DEEP LEARNING FOR MUSIC GENERATION

This file presents the State of the Art of Music Generation.

README.pdf

Table of Contents

2. Algorithmic Composition

1992

HARMONET

Hild, H., Feulner, J., & Menzel, W. (1992). HARMONET: A neural net for harmonizing chorales in the style of JS Bach. In Advances in neural information processing systems (pp. 267-274). Paper

Books

- Westergaard, P. (1959). Experimental Music. Composition with an Electronic Computer.
- Cope, D. (2000). The algorithmic composer (Vol. 16). AR Editions, Inc..
- Nierhaus, G. (2009). Algorithmic composition: paradigms of automated music generation. Springer Science & Business Media.
- Müller, M. (2015). Fundamentals of music processing: Audio, analysis, algorithms, applications. Springer.
- McLean, A., & Dean, R. T. (Eds.). (2018). The Oxford handbook of algorithmic music. Oxford University Press.

2. Neural Network Architectures

NN Architecture	Year	Authors	Link to original paper	Slides

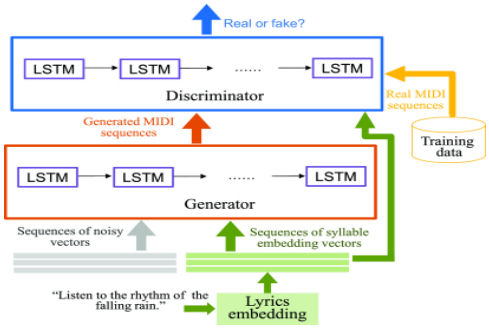
NN Architecture	Year	Authors	Link to original paper	Slides
Long Short-Term Memory (LSTM)	1997	Sepp Hochreiter, Jürgen Schmidhuber	http://www.bioinf.jku.at/publications/older/2604.pdf	LSTM.pdf
Convolutional Neural Network (CNN)	1998	Yann LeCun, Léon Bottou, YoshuaBengio, Patrick Haffner	http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf	
Variational Auto Encoder (VAE)	2013	Diederik P. Kingma, Max Welling	https://arxiv.org/pdf/1312.6114.pdf	
Generative Adversarial Networks (GAN)	2014	Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio	https://arxiv.org/pdf/1406.2661.pdf	
Transformer	2017	Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin	https://arxiv.org/pdf/1706.03762.pdf	

3. Deep Learning Models for Music Generation

2021

Melody Generation from Lyrics

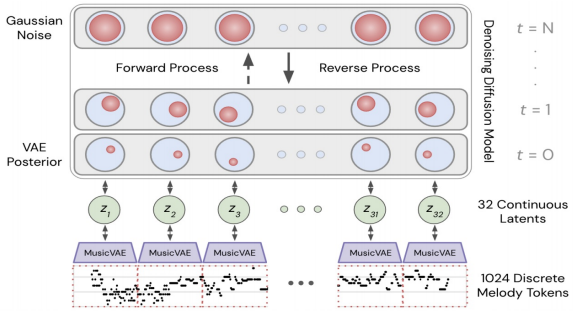
Yu, Y., Srivastava, A., & Canales, S. (2021). Conditional lstm-gan for melody generation from lyrics. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 17(1), 1-20.



Paper

Music Generation with Diffusion Models

Mittal, G., Engel, J., Hawthorne, C., & Simon, I. (2021). Symbolic music generation with diffusion models. arXiv preprint arXiv:2103.16091.

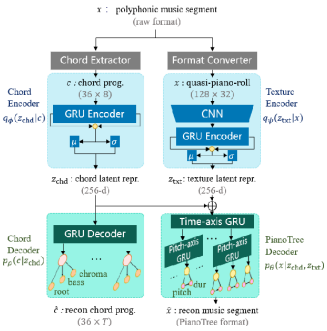


Paper

2020

Controllable Polyphonic Music Generation

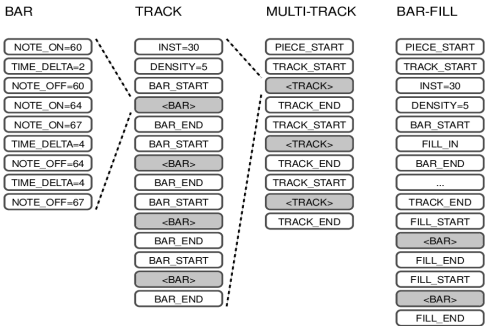
Wang, Z., Wang, D., Zhang, Y., & Xia, G. (2020). Learning interpretable representation for controllable polyphonic music generation. arXiv preprint arXiv:2008.07122.



