

Цифровая обработка изображения

9. Генеративные конкурирующие сети (GAN)

Super resolution

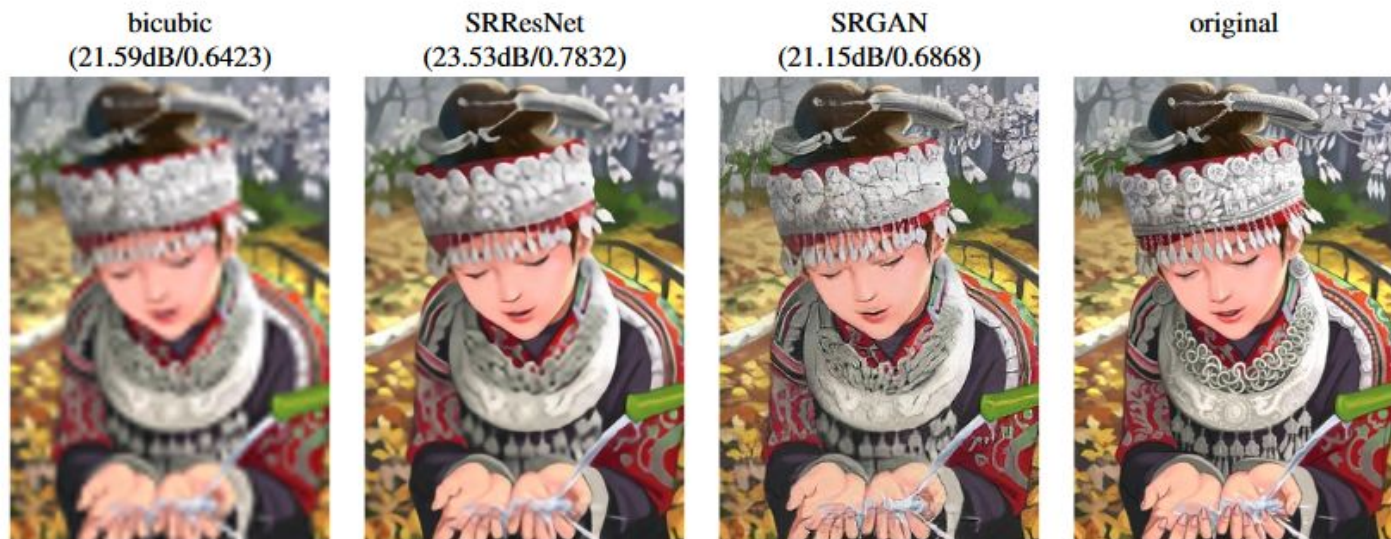


Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

Image-to-image

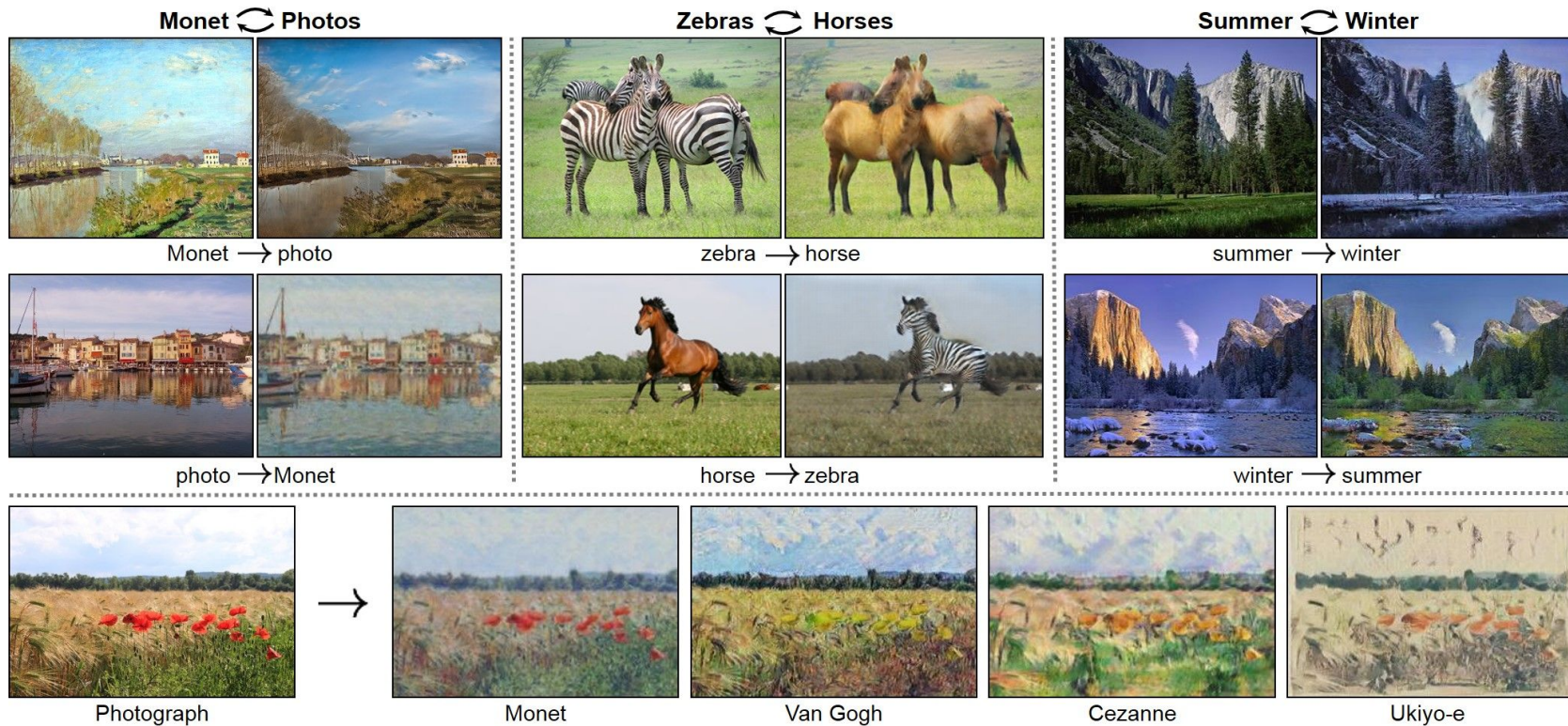
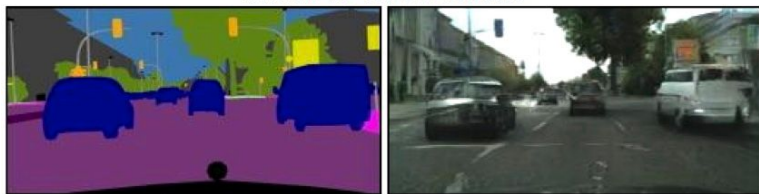


Image-to-image

Labels to Street Scene



input

output

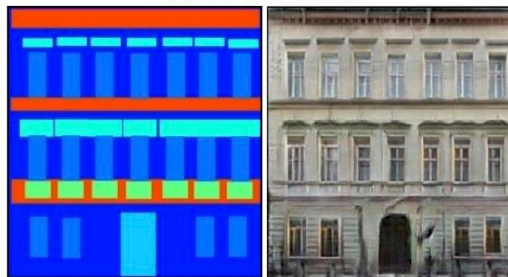
Aerial to Map



input

output

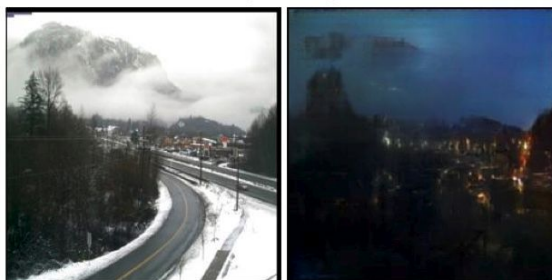
Labels to Facade



input

output

Day to Night



input

output

BW to Color



input

output

Edges to Photo



input

output

Image-to-image

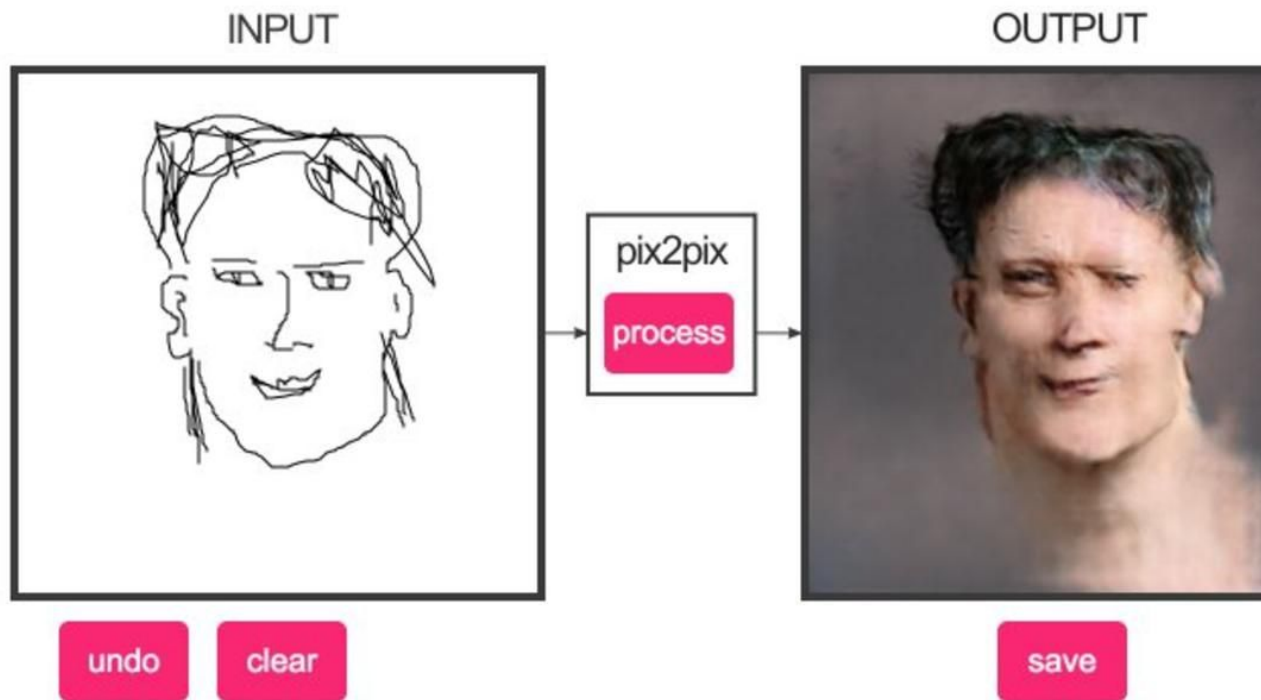
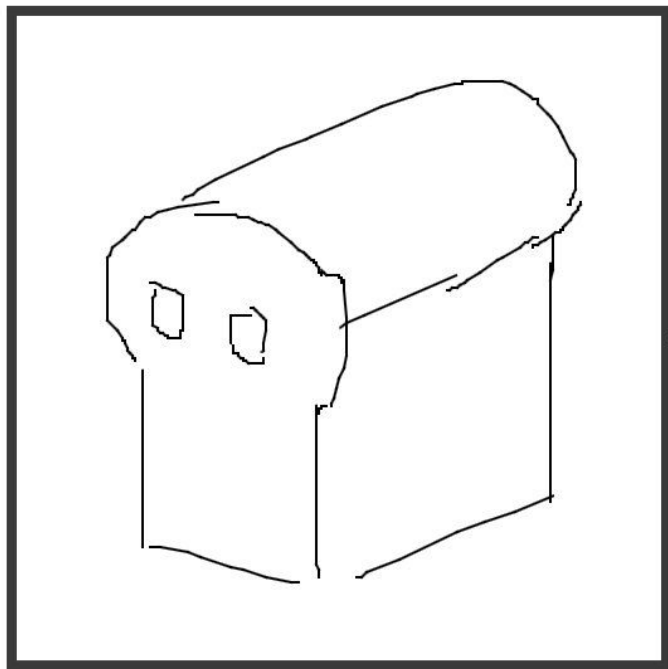


