Deep learning models for music generation

1 Description

One of the interdisciplinary applications of deep neural networks has been in the field of Art. Adversarial Networks [1], Variational Autoencoders [2] and other generative DNNs have been successfully applied to create artificial images that resemble real images [3, 4, 5, 6, 7] and also for generating music according to different styles [8, 9, 10, 11, 12].

2 Objectives

The goal of this project is to investigate the use of deep generator models (GANs, VAEs, or any other combination of neural models) for music analysis (either music generation, music style recognition or or music style transfer. The student will be free to use the music data representation, the datasets and the models.

The student should: 1) Clearly state the problem addressed. 2) Select an appropriate DNN architecture. 2) Select convenient music datasets. 3) Implement the model learning procedure. 4) Evaluating the quality of generated music may be a difficult question with a strong subjective component. Therefore, for the evaluation the student could select a number of exemplars output of the model, and explain the criteria considered to evaluate the results (e.g., asking humans to judge the realism or quality of the generated music).

As in other projects, a report should describe the characteristics of the design, implementation, and results. A Jupyter notebook should include calls to the implemented function that illustrate the way it works.

3 Suggestions

- Read the relevant bibliography about music generation [8, 9, 10, 11, 12], including potential available datasets.
- Implementations can use any Python library that implements DNNs.

References

[1] 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.

[2] Diederik P Kingma and Max Welling. Auto-encoding variational Bayes. *CoRR*, abs/1312.6114, 2013.

[3] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and Jürgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, page 1237. Barcelona, Spain, 2011.

[4] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature

embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 675–678. ACM, 2014.

[5] Yoon Kim. Convolutional neural networks for sentence classification. *CoRR*, abs/1408.5882, 2014.

[6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[7] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[8] Gino Brunner, Yuyi Wang, Roger Wattenhofer, and Sumu Zhao. Symbolic music genre transfer with CycleGAN. *CoRR*, abs/1809.07575, 2018.

[9] Sander Dieleman, Aäron van den Oord, and Karen Simonyan. The challenge of realistic music generation: modelling raw audio at scale. *CoRR*, abs/1806.10474, 2018.

[10] Andrew Pfalz. *Generating Audio Using Recurrent Neural Networks*. PhD thesis, Louisiana State University and Agricultural and Mechanical College, 2018.

[11] Gaëtan Hadjeres, François Pachet, and Frank Nielsen. Deepbach: a steerable model for Bach chorales generation. *CoRR*, abs/1612.01010, 2016.

[12] Alexey Tikhonov and Ivan P Yamshchikov. Music generation with variational recurrent autoencoder supported by history. *CoRR*, abs/1705.05458, 2017.