



Генерация изображений лиц людей с заданными свойствами

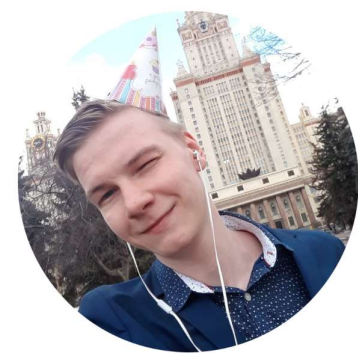
Федотова Алёна



Господинов Георгий



Храмов Алексей



Москва
2018

Постановка задачи

Блондинка, молодая и симпатичная

Постановка задачи

Блондинка, молодая и симпатичная



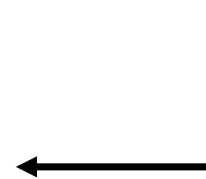
```
young:  1.0  
blond:  1.0  
male:   0.0  
attractive: 1.0
```

Постановка задачи

Блондинка, молодая и симпатичная



```
young:  1.0  
blond:  1.0  
male:   0.0  
attractive: 1.0
```

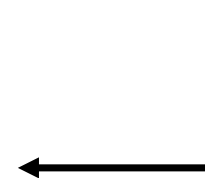


Постановка задачи

Блондинка, молодая и симпатичная



```
young: 1.0  
blond: 1.0  
male: 0.0  
attractive: 1.0
```



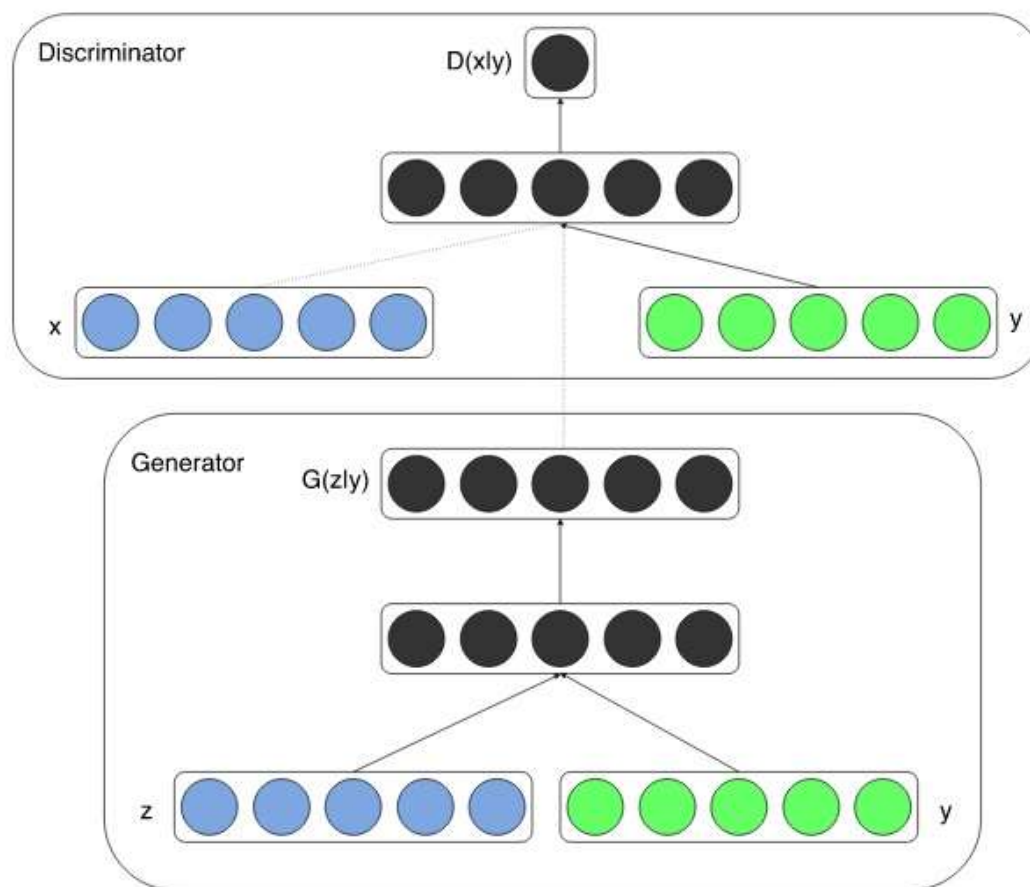
- Данные: CelebA(~200k фото, 40 признаков)
- Код проекта: <https://github.com/ga-gospodinov/FaceGeneration>

Использованные подходы

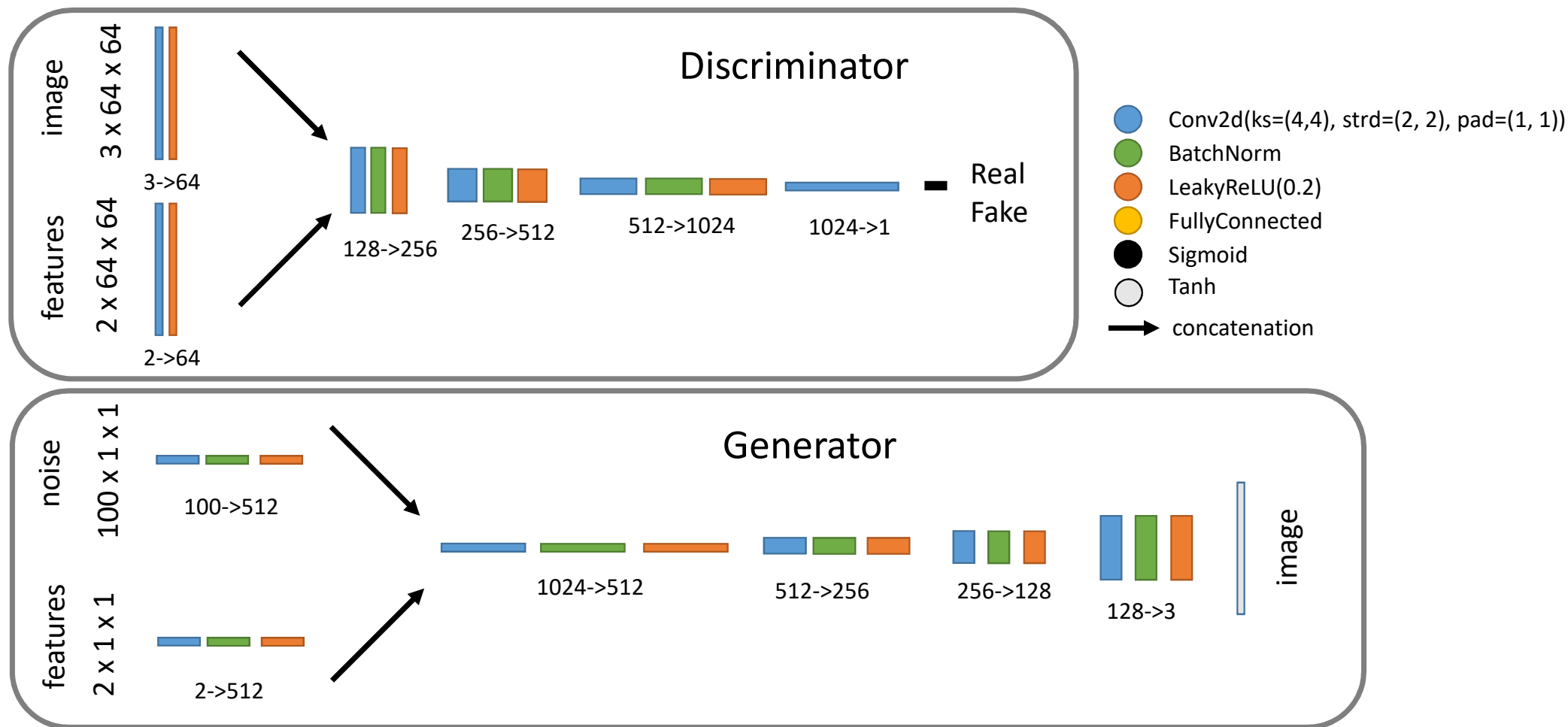
- Conditional GAN
- Fader Network (convolutional adversarial AE)

Conditional GAN

Идея



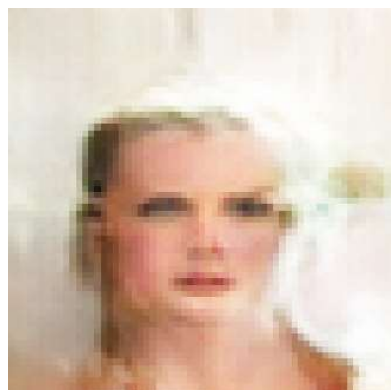
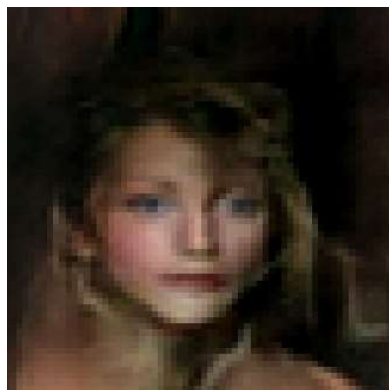
Архитектура



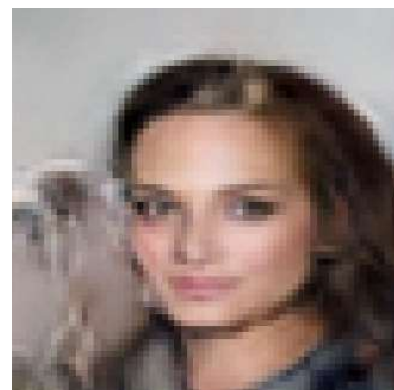
Использованные данные

- Blond_hair, male
- Вся CelebA: > 200k картинок

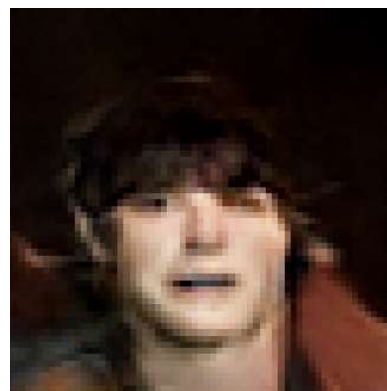
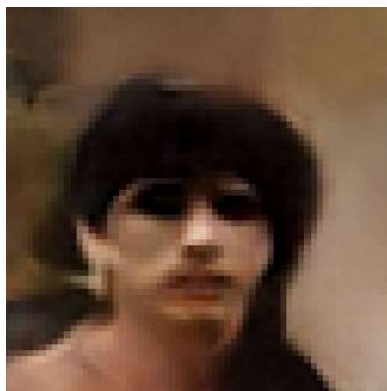
Примеры генерации: блондинки