Image-to-image

INPUT



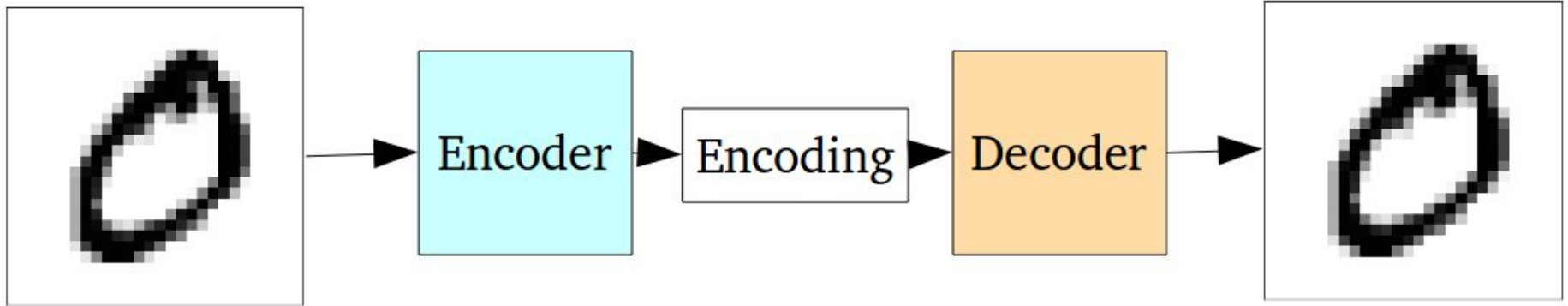
OUTPUT



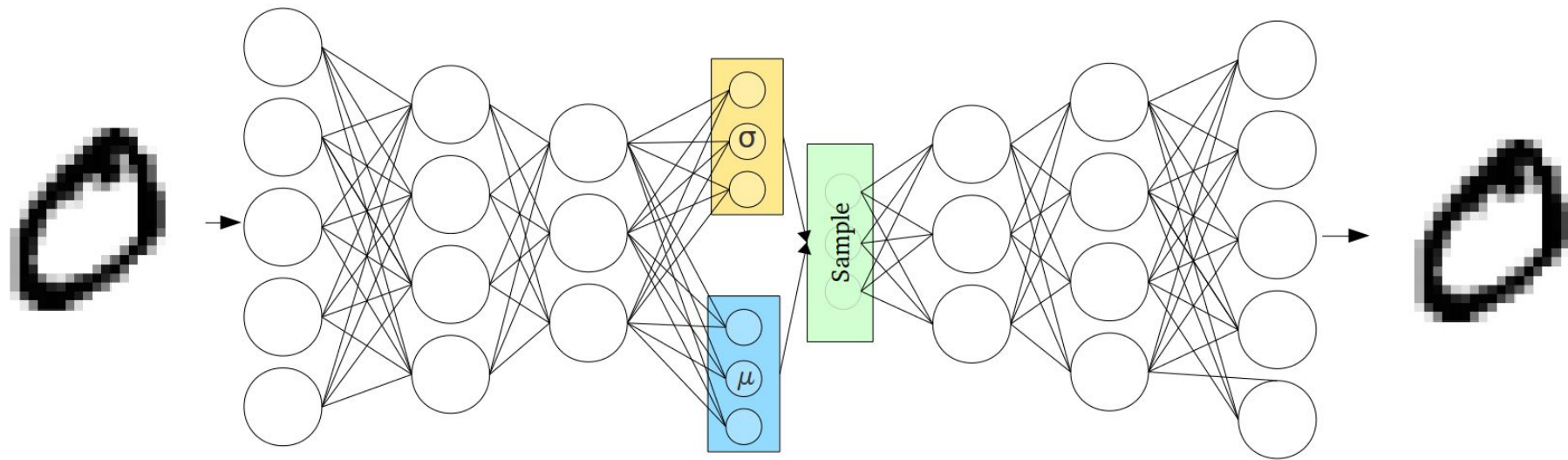
Variational Auto Encoding

[Auto-Encoding Variational Bayes](#)

