Building style-aware neural MIDI synthesizers using simplified differentiable DSP approach

Sergey Grechin, Ryan Groves

@Infinite Album https://infinitealbum.io

Goals

- Going as simple as possible,
- Create a model that can **“learn” the sound of real instruments**
- MIDI controllable
- Client-side inference in mind
- Capture the **“style”** of playing
- Develop a workflow, not only a model
- Apply the approach to modelling of electric guitar sound

- *DDSP was a way to go*

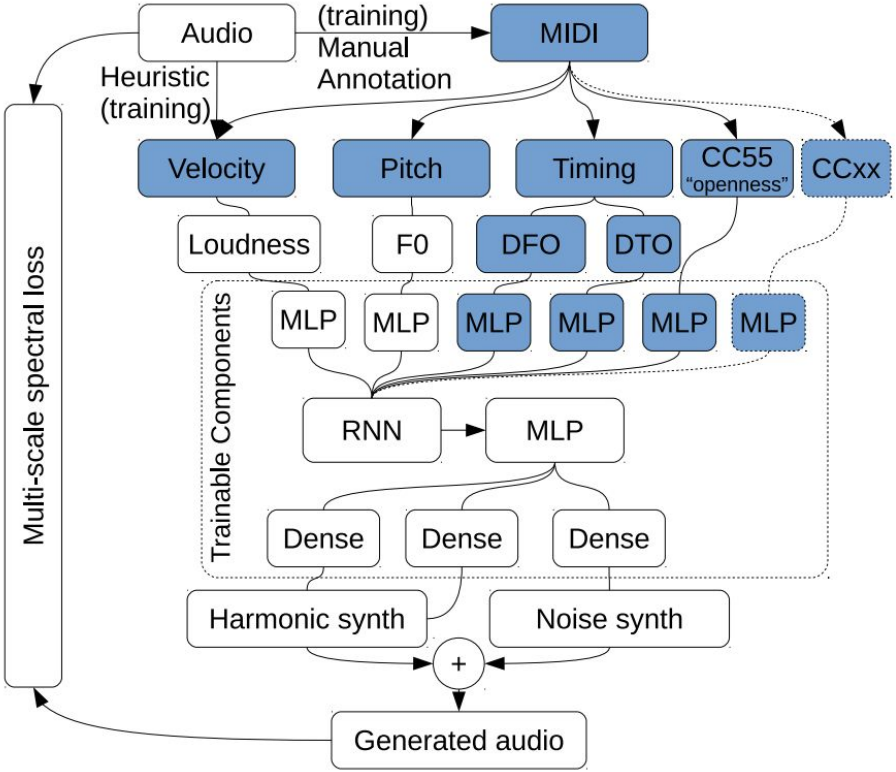
Existing research

Towards realistic MIDI instrument synthesizers

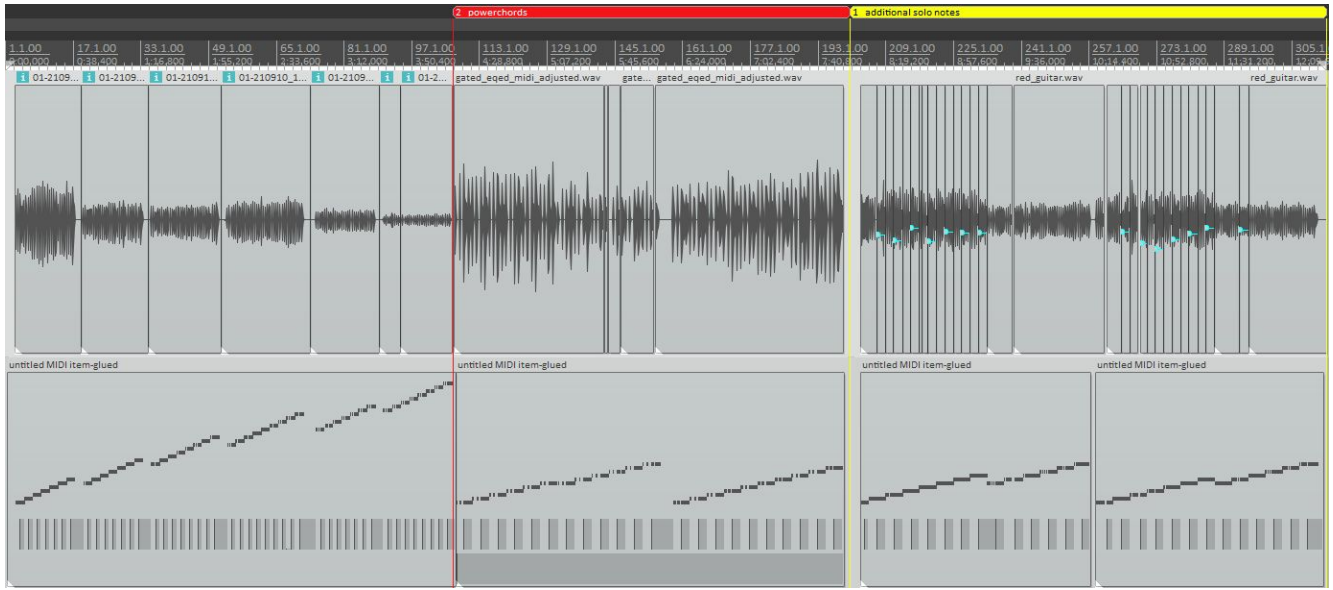
The control-synthesis approach for making expressive and controllable neural music synthesizers

MIDI-DDSP: DETAILED CONTROL OF MUSICAL PERFORMANCE VIA HIERARCHICAL MODELING

Model

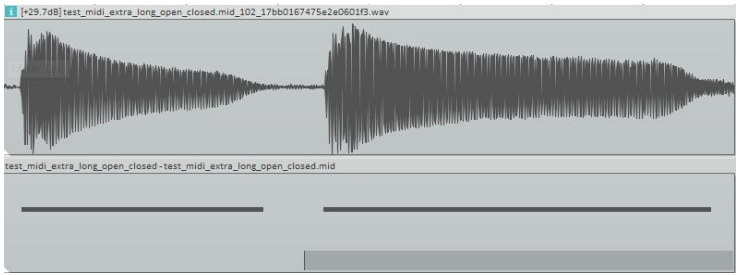

Audio
(training)