Примеры генерации: брюнетки



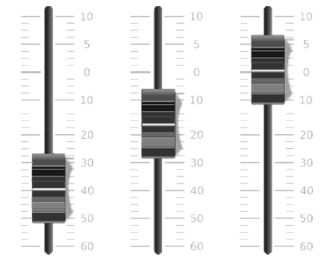
Примеры генерации: брюнеты



Проблемы



Fader Network



Young → Old



Old → Young



Male → Female

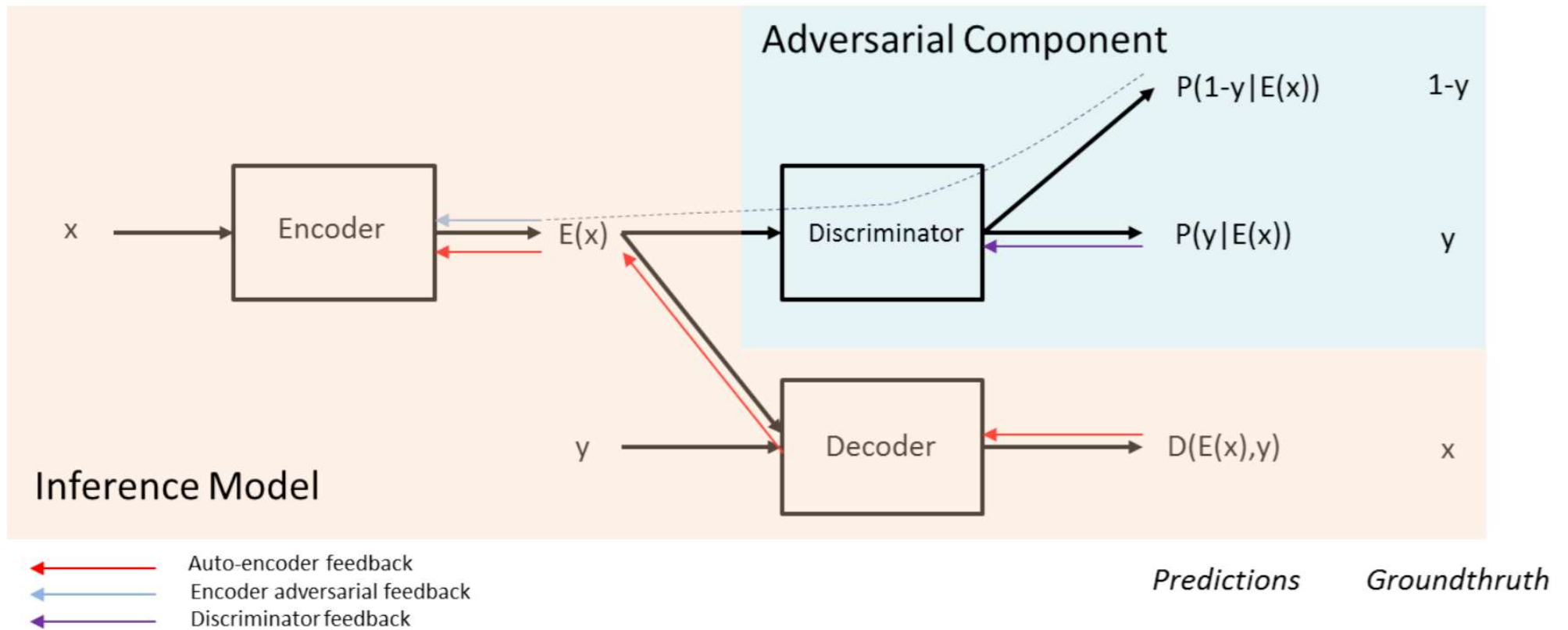


Female → Male



<https://arxiv.org/pdf/1706.00409.pdf>

Идея



Функции потерь

$$\mathcal{L}_{\text{dis}}(\theta_{\text{dis}}|\theta_{\text{enc}}) = -\frac{1}{m} \sum_{(x,y) \in \mathcal{D}} \log P_{\theta_{\text{dis}}}(y|E_{\theta_{\text{enc}}}(x))$$

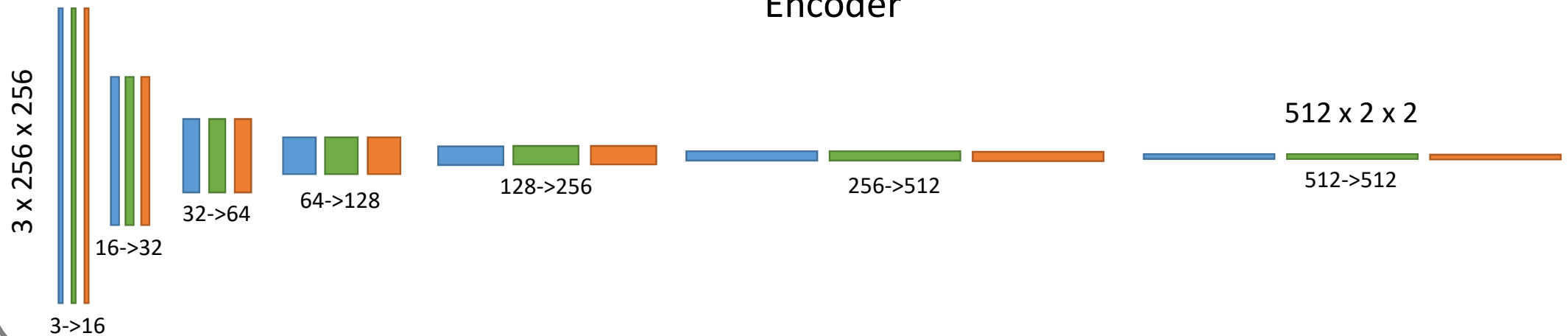
Дискриминатор максимизирует вероятность определить атрибуты изображения по латентному представлению

$$\mathcal{L}(\theta_{\text{enc}}, \theta_{\text{dec}}|\theta_{\text{dis}}) = \frac{1}{m} \sum_{(x,y) \in \mathcal{D}} \|D_{\theta_{\text{dec}}}(E_{\theta_{\text{enc}}}(x), y) - x\|_2^2 - \lambda_E \log P_{\theta_{\text{dis}}}(1 - y|E_{\theta_{\text{enc}}}(x))$$

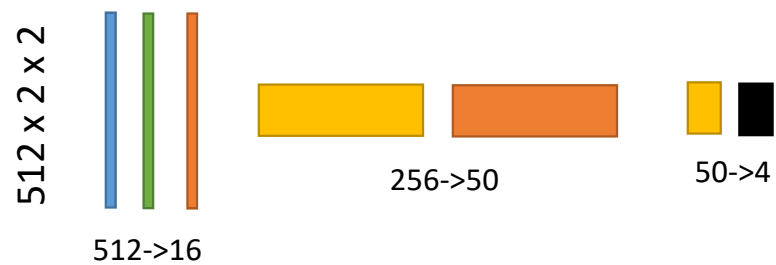
Автоэнкодер минимизирует ошибку восстановления изображения по латентному представлению и атрибутам и максимизирует вероятность ошибки дискриминатора

