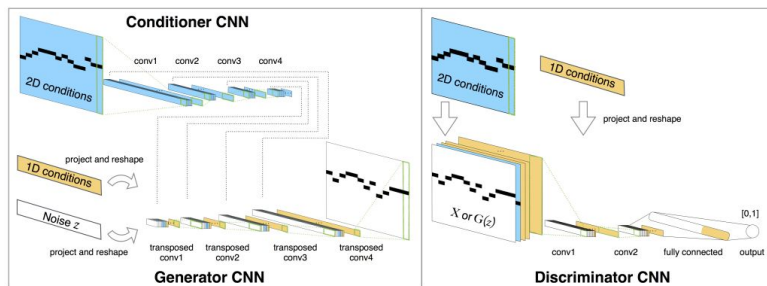


Jukebox: A Generative Model for Music

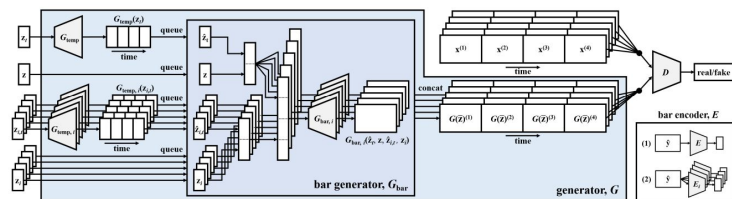
Спирин Иван. Рецензия

Модели для генерации MIDI



MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation (2017)

[\[https://arxiv.org/abs/1703.10847\]](https://arxiv.org/abs/1703.10847)



MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment (2017)

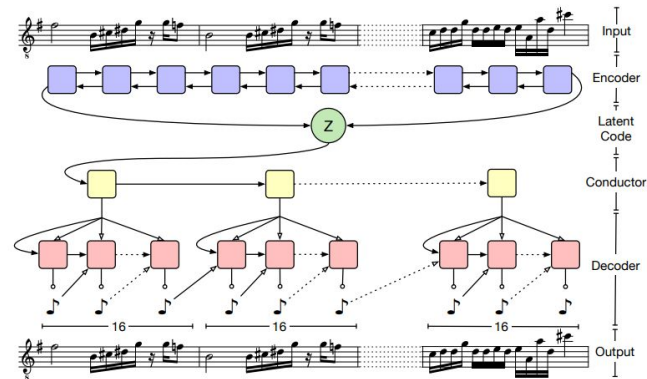
[\[https://arxiv.org/abs/1709.06298\]](https://arxiv.org/abs/1709.06298)

Модели для генерации MIDI



MuseNet : Music Generation using Abstractive and Generative Methods (2019)

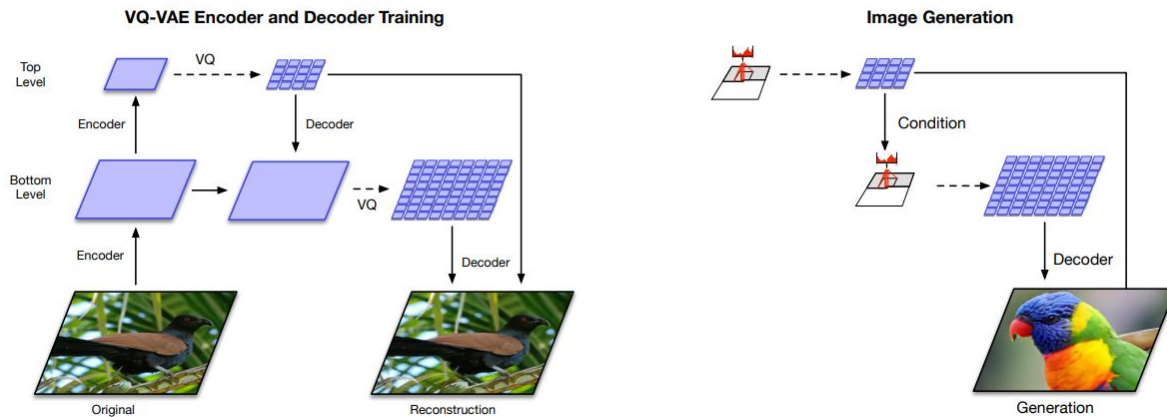
[<https://openai.com/research/musenet>]



A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music (2019)

[<https://arxiv.org/abs/1803.05428>]

VQ-VAE

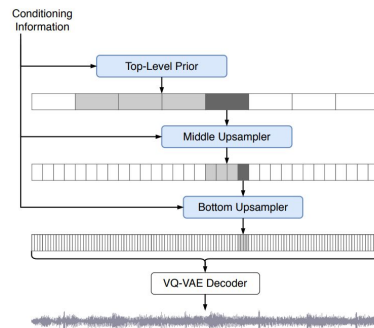
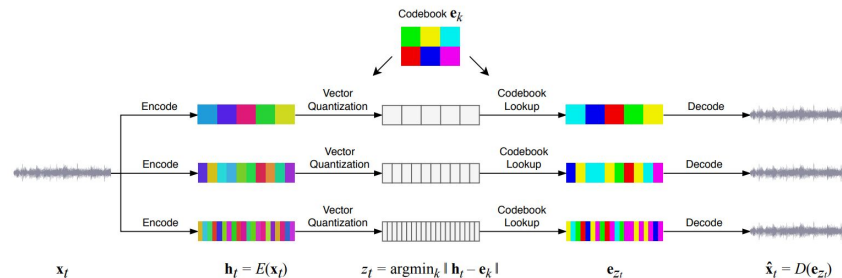


Generating Diverse High-Fidelity Images with VQ-VAE-2 (2019)

[<https://arxiv.org/abs/1906.00446>]

Jukebox

- Hierarchical VQ-VAE
- Autoregressive sparse transformer for compressing
- WaveNet for upsampling
- Artist, genre, lyrics conditioning
- Capable of generating pieces that are multiple minutes long



[<https://arxiv.org/abs/2005.00341>]

Авторы



Prafulla Dhariwal

Researcher,
OpenAI



Alec Radford

Researcher,
OpenAI

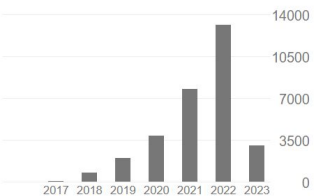


Jong Wook Kim

Researcher,
OpenAI

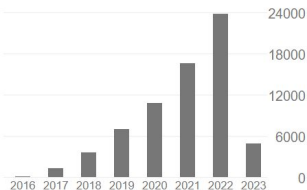
Процитировано

	Все	Начиная с 2018 г.
Статистика цитирования	31200	30929
h-индекс	15	15
i10-индекс	15	15



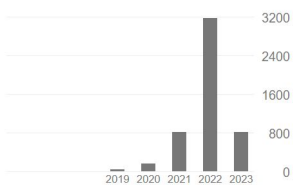
Процитировано

	Все	Начиная с 2018 г.
Статистика цитирования	69535	67568
h-индекс	31	31
i10-индекс	32	32



Процитировано

	Все	Начиная с 2018 г.
Статистика цитирования	5126	5115
h-индекс	11	11
i10-индекс	11	11



Плюсы

- Лучшее качество генерации музыки на момент выхода статьи
- Генерирует вокал поверх мелодии
- Возможность продолжить любую композицию
- Стабильная генерация в течении нескольких минут
- Все в открытом доступе

Минусы

- Тяжелая модель (5B параметров) и долгая генерация (8 часов на 1 минуту аудио)
- У музыки на выходе нет композиции
- Много шумов на выходе
- Вокал только на английском
- Нет метрик и сравнения с другими моделями

Генерация музыки сегодня

- Коммерческие сервисы: Mubert, Soundful, Soundraw, Boomy
- Riffusion (2022) - Stable Diffusion + riff
- MusicLM: Generating Music From Text (2023)
- Noise2Music: Text-conditioned Music Generation with Diffusion Models (2023)
- ArchiSound: Audio Generation with Diffusion (2023)