Manual
Annotation
MIDI
Heuristic
(training)
Velocity
Pitch
Timing
CC55
"openness"
CCxx
Loudness
F0
DFO
DTO
MLP
MLP
MLP
MLP
MLP
MLP
Trainable Components
RNN
MLP
Dense
Dense
Dense
Harmonic synth
Noise synth
+
Generated audio
Multi-scale spectral loss

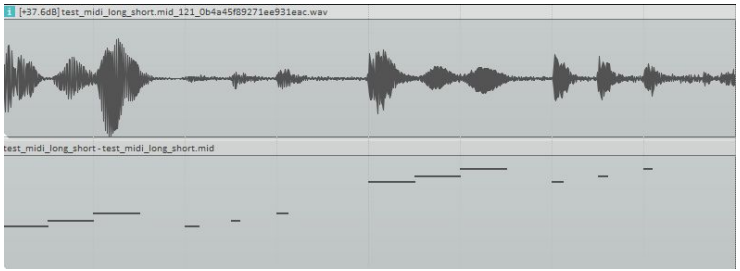
Dataset



Manually annotating in Reaper. Algorithmically augmented: pitch +- 3 semitones

Reconstruction examples





Conclusions/Future work

- The model does learn something
- Future
  - Learning chord sound
    - More fine-grained loss needed
  - Other style features for different instruments
    - Vibrato etc

Thank you!

grechin.sergey@gmail.com

Repo:
https://github.com/hq9000/neural-midi-synthesizer

Dataset:
https://drive.google.com/drive/folders/10wBXOffseRzjnAhv7dg6Ha71VF_t6BoJ

Online supplement:
https://grechin.org/neural_synthesizers_with_simplified_ddsp.html