Архитектура

Encoder



Discriminator



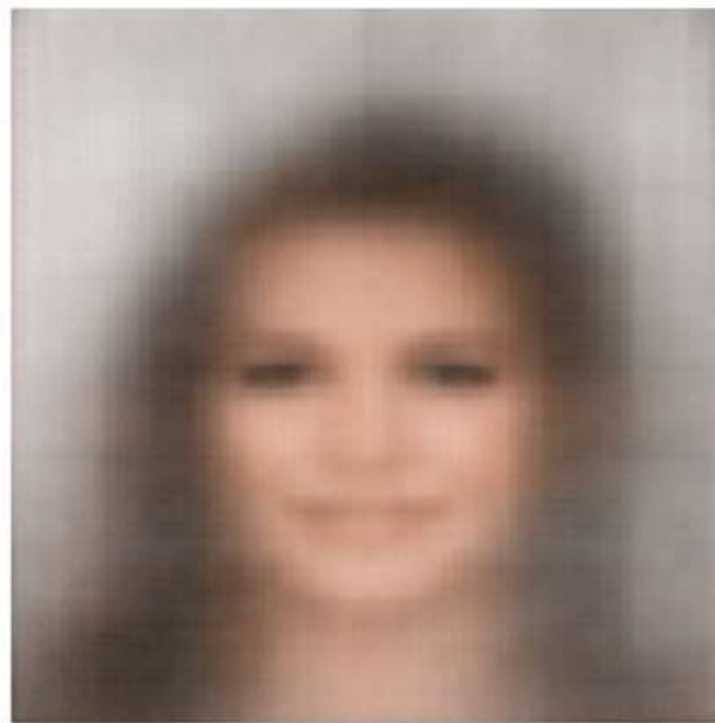
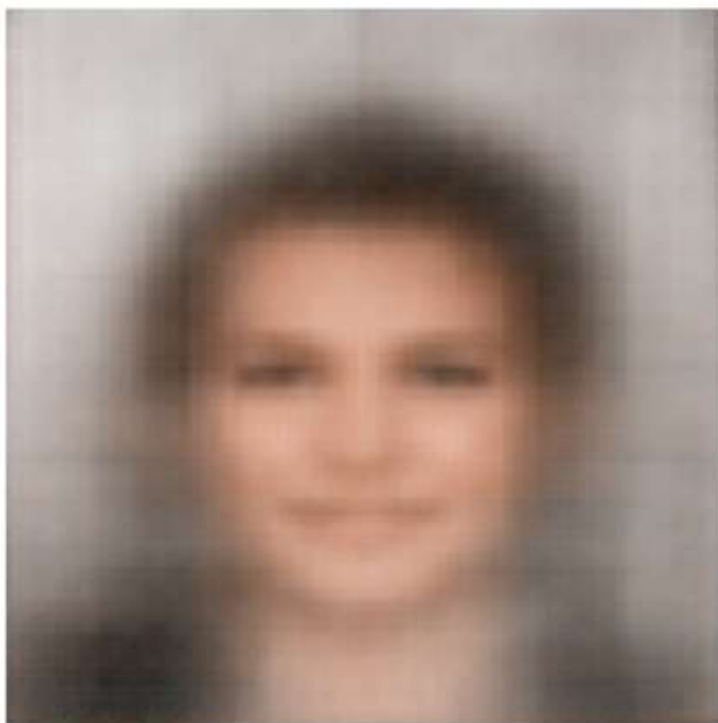
- Conv2d(kernel=(4,4), stride=(2, 2), padding=(1, 1))
- BatchNorm
- LeakyReLU(0.2)
- FullyConnected
- Sigmoid

Использованные данные

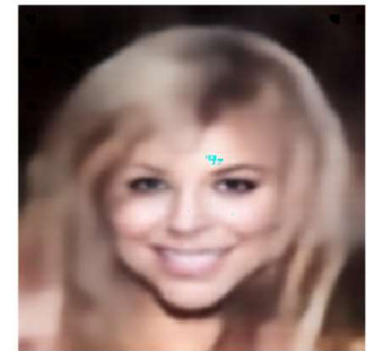
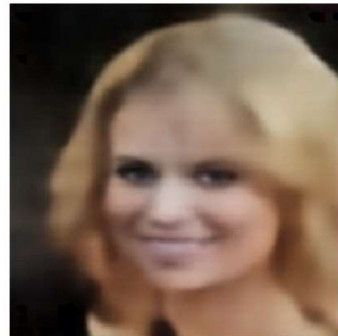
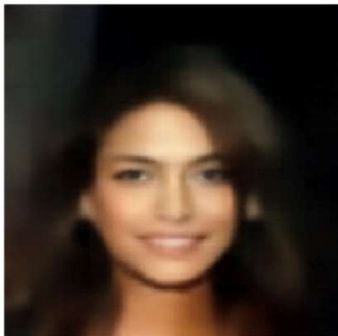
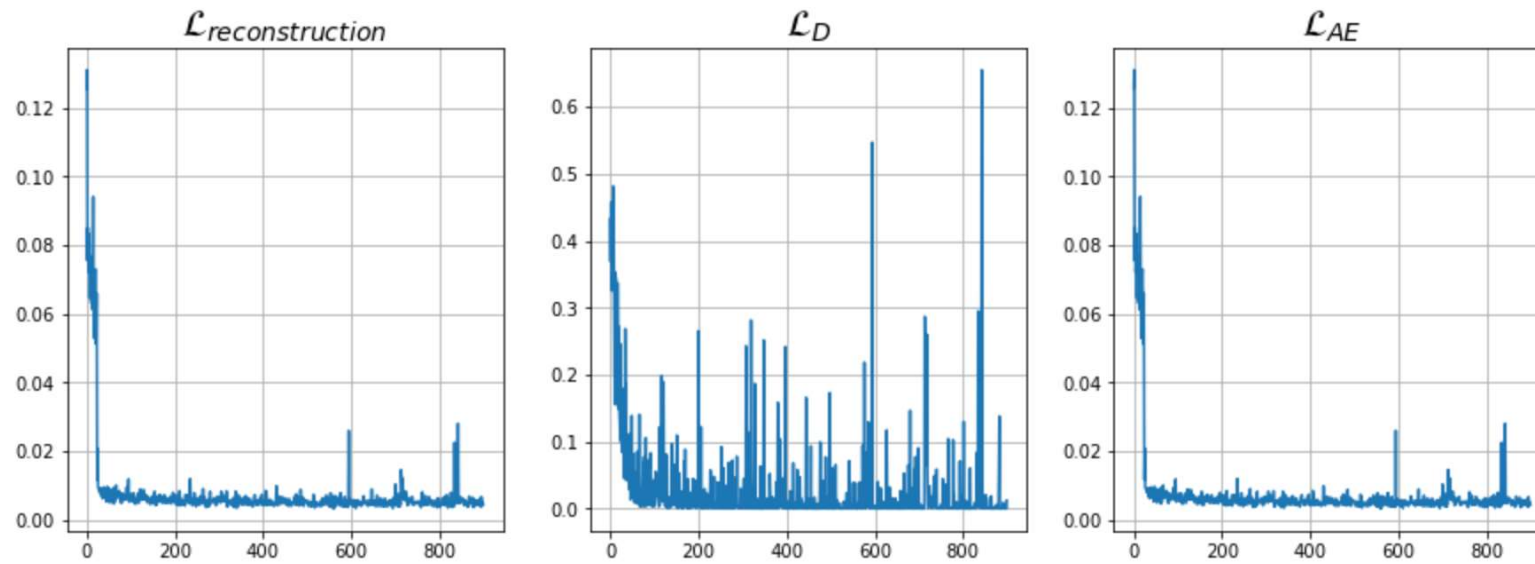
- Blond_hair, attractive, male, young