Paper Web Video

MMM: Multitrack Music Generation

Ens, J., & Pasquier, P. (2020). Mmm: Exploring conditional multi-track music generation with the transformer. arXiv preprint arXiv:2008.06048.



Paper Web Colab Github (AI Guru)

Transformer-XL

Wu, X., Wang, C., & Lei, Q. (2020). Transformer-XL Based Music Generation with Multiple Sequences of Time-valued Notes. arXiv preprint arXiv:2007.07244.

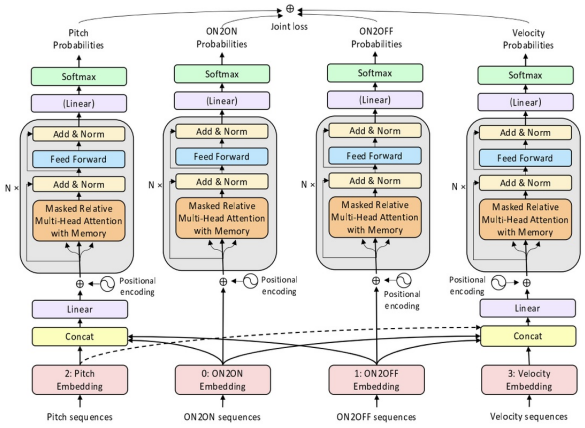
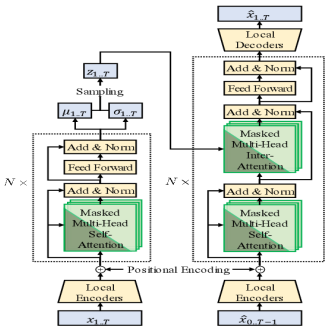


Figure 2: Our proposed framework with four transformer-xl networks.

Paper

Transformer VAE

Jiang, J., Xia, G. G., Carlton, D. B., Anderson, C. N., & Miyakawa, R. H. (2020, May). Transformer vae: A hierarchical model for structure-aware and interpretable music representation learning. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 516-520). IEEE.

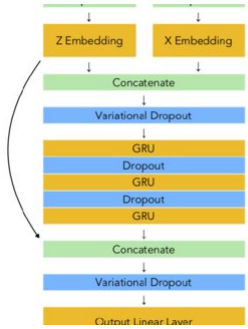


Paper

2019

TonicNet

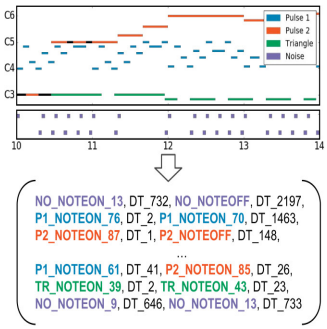
Peracha, O. (2019). Improving polyphonic music models with feature-rich encoding. arXiv preprint arXiv:1911.11775.



Paper

LakhNES

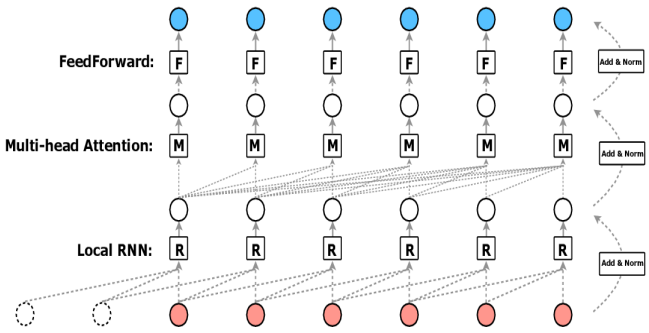
Donahue, C., Mao, H. H., Li, Y. E., Cottrell, G. W., & McAuley, J. (2019). LakhNES: Improving multi-instrumental music generation with cross-domain pre-training. arXiv preprint arXiv:1907.04868.