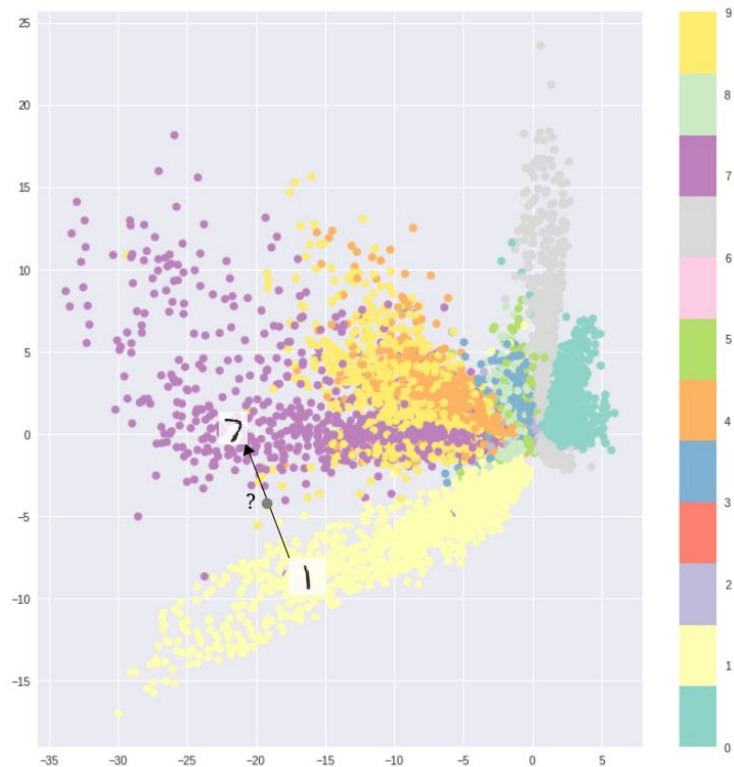
Variational Auto Encoding



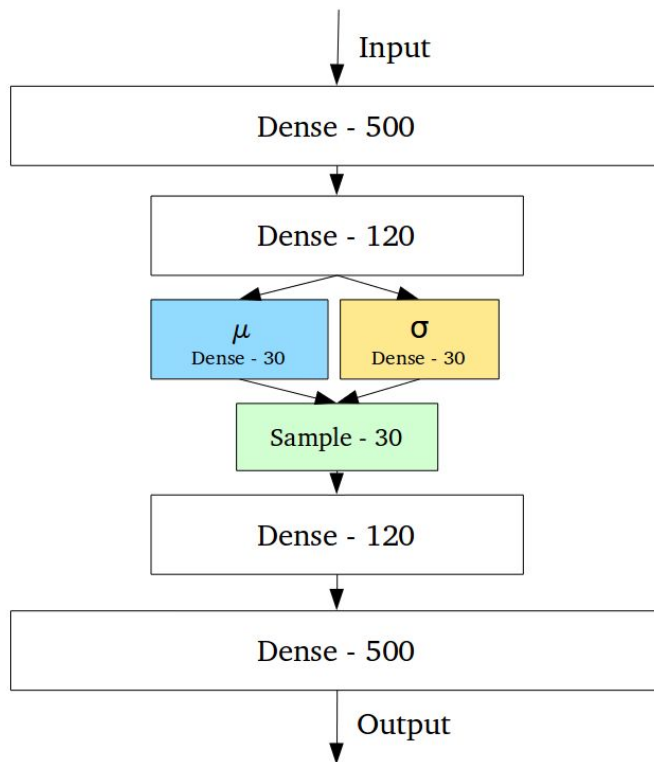
Variational Auto Encoding



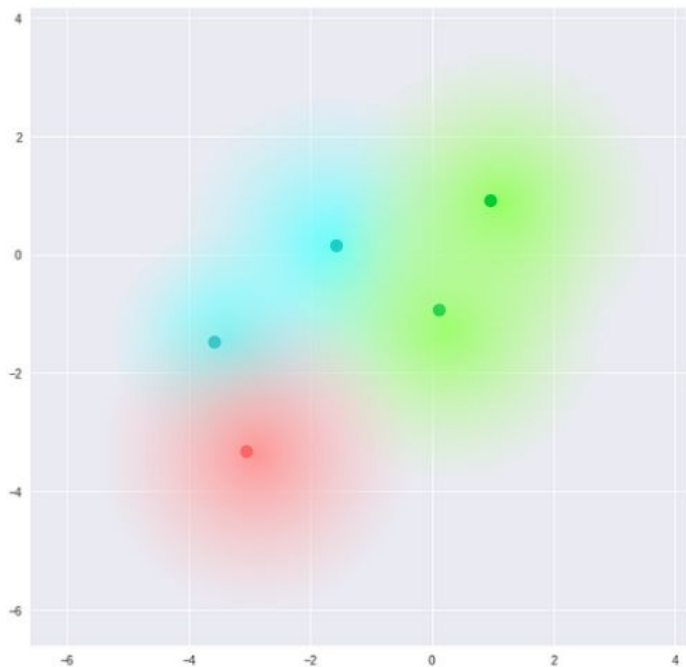
