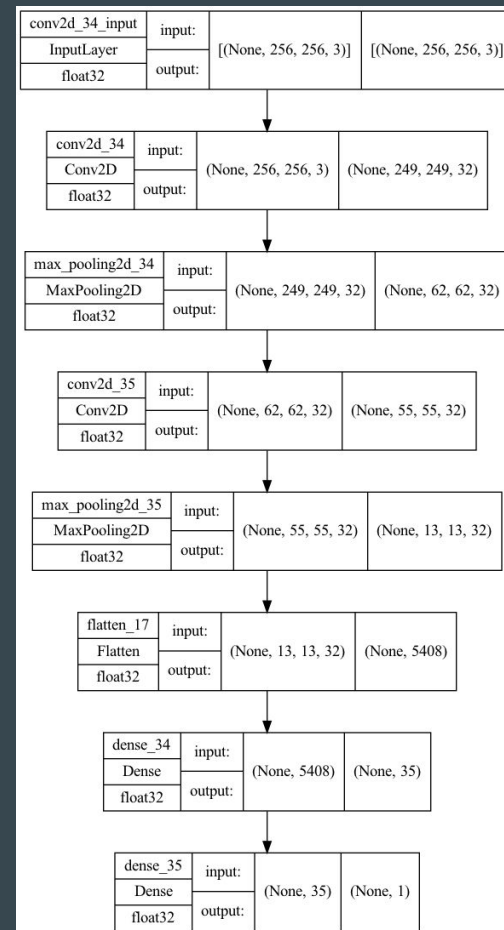
CNN - Multi-label Classification of Waveforms

	Score: Test Set
Recall	78.6%
Precision	77.7%
AUC	77.3%

CNNs - Estimating ADSR Envelope

	Duration Model: Test Set	Sustain Level Model: Test Set
Mean Squared Error	0.383*	0.008
Mean Absolute Error	0.436	0.055*

*Loss function for Model

Streamlit Demo