Paper

R-Transformer

Wang, Z., Ma, Y., Liu, Z., & Tang, J. (2019). R-transformer: Recurrent neural network enhanced transformer. arXiv preprint arXiv:1907.05572.

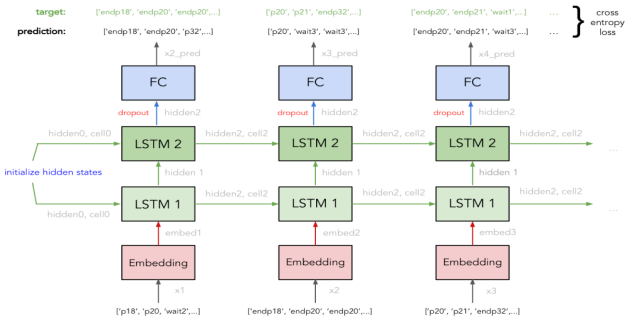


Paper

MuseNet - OpenAI

Web

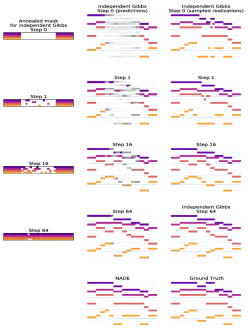
Maia Music Generator



Web

Coconet: Counterpoint by Convolution

Huang, C. Z. A., Cooijmans, T., Roberts, A., Courville, A., & Eck, D. (2019). Counterpoint by convolution. arXiv preprint arXiv:1903.07227.

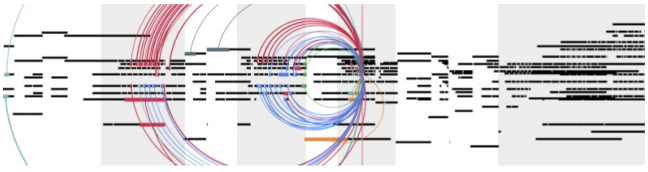


Paper Web

2018

Music Transformer - Google Magenta

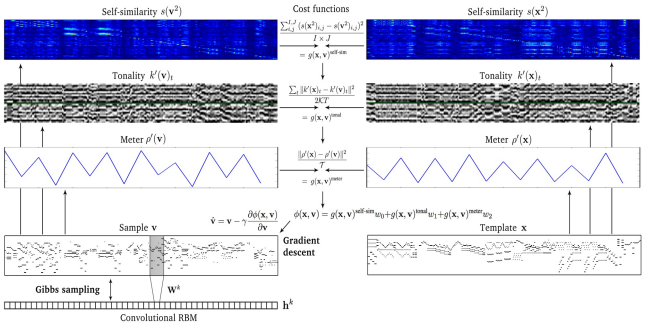
Huang, C. Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, et al. (2018). Music transformer. arXiv preprint arXiv:1809.04281.



Web Poster Paper

Imposing Higher-level Structure in Polyphonic Music

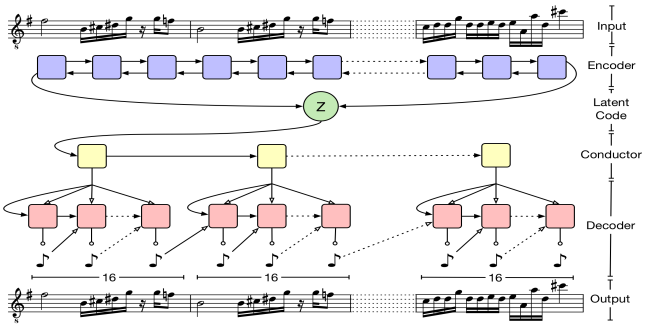
Lattner, S., Grachten, M., & Widmer, G. (2018). Imposing higher-level structure in polyphonic music generation using convolutional restricted boltzmann machines and constraints. Journal of Creative Music Systems, 2, 1-31.



