

ГАНы

Пак Ди Ун, ПМИ-182

15 марта 2022

Отчет

Задача 1. *StarGan V1*

Результаты работы

Black hair. Работает с кайфом. Примеры:

Target domain: Black_Hair



Target domain: Black_Hair



Target domain: Black_Hair



Target domain: Black_Hair



Brown hair. Тоже, вроде, с кайфом:

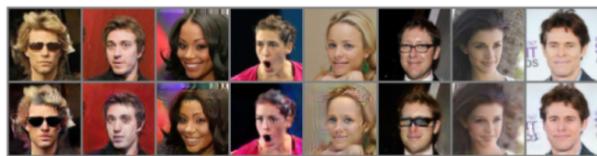
Target domain: Brown_Hair



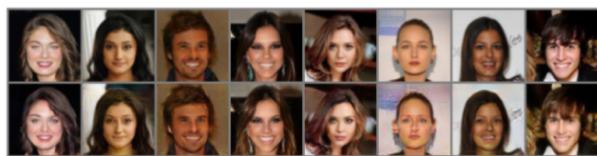
Target domain: Brown_Hair



Target domain: Brown_Hair



Target domain: Brown_Hair



Blond hair. Работает с переменным успехом. Больше похоже на red hair имхо:

Target domain: Blond_Hair



Target domain: Blond_Hair



Target domain: Blond_Hair

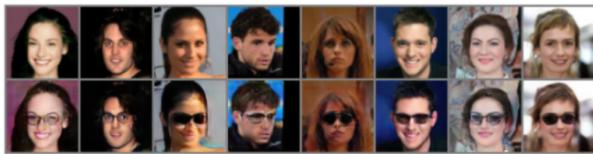


Target domain: Blond_Hair



Eyeeglasses. Ну что за POG:

Target domain: Eyeglasses



Target domain: Eyeglasses



Target domain: Eyeglasses



Target domain: Eyeglasses



Goatee. Работает очень хорошо на удивление. Даже на девушках смотрится:

Target domain: Goatee



Target domain: Goatee



Target domain: Goatee



Target domain: Goatee



Male. Это гигапог, нечего добавить:

Target domain: Male



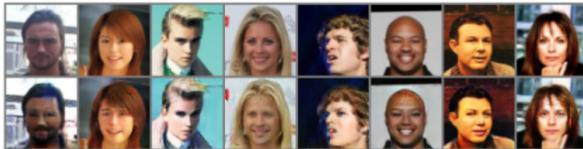
Target domain: Male



Target domain: Male



Target domain: Male



Mustache. Работает хорошо, но усыки у всех одинаковые, что может потом выстрелить в ногу:

Target domain: Mustache



Target domain: Mustache



Target domain: Mustache



Target domain: Mustache



Receding hairline. Супер переменный успех:

Target domain: Receding_Hairline



Target domain: Receding_Hairline



Target domain: Receding_Hairline



Target domain: Receding_Hairline



Wearing hat. Попытки добавления бандан и панам видны, но не слишком успешны:

Target domain: Wearing_Hat



Target domain: Wearing_Hat



Target domain: Wearing_Hat



Target domain: Wearing_Hat



Young. В большинстве случаев ничего не делает, но иногда из молодых хоропо рисует старых (пример (0, 0) и (0, 3)):

Target domain: Young



Target domain: Young



Target domain: Young



Target domain: Young



КРОР. Шо тут сказать, волосы красит хорошо; очки, бородка и усыки прошли, остальное реально не очень:



КРОР. Аналогично, красит волосы норм; очки, бородка и усыки различимы; повязка на голове кайнда понятна, остальное фигня:



Rick. Девушка Рик и очкастый Рик горячи. Остальных чуточку пошакалило, но усы+бородка работают. Волосы красятся плохо:



Shrek. Shrek Is Love, Shrek Is Life:



Trump. Забавно, вроде, все фотки превратили Трампа в Джастина Бибера with extra steps:



Наблюдения

В целом, модель работает, причем на реальных данных тоже что-то выдает. Качество генерации немного проседает, это видно, но объясняется тем, что я

повысил коэф. у функции потерь, отвечающей за классификацию домена. Это сделано было, потому что первые 4 модели очень хорошо восстанавливали картинку, но не меняли домен, что было грустно с точки зрения предназначения ГАНа)

Использовал идею PatchGAN'a (также пытался без нее, но не взлетело), градиентный штраф (без него тоже не заработало), а также пробовал крутить веса лоссов (получил финальную хорошую модель).

По коду, было не так сложно запрограммировать, сколько дождаться обучения. Было несколько багов на пару дней, но ничего суперсерьезного, все заработало в итоге. 0.07099916191110688 LPIPS на честном флипе.

Задача 2. StarGan V2

Результаты работы

Из одного домена в произвольные. Без черрипикинга)

Black hair. Эти примеры хороши даже без черрипикинга. Примеры различные, по стилю очень много где совпадали, как и +- по референсу (поза, контент).



Blond hair. Тут тоже все хорошо, только иногда части изображений плывут.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Brown hair. Иногда непонятно, что взято от референса, но мы и не хардкодили это, то есть контент выбирался по предположениям генератора о контенте...



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Еyeglasses. С очками результаты как-то пополыли сильно, не справляется моделька.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Goatee. С бородкой боремся с переменным успехом, иногда повторяя стиль, что интересно.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Male. Мужички справились с работой, но это просто рандом, так как для разных полов у нас по факту разные ветки работают.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Mustache. Для стандартных фоток работает, но иногда сильно страдает.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Receding Hairline. Броде, как везде. Есть стремные варианты, есть прямо по кайфу.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Wearing Hat. Много изображений без явной корреляции.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Young. Вообще подозрительный признак, тут все как везде.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Из произвольного домена в данный. Отсмотрев все эти картинки, нельзя прийти к какому-то заключению, что модель считает стилем/контентом. Иногда видно корреляция с оригиналом/референсом, но тяжело сказать, насколько она интерпретируется в целом.

