

ГАНы

Пак Ди Ун, ПМИ-182

29 марта 2022

Отчет

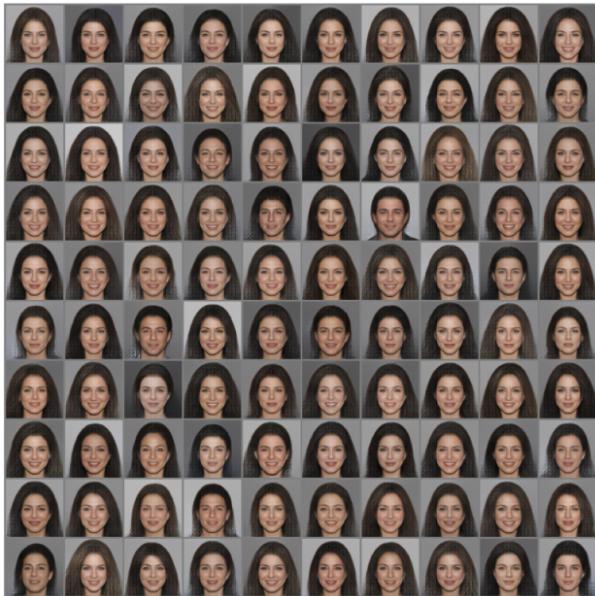
Задача 1. *GLOW*

Модель

Написал свою модель чисто по статье за пару часов, отладил и поставил на обучение. Оно не училось, пару дней из-за этого дебагал (начал писать, как появилась тетрадка). Нашел рабочую реализацию и начал использовать в качестве дебаг помощника. Там в issues'ах было несколько обсуждено несколько моментов из статьи, в которых реализация отличалась для увеличения скорости сходимости, эти хаки я тоже взял с собой. Также один из своих классов в итоге не заработал, поэтому частично его забрал (reverse класса Block). Также для дебага однокурсники посоветовали использовать не просто nll, а его попиксельную нормализацию, константы тоже забрал из репа. Дальше завел логгирование в wandb. На всю модель ушло примерно 4 дня и 40 гпу часов (личный комп работал ночью). В основном эксперименты были с размером сетки (8 на 4 в итоге), квантизацией и размером батча. Итоговая модель училась 12 часов. FID: 3.7805.

Результаты:

Генерация без черрипикинга. Температура 0.2. Однаковые красивые лица. Война клонов:



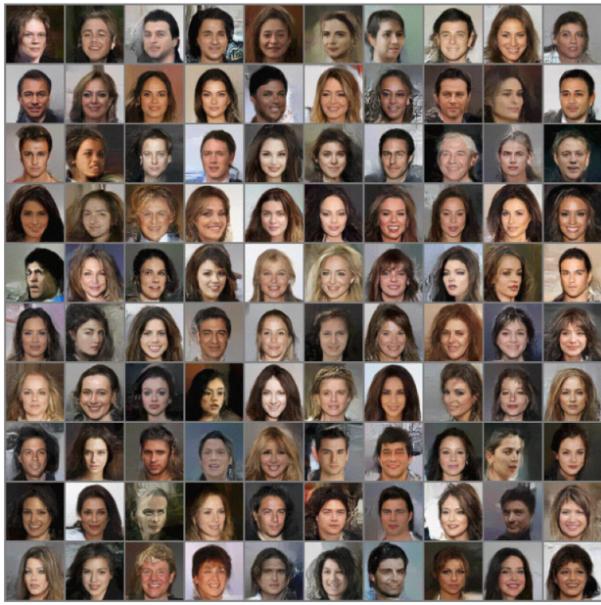
Генерация без черрипикинга. Температура 0.5. Все еще красивые и похожие, но есть солидное разнообразие:



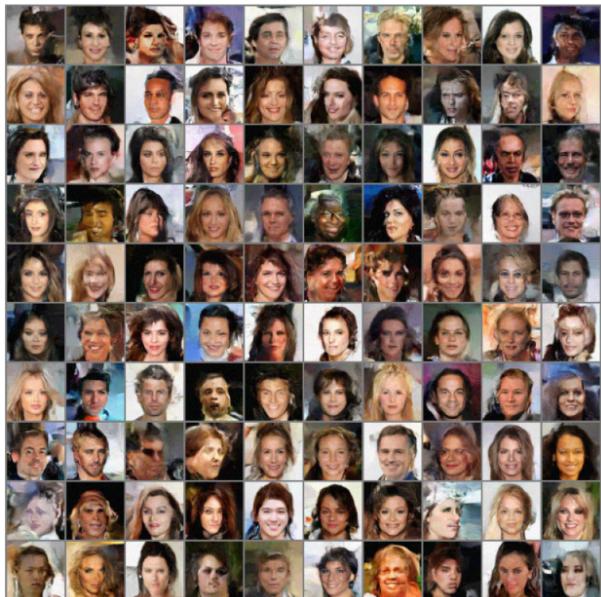
Генерация без черрипикинга. Температура 0.6. Красота:



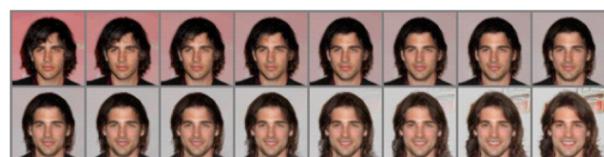
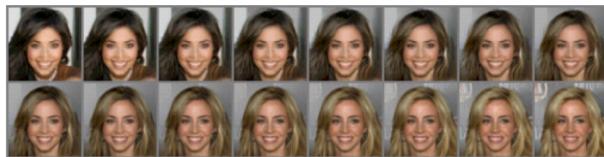
Генерация без черрипикинга. Температура 0.7. Самый кайф, хорошее разнообразие, но появляются артефакты:



Генерация без черрипикинга. Температура 1. Начинаются стремные штуки из эпохи возрождения:



Латентная интерполяция. Одежда и разные особенности переключаются хорошо и резко. Все промежуточные лица выглядят как реальные:



Ближайшие в латенте картинки. На реальной фотке не сработало, но фотка Тома и первая найденная +- похожи:



Задача 2. VAE

Чисто по статье накидал код модели, прописал нужные функции за час, наладил пайплайн обучения. Потратил пару часов на подбор веса перед KL, так как KL сильно перевешивал реконструкцию, приводя к однообразным лицам. Когда вес был слишком маленьким, семплирование отпадало, так как распределение латента было слишком плохим. Также изображения мыльноватые получились, поиграл с размерностью латента в адекватных границах (не больше 128, та как картинки 64x64), адекватное качество получил при 64 (все равно мыльно и не очень, но приемлемо). Лица получились чуть более размытыми, чем у реализации из нашего репа, но незначительно. FID: 2.7141. Результаты работы:

Генерация без черрипикинга. Температура 0.2. Однократные стереотипные лица:



Генерация без черрипикинга. Температура 0.5. Лица одинаковые, но меняется пол и цвет волос:



Генерация без черрипикинга. Температура 0.6. Меняются формы лица:



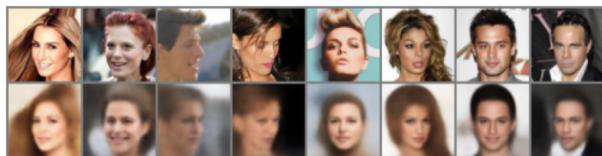
Генерация без черрипикинга. Температура 0.7. Тут уже все меняется, все красиво:



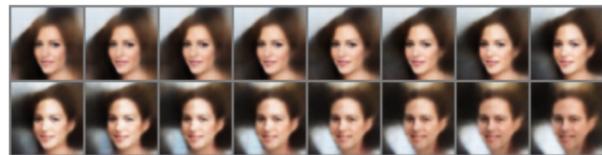
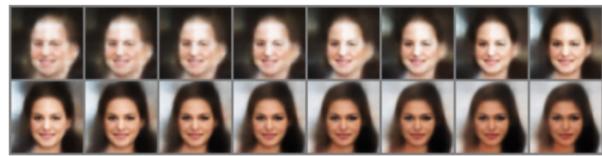
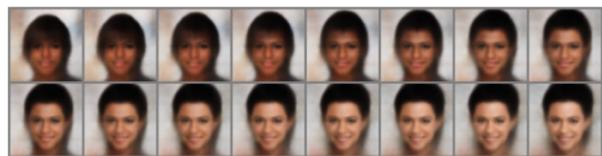
Генерация без черрипикинга. Температура 1. Даже практически нет артифактов, все по кайфу:



Реконструкция. Красиво и правдоподобно, но все мылится:

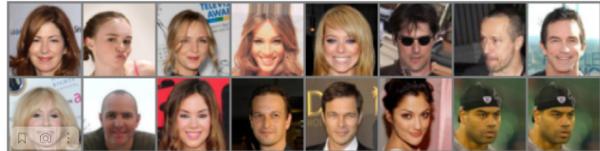


Латентная интерполяция. Все чуток смазано, поэтому интерполяция тоже стремяная. На промежуточных итерациях лица еще хуже:





Ближайшие в латенте картинки (также не особо скоррелированный результат получился):



Задача 3. Сравнение

В латенте лучше из двух моделей справляется Glow, каждая точка выглядит как реальное лицо, медленно меняющееся. VAE же в этом случае просто морфит изображения как в фотошопе.

Поиск ближайших картинок в латенте не очень хорошо отработал у обеих моделей. Подозреваю, это из-за того, что датасет Celeba сильно отличается от

обычных фотографий (также было обнаружено в предыдущей домашке).
Лучшие мемы:

Не мем, но искусство:

