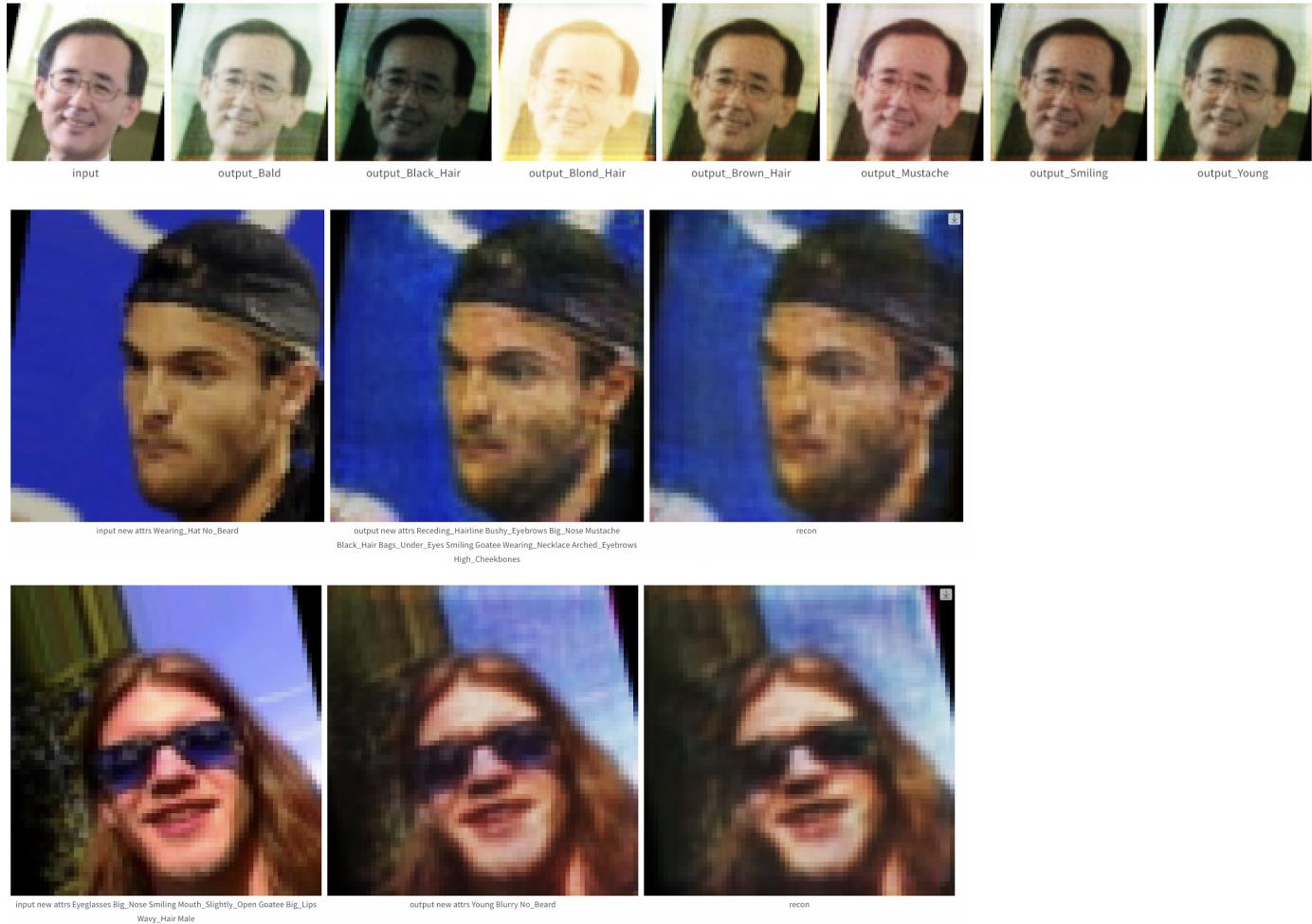
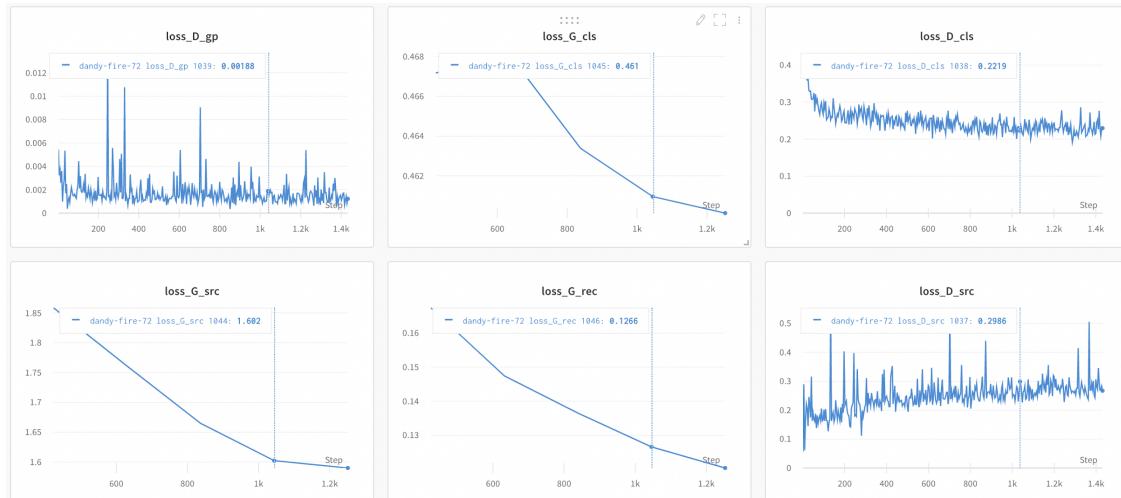