#1 Обучение одновременно АЕ и дискриминатора



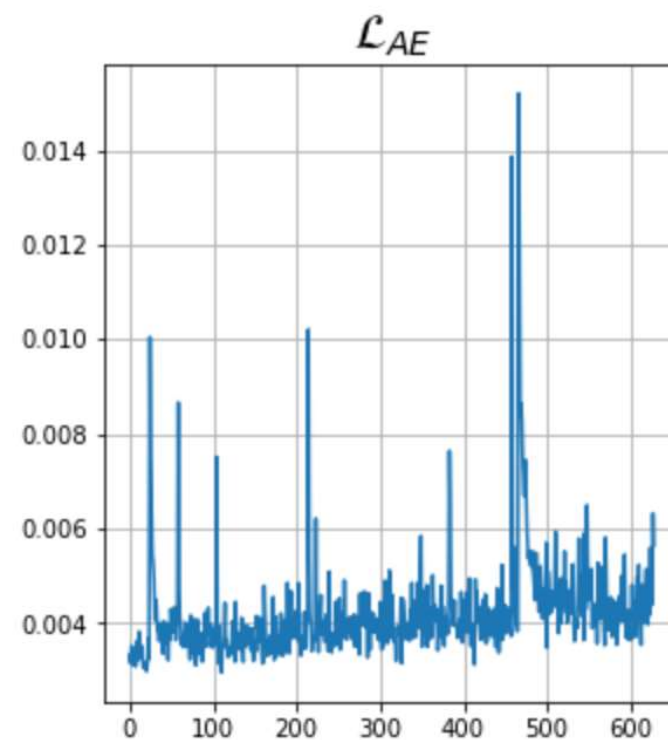
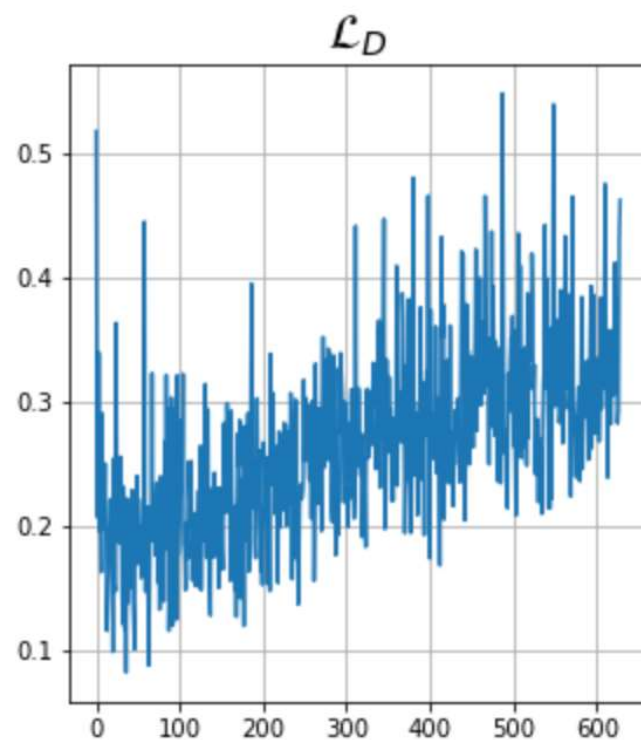
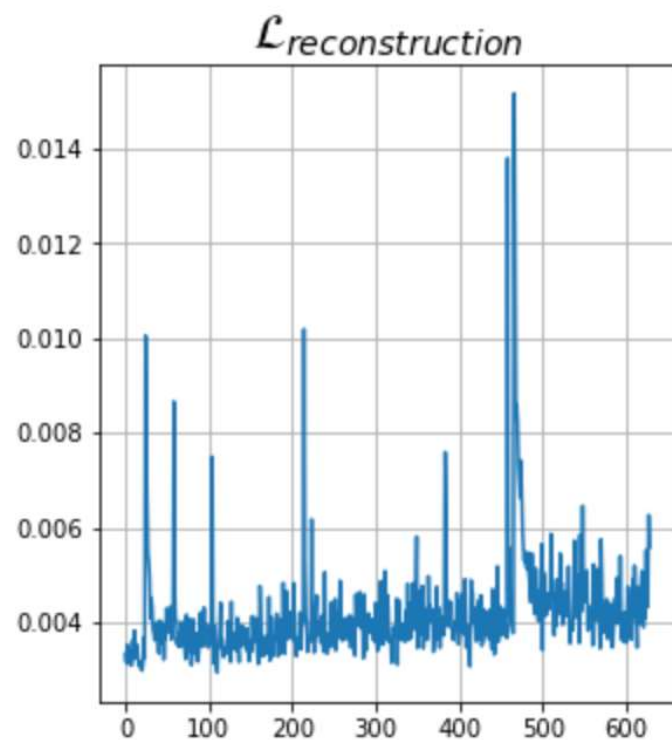
#2.1 Обучение АЕ с $\lambda = 0$



