Project Proposals

Jacopo Pegoraro

Jacopo.pegoraro@unipd.it

Daniele Mari

Daniele.mari@phd.unipd.it

Riccardo Mazzieri

riccardo.mazzieri@phd.unipd.it

2023-2024

Symbolic-domain music generation

- **Reference paper:** Yang, Li-Chia, Szu-Yu Chou, and Yi-Hsuan Yang. "MidiNet: A convolutional generative adversarial network for symbolic-domain music generation." arXiv preprint arXiv:1703.10847 (2017).

Symbolic-domain music generation

- Music generation is the task of generating a piece of music using a neural network.
- The most common approach in the literature is to generate and learn from **symbolic representations** of musical pieces (the set of notes played in the musical piece, the instruments, tempo, etc…), and to synthesize (or play) the actual music at a later stage.
- Those symbolic representations are most commonly stored in **MIDI (Musical Instrument Digital Interface) files**

MIDI files

MIDI files contain musical data, instead of audio.

- They are split in multiple **channels,** one for each instrument
- Each channel contains the ordered set of **musical notes** played by the instrument (the *piano roll*), each with its own duration and intensity (*velocity*)
- Contain useful **metadata** about the musical piece:
  - Tempo (BPM)
  - Key Signature
  - Time Signature

MIDI files

The *piano roll* can be represented as a matrix of shape M by N.
Each entry of the matrix is filled if the note is present or not. The value of the entry can also represent the intensity of the musical note.

M = number of notes (ordered from low to high pitches)

MIDI piano roll

N = Number of timesteps
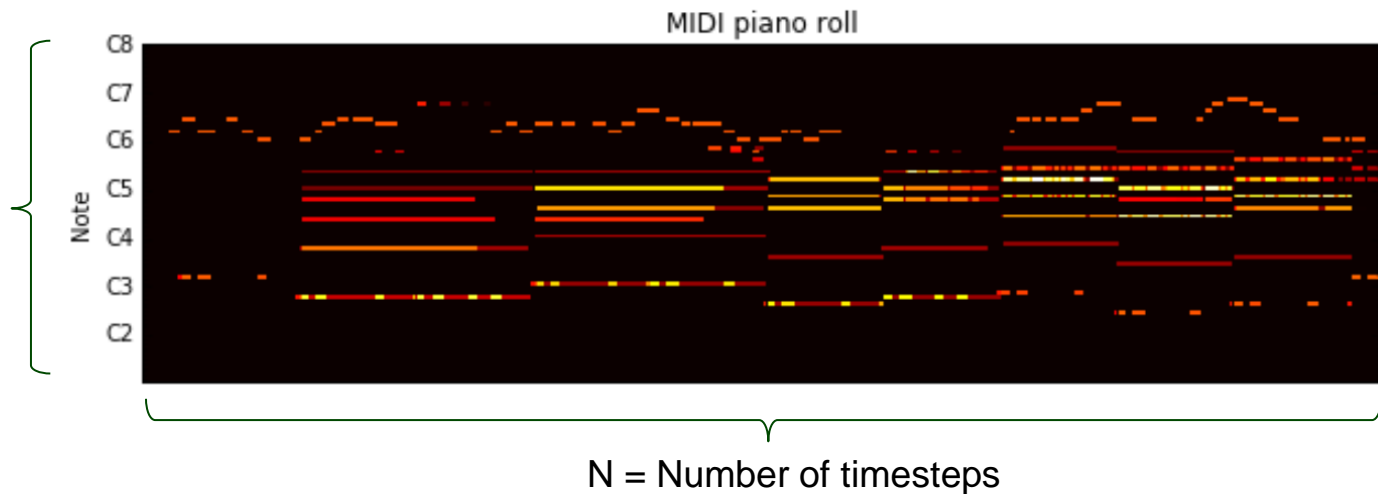
Datasets

You can choose among different datasets. Our suggestions are:

- Lakh MIDI Dataset: The Lakh MIDI dataset is a collection of 176,581 unique MIDI files, 45,129 of which have been matched and aligned to entries in the Million Song Dataset. We suggest to use the "Clean-MIDI subset".
- MAESTRO Dataset: MAESTRO (MIDI and Audio Edited for Synchronous TRacks and Organization) is a dataset composed of about 200 hours of virtuosic piano performances.

These dataset also provide alignments of MIDI files and audio waveforms, but you will need just the MIDI files for the project.

Suggested tasks

Possible tasks for the project are (in order of difficulty):
- Melody generation*
- Melody generation conditioned on previous notes*
- Melody generation conditioned on underlying chords*
- Musical piece generation (melody + chords)

* tackled by the reference paper

Suggestions

1. Spend time on data exploration and preprocessing:
   - Find strategies to isolate melodies or chords from the complete MIDI files
   - Explore data augmentation strategies
2. Try to use some complex CNN architectures for processing the piano rolls, (e.g. Inception or ResNet).
3. You might try different generative architectures such as VAE

Note: some background knowledge on music theory might be useful to better understand the data you are working with and literature on the topic.

Musical terminology (simplified)

- **Melody**: a melody is a series of notes laid out in time. In a melody, only a single note is playing at each timestep.
- **Chords**: chords are musical events where multiple notes play at the same time, forming vertical stacks of notes.
- **Harmony**: the specific sequence of chords in a piece of music determines its harmony.

The overlapping of a melody with underlying chords forms a *very basic* definition of a musical piece. (Note that more complex pieces of music might not follow this structure)

Example: a sequence of *chords*

Example: a sequence of single notes forms a *melody*