Скрытое представление



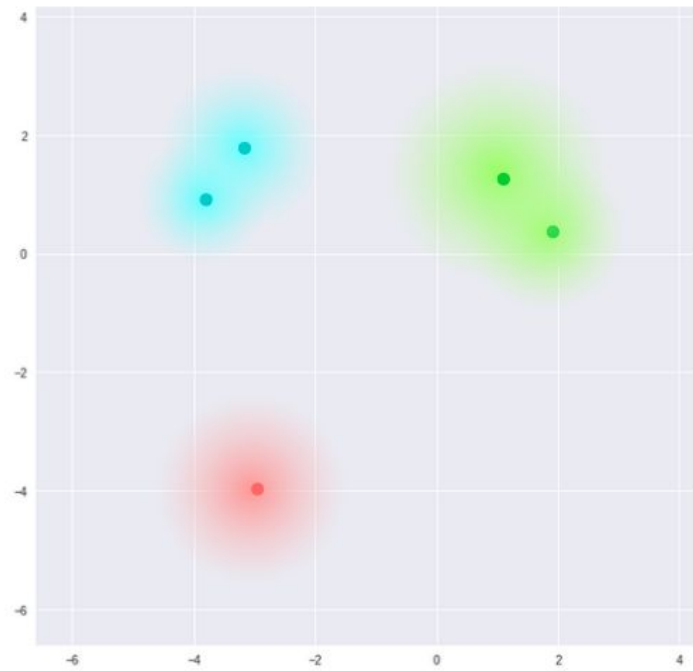
Variational Auto Encoding



Variational Auto Encoding



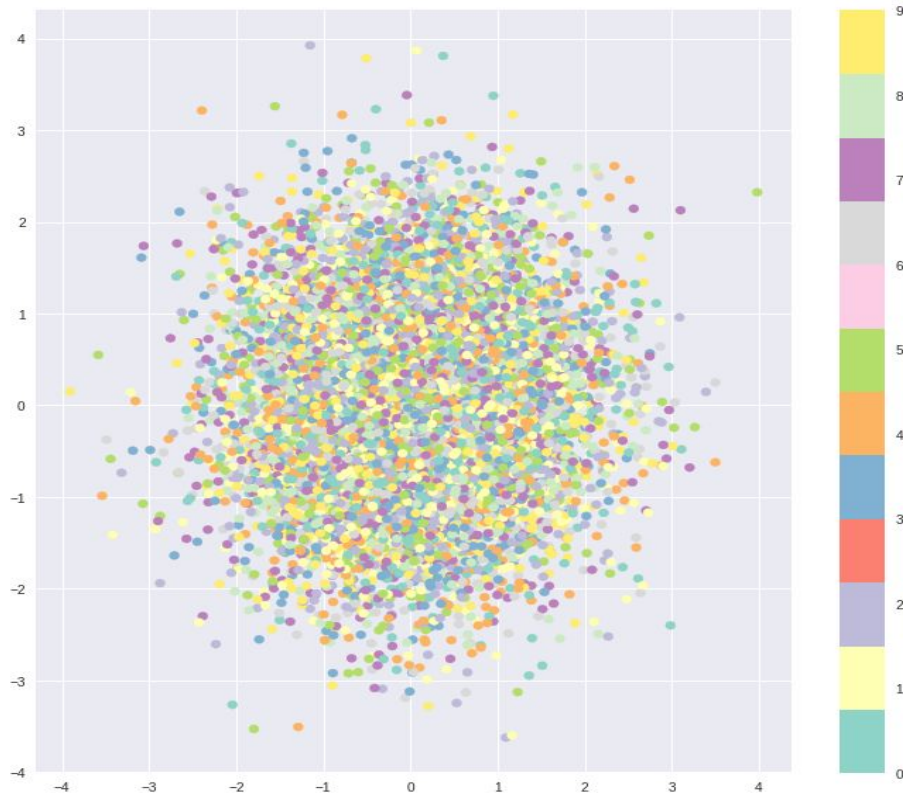