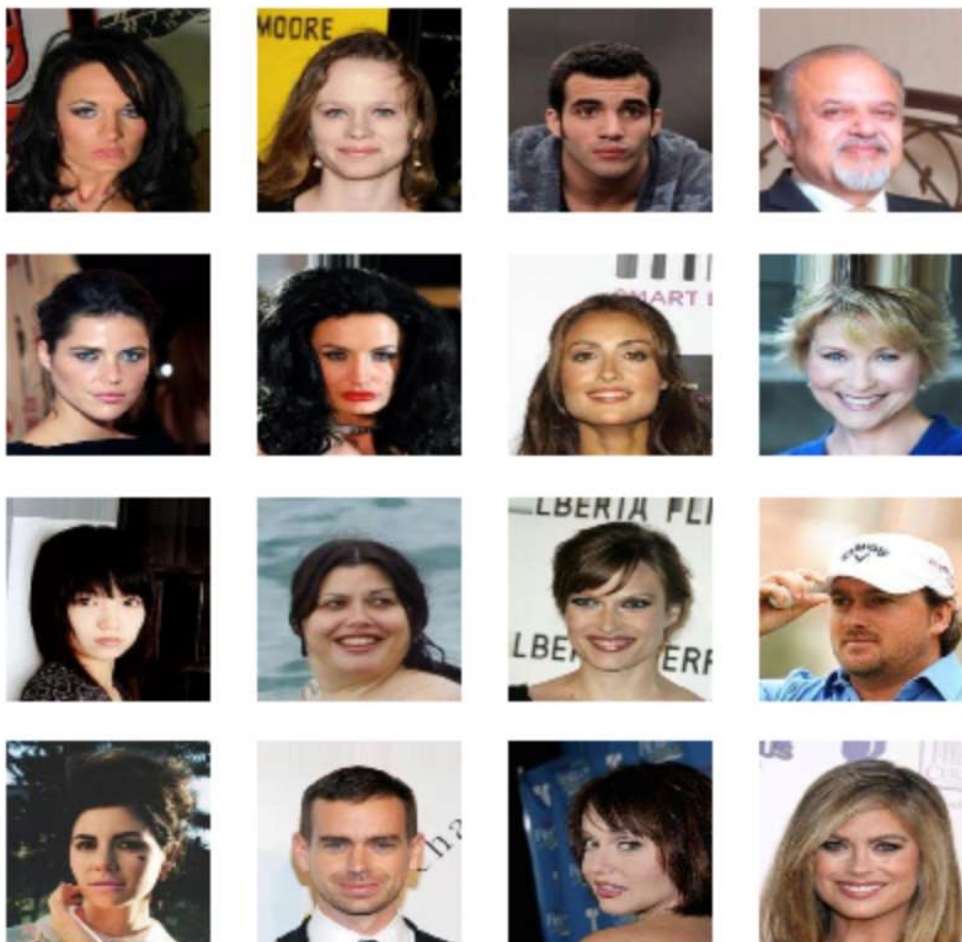
Модель делает милее... и красивее...



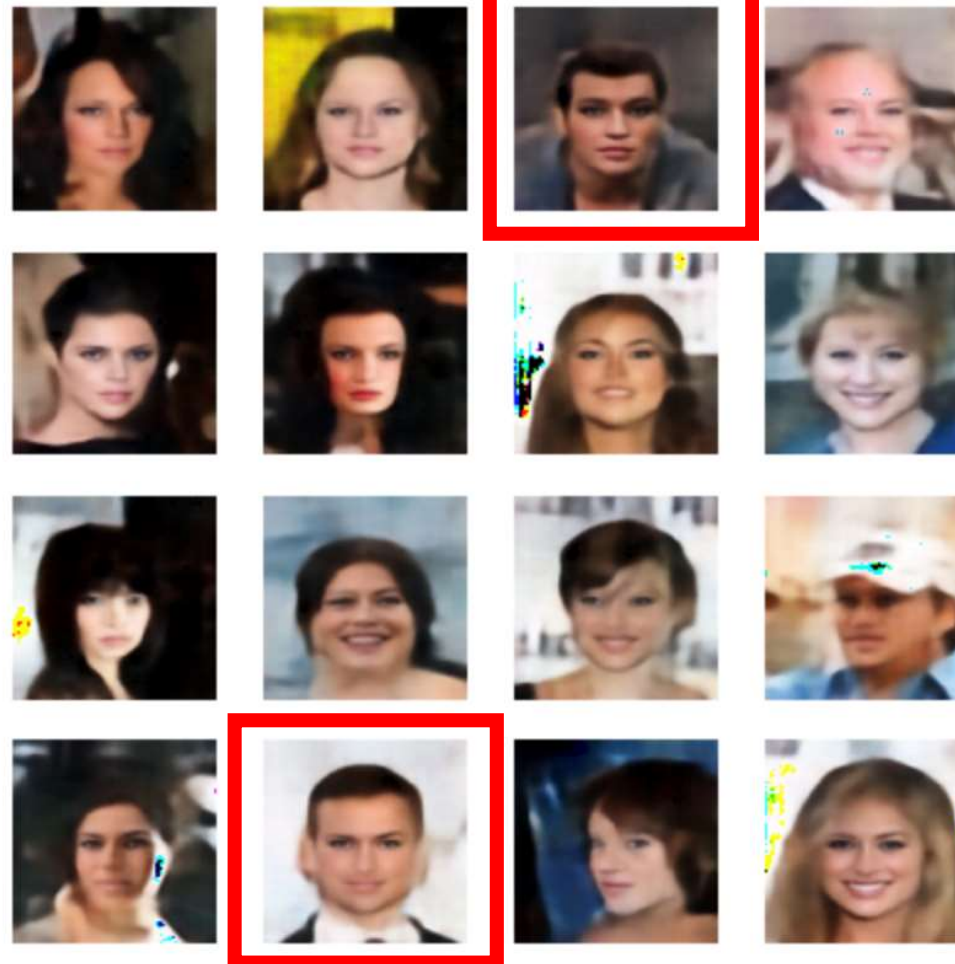
#2.2 Дообучение с линейным увеличением λ