Paper

MusicVAE - Google Magenta

Roberts, A., Engel, J., Raffel, C., Hawthorne, C., & Eck, D. (2018, July). A hierarchical latent vector model for learning long-term structure in music. In International Conference on Machine Learning (pp. 4364-4373). PMLR.

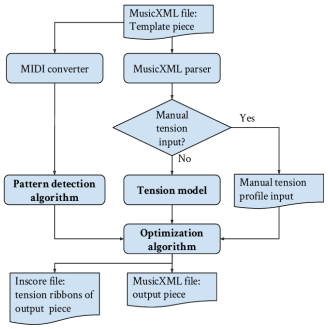


Web Paper Code Google Colab Explanation

2017

MorpheuS

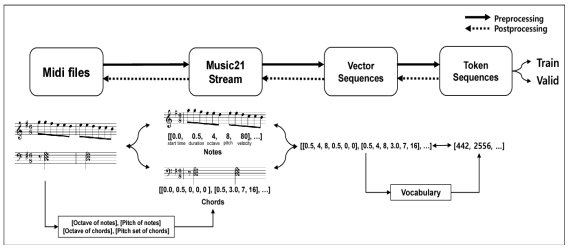
Herremans, D., & Chew, E. (2017). MorpheuS: generating structured music with constrained patterns and tension. IEEE Transactions on Affective Computing, 10(4), 510-523.

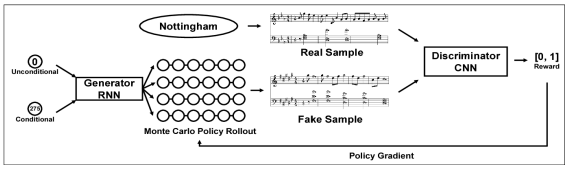


Paper

Polyphonic GAN

Lee, S. G., Hwang, U., Min, S., & Yoon, S. (2017). Polyphonic music generation with sequence generative adversarial networks. arXiv preprint arXiv:1710.11418.

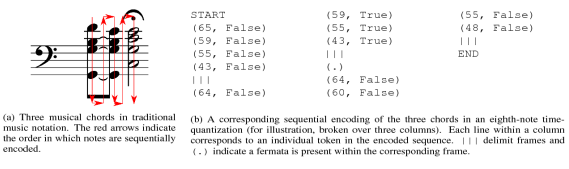




Paper

BachBot - Microsoft

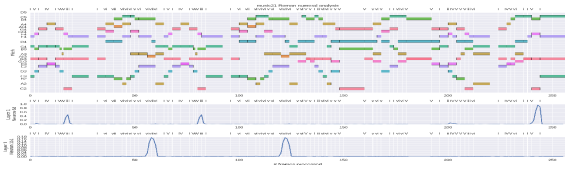
Liang, F. T., Gotham, M., Johnson, M., & Shotton, J. (2017, October). Automatic Stylistic Composition of Bach Chorales with Deep LSTM. In ISMIR (pp. 449-456).



(a) Three musical chords in traditional music notation. The red arrows indicate the order in which notes are sequentially encoded.

```
START           (59, True)      (55, False)
(65, False)     (55, True)      (48, False)
(59, False)     (43, True)      |||
(55, False)     |||             END
(43, False)     (.)
|||             (64, False)
(64, False)     (60, False)
```

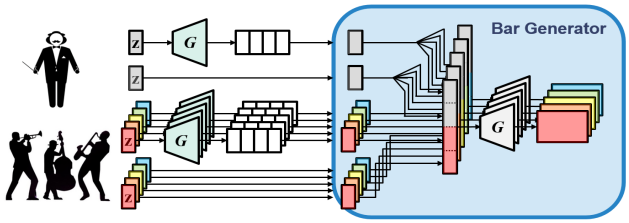
(b) A corresponding sequential encoding of the three chords in an eighth-note time-quantization (for illustration, broken over three columns). Each line within a column corresponds to an individual token in the encoded sequence. ||| delimit frames and (.) indicate a fermata is present within the corresponding frame.



Paper Liang Master Thesis 2016

MuseGAN

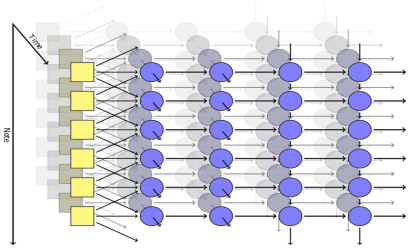
Dong, H. W., Hsiao, W. Y., Yang, L. C., & Yang, Y. H. (2018, April). Musegan: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).



Web Paper Poster GitHub

Composing Music with LSTM

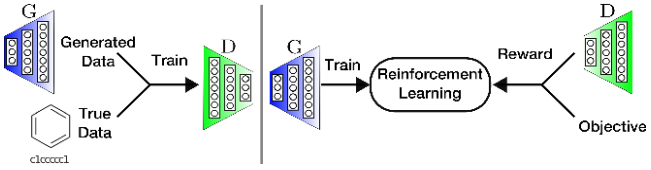
Johnson, D. D. (2017, April). Generating polyphonic music using tied parallel networks. In International conference on evolutionary and biologically inspired music and art (pp. 128-143). Springer, Cham.



Paper Web GitHub Blog

ORGAN

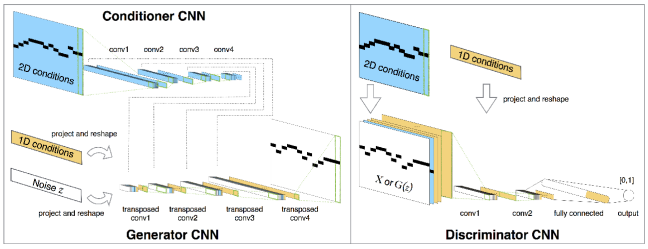
Guimaraes, G. L., Sanchez-Lengeling, B., Outeiral, C., Farias, P. L. C., & Aspuru-Guzik, A. (2017). Objective-reinforced generative adversarial networks (ORGAN) for sequence generation models. arXiv preprint arXiv:1705.10843.



Paper

MidiNet

Yang, L. C., Chou, S. Y., & Yang, Y. H. (2017). MidiNet: A convolutional generative adversarial network for symbolic-domain music generation. arXiv preprint arXiv:1703.10847.

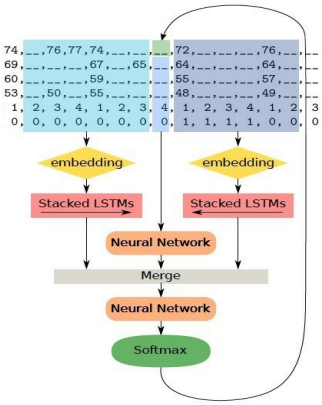


Paper

2016

DeepBach

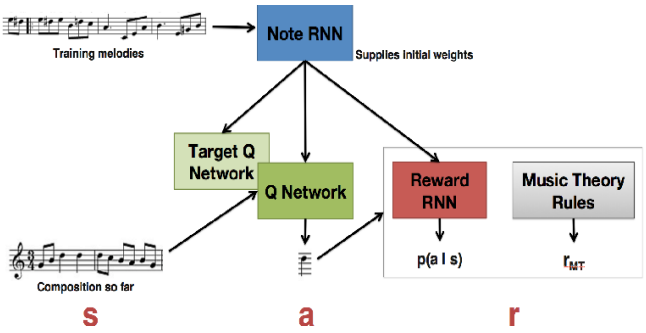
Hadjeres, G., Pachet, F., & Nielsen, F. (2017, July). Deepbach: a steerable model for bach chorales generation. In International Conference on Machine Learning (pp. 1362-1371). PMLR.



