A GAN BASED DRUM PATTERN GENERATION UI PROTOTYPE

Hamid Eghbal-zadeh*[1] **Richard Vogl***[1,2] **Gerhard Widmer**[1] **Peter Knees**[2]
* Equal contribution
[1] Institute of Computational Perception, JKU Linz, Austria
[2] Faculty of Informatics, TU Wien, Vienna, Austria
hamid.eghbal-zadeh@jku.at, richard.vogl@tuwien.ac.at

ABSTRACT

Tools which support artists with the generation of drum tracks for music productions have gained popularity in recent years. Especially when utilizing digital audio workstations, using e.g. predefined drum loops can speed up drafting a musical idea. However, regarding variation and originality of drum patterns, methods based on generative models would be more suited for this task. In this work we present a drum pattern generation prototype based on a step sequencer interface which allows to explore the suitability of generative adversarial networks for this task.

1. INTRODUCTION

Automatic music generation has been an active field of research for several years. While early attempts mostly use rule-based and probabilistic models, recent systems focus on machine learning and artificial intelligence based methods [6,14]. These methods can be divided into two groups: *i)* methods focusing on generating audio signals and *ii)* methods that create symbolic music. Generative systems that focus on full music tracks may find application in the context of media arts or for automatic sound track generation e.g. for video games. In this work we focus on generating only drum tracks which is a relevant task in the context of digital music production. Generating a symbolic drum track can be a labor intensive task, while involving repetitive steps. To draft musical ideas quickly, it is often desirable to have a simple and fast method to create a basic drum track. Some digital audio workstations (DAWs) and drum sequencers thus provide predefined drum loops to enable such a workflow. Apple's Logic Pro X DAW features a so called *Drummer* plugin to create drum tracks, which allows to interactively vary loudness and complexity, besides other parameters. While for the early stages of writing music, these approaches help to speed up the process, musicians and producers often refrain from using these technologies for the end product. This is due to the fact that out-of-the-box patterns bear the danger of sounding unoriginal.

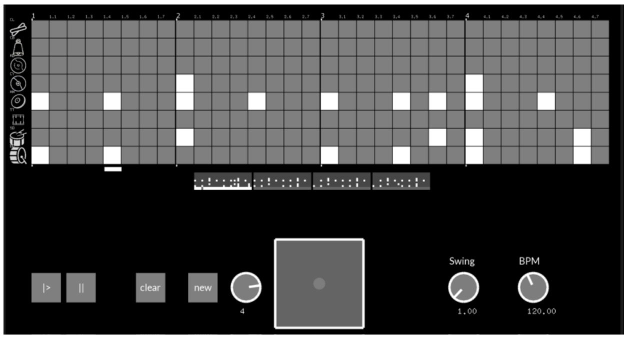

Figure 1. UI of the proposed prototype.

To incorporate more natural variation and thus overcome these downsides, generative approaches for drum tracks can be used. Early methods for symbolic drum track generation built on genetic algorithms (GAs) [5, 7, 8, 11]. Vogl et al. [12] use restricted Boltzmann machines (RBMs) to achieve this. Recent works show that generative adversarial networks (GANs [4]) can be used for music generation; both for symbolic music [3, 14] as well as audio [2, 10]. In this work, GANs are used to directly generate symbolic drum tracks, parameterizable and controllable in a similar fashion as the *Drummer* plugin in Logic Pro X.

2. METHOD

The drum pattern variation engine is implemented using a GAN. The network is trained on sets of 8-by-32 matrix representations of drum patterns. Each of these matrices represent a bar of music, while the eight rows represent different drum instruments and the 32 columns individual discrete time steps (32nd notes). For this work dynamics (changes in volume for each onset) are ignored as a simplification.

2.1 Network Structure and Training

The structure of the GAN is based on a convolutional recurrent design. Four deconvolution layers with 3x3 filters are used to generate a bar (8x32 matrix), while recurrent connections over bars are used to model the temporal evolution of patterns. This way, a varying number of bars can be considered during training or inference.

As loss for GAN training a Wasserstein loss [1] in combination with additional conditioning on genre, and two features extracted from the one-bar drum patterns (complexity and loudness) are used.