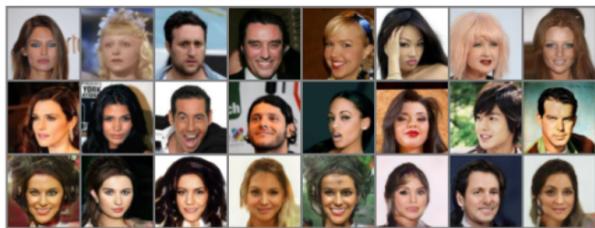
Black hair.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



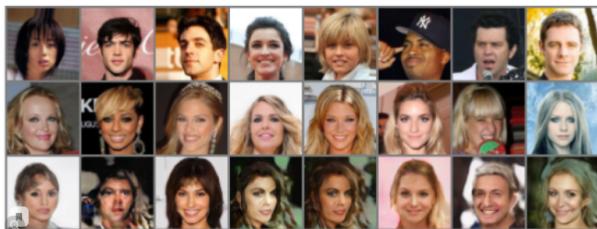
Blond hair.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Brown hair.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



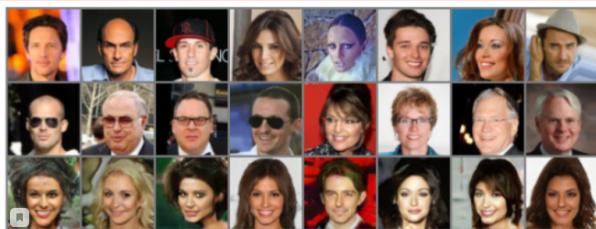
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Eyeglasses.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Goatee.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



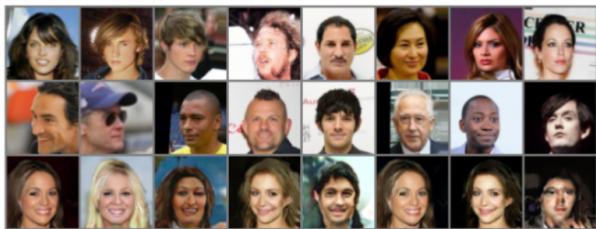
Male.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Mustache.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



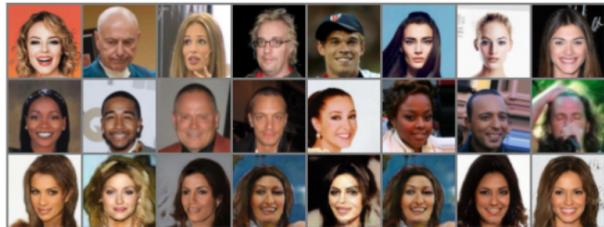
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Receding Hairline.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Wearing Hat.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Young.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Интерполяция стиля. Взял 2 изображения, получил стиль и нагенерил промежуточные варианты с смешанным стилем. Получилось очень даже сочно и интерпретируемо.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



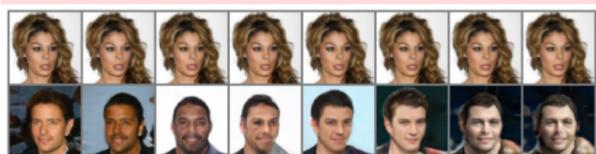
Интерполяция z . Взял 2 z , интерполировал и создал стиль. Также видно, что оно внутри гладкое.



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).



Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

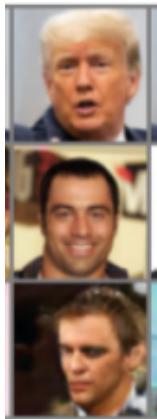


Свои картинки. Оно работает!!!:





Любимая картинка. Раскрыт секрет личности Навального:



Наблюдения

Задачу решал тяжело и долго. Сначала накидал архитектуру из статьи с нуля, потом накидал каркас лосса. Первые несколько часов дебагал лосс и подгонял размерности, потом запустил и понял, что ничего не учится. Нашел багос в своем AdaIN, переписал и снова запустил. Утром ничего не получил, решил перечитать статью по AdaIN. Вроде, ничего не упустил, но все равно не работало. В итоге сдался и полез в оф. реп StarGanV2, забрал AdaIN оттуда. Начало как-то работать. Потом сверил лоссы, понял, что штраф градиента тоже неправильно

был написан, поправил. Дальше полтора дня учил и получил неплохую модельку. Также затем прикрутил moving average. Прирост наибольший дал AdaIN во время обучения, проверял, что без ничего не учится, но это ожидалось. Также пытался отрезать ветку с mapping сеткой, там училось, но медленнее. Лучшее качество визуально дал Moving Average. Итоговая модель показывает 0.25 LPIPS, что выглядит не очень впечатляюще, но визуально все очень неплохо. Если честно, не очень шарю за гигаподходы в ГАНах, поэтому могу предложить только каким-то образом попытаться учить не просто нормальное распределение в латенте, а пытаться его на некоторые штуки обусловить. Ну или генерить многомерное для повышенной интерпретируемости. Также можно посмотреть размеры сеток, так как, кажется, для нашего ужатого датасета можно и сетку пожать. Также я бы поигрался с lr, потому что, кажется, сначала учится слишком плохо, а потом слишком быстро (пролетая нужные минимумы, если судить по изображениям). Шедулер – сила!