Web Paper Code

Fine-Tuning with RL

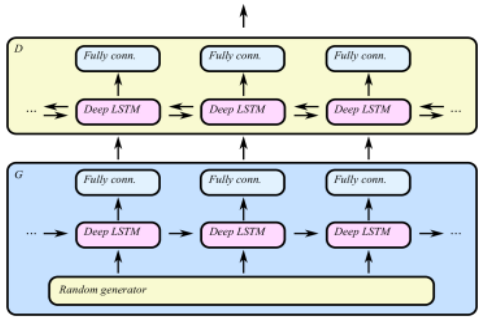
Jaques, N., Gu, S., Turner, R. E., & Eck, D. (2016). Generating music by fine-tuning recurrent neural networks with reinforcement learning.



Paper

C-RNN-GAN

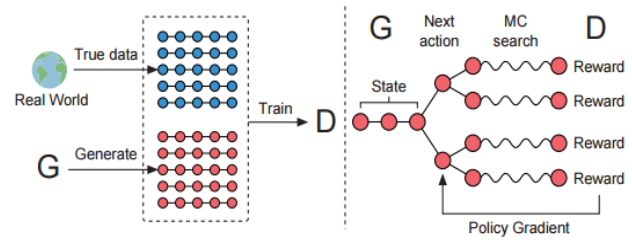
Mogren, O. (2016). C-RNN-GAN: Continuous recurrent neural networks with adversarial training. arXiv preprint arXiv:1611.09904.



Paper

SeqGAN

Yu, L., Zhang, W., Wang, J., & Yu, Y. (2017, February). Seqgan: Sequence generative adversarial nets with policy gradient. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, No. 1).



Paper

2002

Temporal Structure in Music

Eck, D., & Schmidhuber, J. (2002, September). Finding temporal structure in music: Blues improvisation with LSTM recurrent networks. In Proceedings of the 12th IEEE workshop on neural networks for signal processing (pp. 747-756). IEEE.

Paper

Books and Reviews

Books

- Briot, J. P., Hadjeres, G., & Pachet, F. (2020). Deep learning techniques for music generation (pp. 1-249). Springer.

Reviews

- Ji, S., Luo, J., & Yang, X. (2020). A Comprehensive Survey on Deep Music Generation: Multi-level Representations, Algorithms, Evaluations, and Future Directions. arXiv preprint arXiv:2011.06801. Paper
- Briot, J. P., Hadjeres, G., & Pachet, F. D. (2017). Deep learning techniques for music generation--a survey. arXiv preprint arXiv:1709.01620. Paper

4. Datasets

- The Lakh MIDI Dataset v0.1 Web Tutorial IPython

5. Journals and Conferences

- International Society for Music Information Retrieval (ISMIR) Web
- IEEE Signal Processing (ICASSP) Web
- ELSEVIER Signal Processing Journal Web
- Association for the Advancement of Artificial Intelligence (AAAI) Web
- Journal of Artificial Intelligence Research (JAIR) Web
- International Joint Conferences on Artificial Intelligence (IJCAI) Web
- International Conference on Learning Representations (ICLR) Web
- IET Signal Processing Journal Web
- Journal of New Music Research (JNMR) Web
- Audio Engineering Society - Conference on Semantic Audio (AES) Web
- International Conference on Digital Audio Effects (DAFx) Web

6. Authors

- David Cope Web
- Colin Raffel Web
- Jesse Engel Web
- Douglas Eck Web
- François Pachet Web

7. Research Groups and Labs

- Audiolabs Erlangen Web
- Music Informatics Group Web
- Music and Artificial Intelligence Lab Web

8. Apps for Music Generation with AI

- AIVA (paid) Web
- Amper Music (paid) Web
- Ecrett Music (paid) Web
- Humtap (free, iOS) Web
- Amadeus Code (free/paid, iOS) Web
- Computoser (free) Web
- Brain.fm (paid) Web

9. Other Resources

- Bustena (web in spanish to learn harmony theory) Web

Author

Carlos Hernández-Oliván: carloshero@unizar.es