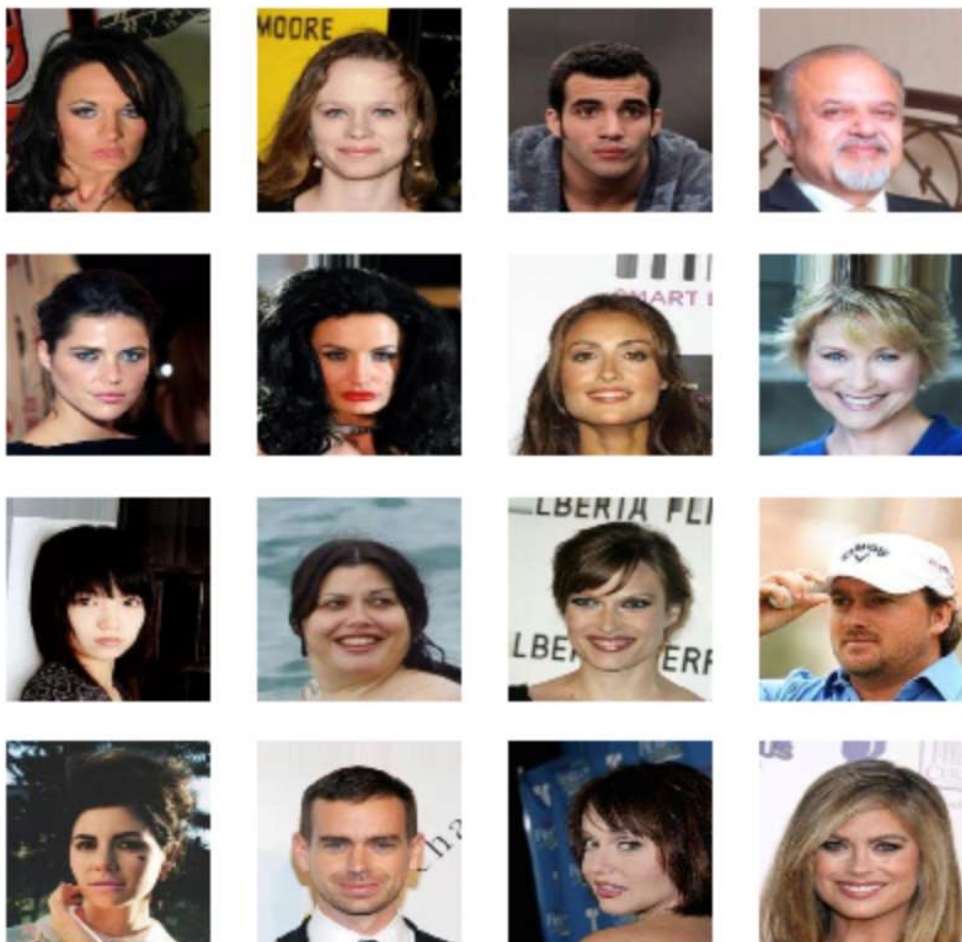
Начальные изображения



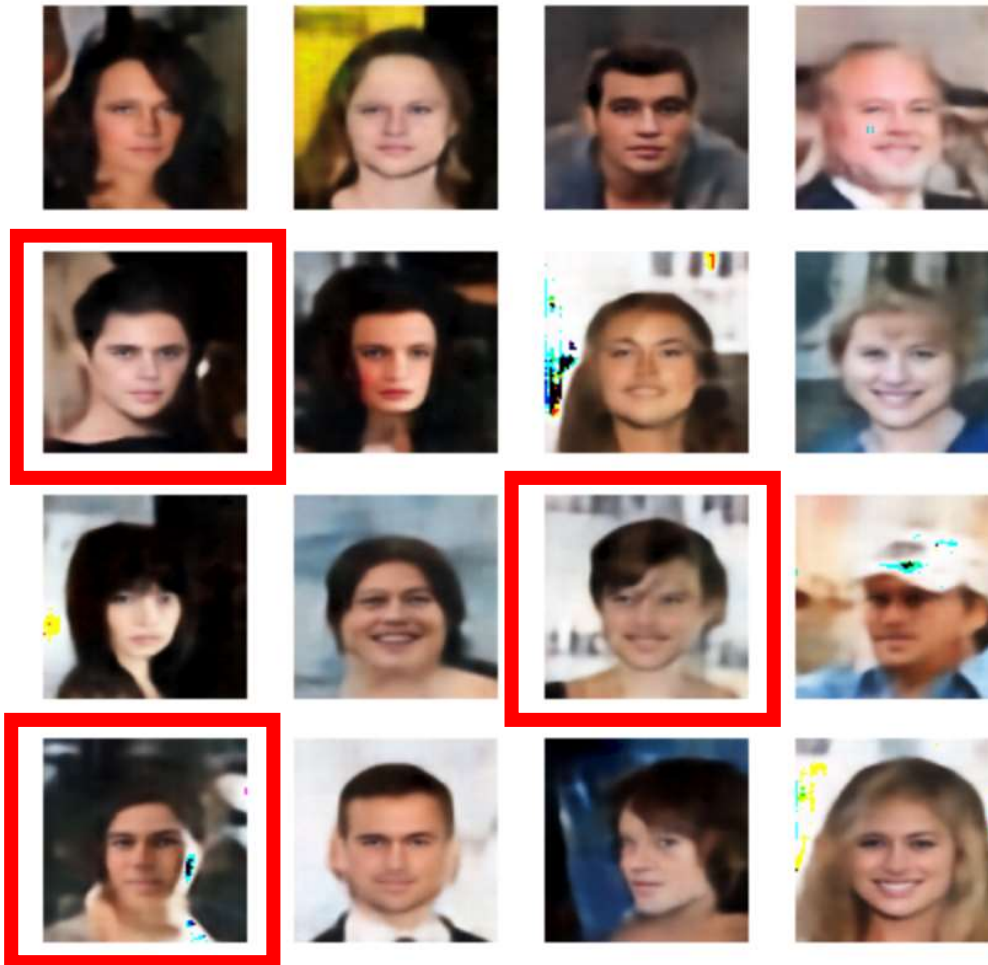
Женский пол

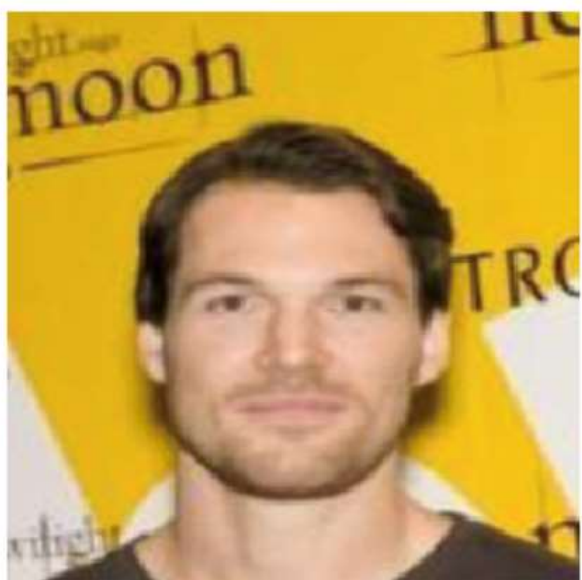
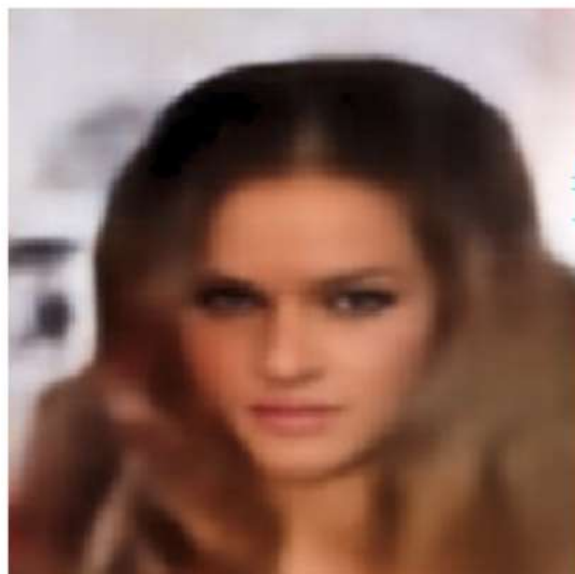


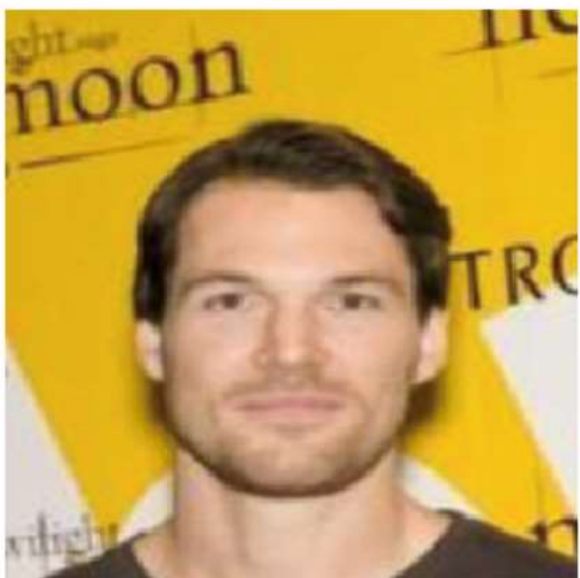
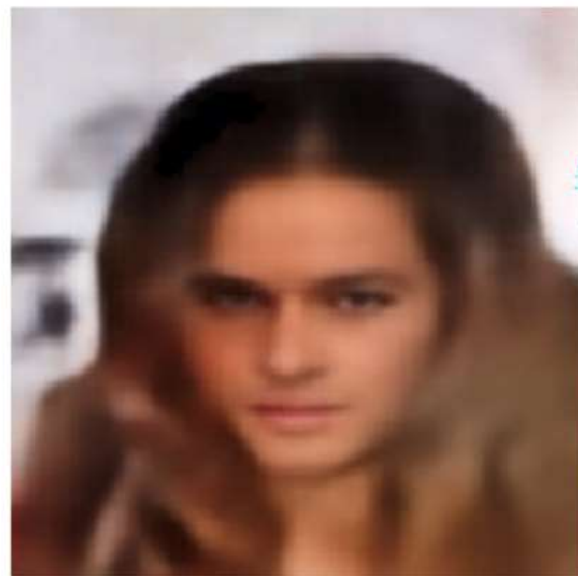
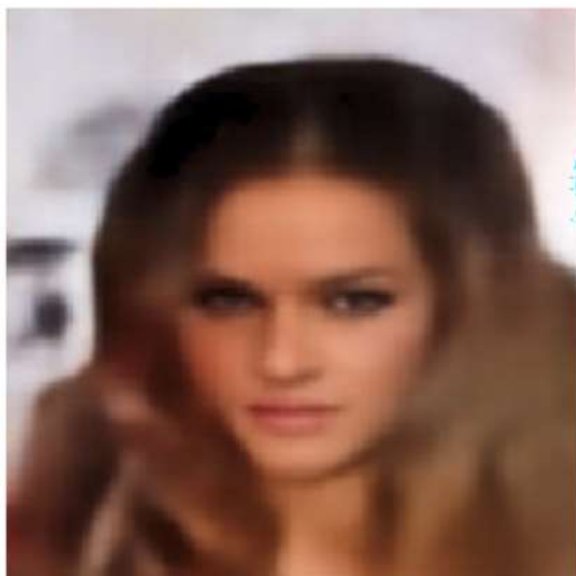
Начальные изображения



Мужской пол







Выводы

- Получена генерация лиц людей с помощью cGAN и FaderNetwork
- FaderNetwork воспроизводит более реалистичные картинки, которые отличаются от исходных, но позволяет переносить только пол
- С помощью cGAN картинки получились уродливее и в худшем разрешении, но такой подход позволил управлять еще и цветом волос
- Что еще можно было бы попробовать: cVAE

Спасибо за внимание!