2.2 Training Data

A MIDI dataset of popular songs published as part of [13] serves as training data for the GAN. First, drum tracks were extracted from the full MIDI files and then converted to the 8-by-32 matrix per bar representation. The genre tags for this data were created by parsing the MIDI file names to extract artist name and utilizing an artist to genre lookup (*Spotify* API [1]). Additionally a large scale genre dataset of two-minute electronic dance music (EDM) samples was used [9]. To extract symbolic drum tracks, a state-of-the-art drum transcription system [13] was applied.

3. UI

Figure 1 shows the basic user interface of the prototype. The main area features a eight-by-32 step-sequencer grid. Step sequencers have been shown to be an effective input method for drum pattern creation, however, they only allow a discrete time grid. Beneath the step-sequencer grid, an x/y pad allows control of complexity and loudness for pattern creation. Additionally controls for playback control, genre, tempo, and swing ratio allow customization of the drum patterns.

4. CONCLUSION

In this work we present a GAN based drum pattern generation prototype, trained using large scale drum pattern datasets. These datasets were extracted from MIDI songs and two minute audio excerpts utilizing a drum transcription system. The UI follows well-known interface paradigms used for drum pattern generation.

5. ACKNOWLEDGMENTS

This work has been partly funded by the Austrian FFG under the BRIDGE 1 project *SmarterJam* (858514), as well as by the Austrian Ministries BMVIT and BMWFW, and the Province of Upper Austria, via the COMET Center SCCH.

6. REFERENCES

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, volume 70 of *Proceedings of Machine Learning Research*, pages 214–223, 2017.

[2] Chris Donahue, Julian McAuley, and Miller Puckette. Synthesizing audio with generative adversarial networks. *arXiv preprint arXiv:1802.04208*, 2018.

[3] Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. Musegan: Demonstration of a convolutional gan based model for generating multi-track piano-rolls. In *Late Breaking/Demos, 18th International Society for Music Information Retrieval Conference (ISMIR)*, 2017.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[5] Damon Horowitz. Generating rhythms with genetic algorithms. In *Proceedings of the 12th National Conference on Artificial Intelligence*, volume 94, page 1459, 1994.

[6] Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Andrew Dai, Matt Hoffman, Curtis Hawthorne, and Douglas Eck. Generating structured music through self-attention. In *The 2018 Joint Workshop on Machine Learning for Music*, 2018.

[7] Maximos A. Kaliakatsos-Papakostas, Andreas Floros, Nikolaos Kanellopoulos, and Michael N. Vrahatis. Genetic evolution of l and fl-systems for the production of rhythmic sequences. In *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation*, GECCO '12, pages 461–468, 2012.

[8] Maximos A. Kaliakatsos-Papakostas, Andreas Floros, and Michael N. Vrahatis. evodrummer: Deriving rhythmic patterns through interactive genetic algorithms. In *Evolutionary and Biologically Inspired Music, Sound, Art and Design*, volume 7834 of *Lecture Notes in Computer Science*, pages 25–36. 2013.

[9] Peter Knees, Angel Faraldo, Perfecto Herrera, Richard Vogl, Sebastian Böck, Florian Hörschläger, and Mickael Le Goff. Two data sets for tempo estimation and key detection in electronic dance music annotated from user corrections. In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, 2015.

[10] Narain Krishnamurthy. Beatgan - generating drum loops via gans. `https://github.com/NarainKrishnamurthy/BeatGAN2.0`, 2018.

[11] Cárthach Ó Nuanáin, Perfecto Herrera, and Sergi Jordà. Target-based rhythmic pattern generation and variation with genetic algorithms. In *Proceedings of the 12th Sound and Music Computing Conference*, 2015.

[12] Richard Vogl and Peter Knees. An intelligent drum machine for electronic dance music production and performance. In *Proceedings of the 17th International Conference on New Interfaces for Musical Expression (NIME)*, Copenhagen, 2017.

[13] Richard Vogl, Gerhard Widmer, and Peter Knees. Towards multi-instrument drum transcription. In *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18)*, 2018.

[14] Li-Chia Yang, Szu-Yu Chou, and Yi-Hsuan Yang. Midinet: A convolutional generative adversarial network for symbolic-domain music generation. *arXiv preprint arXiv:1703.10847*, 2017.

[1] urlhttps://developer.spotify.com/console