## Задача 1

Для начала была реализована чистая модель по статье StarGAN, с такими же функциями потерь и весами, с размером картинок 90, обучал в течение 12 эпох. Пример работы:



Графики функций потерь.



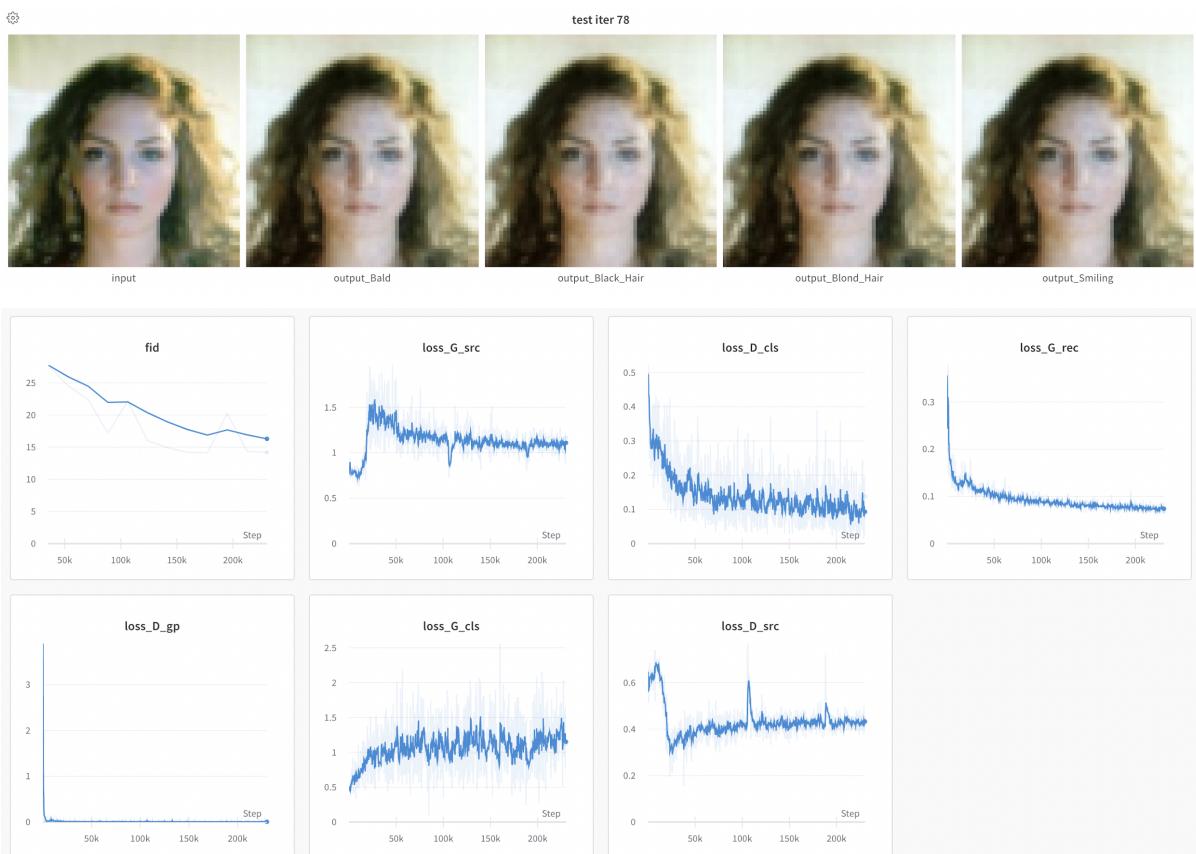
Далее к стандартной архитектуре были применены аугментации (случайные повороты на угол от -15 до 15 и по горизонтали). Визуально результаты на валидационной выборке не сильно изменились, FID метрика была приблизительно равна 40. Далее решил использовать BatchNorm вместо стандартного InstanceNorm (он мешал восстанавливать цвет картинки на всем батче) и заменил ConvTranspose на Upsample+Conv2d, что помогло убрать маленькие квадраты на выходе генератора. Так FID упал до 30, визуально стало чуть лучше, но атрибуты восстанавливались все равно плохо.

