

Applications of deep learning: audio (music and other sounds)

This includes problems such as: sound classification, music source separation, chord classification, audio generation, timbre transfer, music transcription.

6a. GANSynth

Citation: Jesse Engel, Kumar Krishna Agrawal, Shuo Chen, Ishaan Gulrajani, Chris Donahue, Adam Roberts. GANSynth: Adversarial Neural Audio Synthesis. In Proceedings of the 2019 International Conference on Learning Representations (ICLR '19). [[PDF](#)] [[Notebook](#) (official)] [[Examples](#)] [[Github](#)] [[Blog post](#)]

6b. Generating music with singing - Jukebox

Citation: Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, Ilya Sutskever. Jukebox: A Generative Model for Music. arXiv:2005.00341 (2020). [[PDF](#)] [[Github](#) (official)] [[Blog post](#)] [[Examples](#)] [[Notebook](#) (switch to the smaller model to avoid running out of memory on Colab)]

6c. Differentiable Digital Signal Processing + ML

Citation: Jesse Engel, Lamtharn (Hanoi) Hantrakul, Chenjie Gu, Adam Roberts. DDSP: Differentiable Digital Signal Processing. In Proceedings of the 2020 International Conference on Learning Representations (ICLR '20). [[Open review](#)] [[PDF](#)] [[Github](#)] [[Blog post](#)] [[Examples](#)] [[Notebook: timbre transfer](#)] [[Tutorials](#)]

6d. Generating Drum Performances

Citation: Jon Gillick, Adam Roberts, Jesse Engel, Douglas Eck, David Bamman. Jon Gillick, Adam Roberts, Jesse Engel, Douglas Eck, David Bamman. Learning to Groove with Inverse Sequence Transformations. In Proceedings of the 2019 International Conference on Machine Learning (ICML '19). [[PDF](#)] [[Blog post](#)] [[Notebook](#)] [[Examples](#)] [[Github](#)]

6e. Continuing a note sequence, with long-term structure

Citation: Elliot Waite. Generating Long-Term Structure in Songs and Stories (2016). [[Blog post](#)] [[Github](#) (includes training code)] [[Notebook](#): shows how to use pre-trained Melody-RNN. You can modify it to use Lookback and/or Attention]

6f. Music style transfer

Citation: Noam Mor, Lior Wolf, Adam Polyak, Yaniv Taigman. A Universal Music Translation Network. [[PDF](#)] [[Github](#) (official) (note: this is optimized for a specific GPU, so you will have to use the Python-based generation with pre-trained models on Colab)]

Citation: Gino Brunner, Yuyi Wang, Roger Wattenhofer and Sumu Zhao. Symbolic Music Genre Transfer with CycleGAN. In Proceedings of the 30th International Conference on Tools with Artificial Intelligence (ICTAI '18). [[PDF](#)] [[Github v1](#), [v2](#)]

6g. Music source separation

Citation: Stefan Uhlich, Marcello Porcu, Franck Giron, Michael Enenkl, Thomas Kemp, Naoya Takahashi, and Yuki Mitsufuji. Improving music source separation based on deep neural networks through data augmentation and network blending. In

Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '17) [[Paper](#) (use NYU Library account)] [Open unmix: [Github](#), [notebook](#) for pre-trained models, [notebook](#) for training, [tutorials](#)] [Spleeter: [Github](#) for TF, with [notebook](#), pre-trained models] [[Blog post](#)]

6h. Onset and frames: piano transcription

Citation: Curtis Hawthorne, Erich Elsen, Jialin Song, Adam Roberts, Ian Simon, Colin Raffel, Jesse Engel, Sageev Oore, Douglas Eck. Onsets and Frames: Dual-Objective Piano Transcription. In Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR '18), 2018. [[PDF](#)] [[Examples](#)] [[Blog post](#)] [[Github](#) (official)] [[Notebook](#) (official)] [[Notebook](#) in PyTorch]