## Задача 2

FID метрика была реализована с помощью Inception, как писал выше, на обычной архитектуре с небольшими изменениями удалось достичь FID 30.

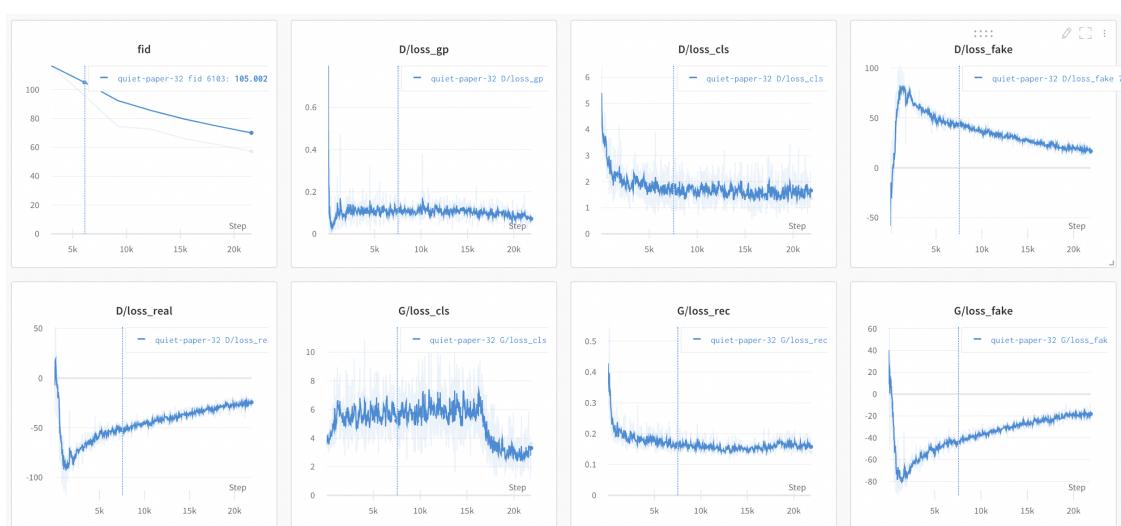
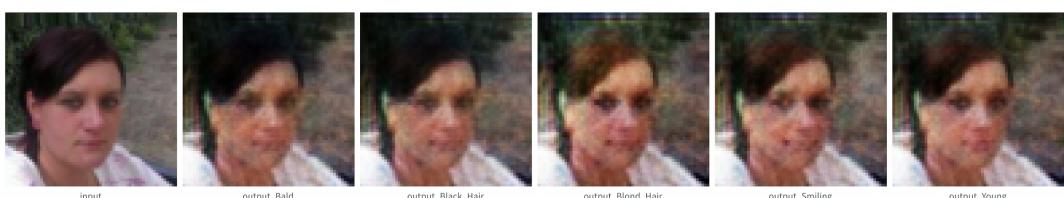
## Задача 3

Для ускорения обучения взял размер картинок 64 и убрал аугментации. В генераторе решил попробовать передавать вектор лейблов не в начале, а перед скрытой частью сети из Residual блоков, т. е. сразу после энкодера. Так за 12 эпох удалось достичь FID 14, при этом генератор научился хорошо восстанавливать изображения, но атрибуты все равно плохо транслировались.



Также провел много экспериментов с незначительными изменениями архитектур генератора и дискриминатора, а также менял количество атрибутов до более простых (например 3 изменения цвета волос + пол + улыбка), надеясь на нормальный перенос атрибутов. FID от этого стал хуже и стал равен ~60 (зато сеть научилась перенесу атрибутов). А также незначительно менял подсчет лоссов. Вот примеры:

8 атрибутов:



40 атрибутов:



Input new attrs Attractive Wearing\_Necklace Young Rosy\_Cheeks Blond\_Hair Heavy\_Makeup No\_Beard Wearing\_Lipstick Oval\_Face Wearing\_Earrings

output new attrs Attractive High\_Cheekbones Smiling Young Mouth\_Slightly\_Open Heavy\_Makeup No\_Beard Wearing\_Lipstick Oval\_Face Wearing\_Earrings

recon



Input new attrs Attractive Bushy\_Eyebrows Big\_Nose Young Goatee Male

output new attrs Black\_Hair Young No\_Beard Pointy\_Nose

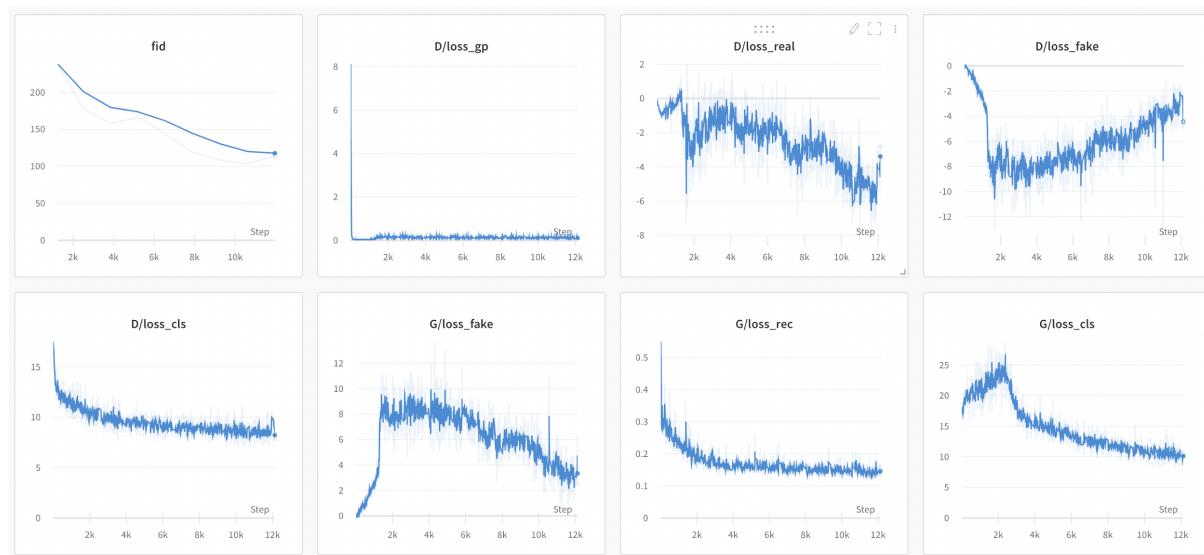
recon



Input new attrs Attractive Bushy\_Eyebrows Black\_Hair High\_Cheekbones Big\_Nose Smiling Young No\_Beard Wearing\_Lipstick Oval\_Face Wearing\_Earrings

output new attrs Attractive High\_Cheekbones Young Pale\_Skin Heavy\_Makeup No\_Beard Wearing\_Lipstick Oval\_Face Wavy\_Hair

recon



По всем графикам видно, что сеть обучается и довольно таки медленно сходится к 0, что особо не наблюдалось в случае с ванильным StarGAN'ом.

Видно что с 40 классами сети сложнее менять их, и скорее всего на обучение нужно больше времени, это видно о по метрике FID, в этом случае она приблизительно равна 100.