What we require



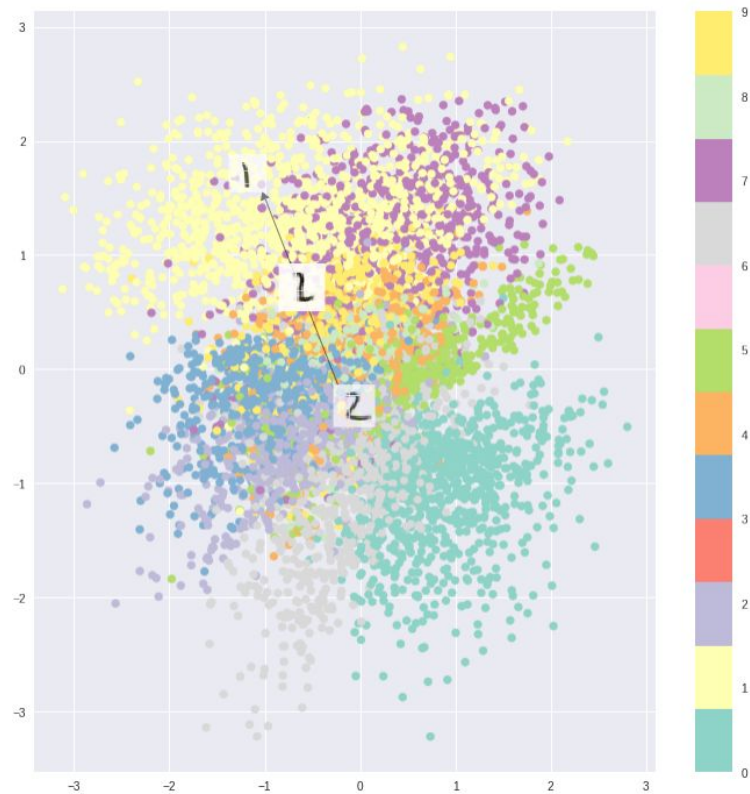
What we may inadvertently end up with

KL-divergence

$$\sum_{i=1}^n \sigma_i^2 + \mu_i^2 - \log(\sigma_i) - 1$$

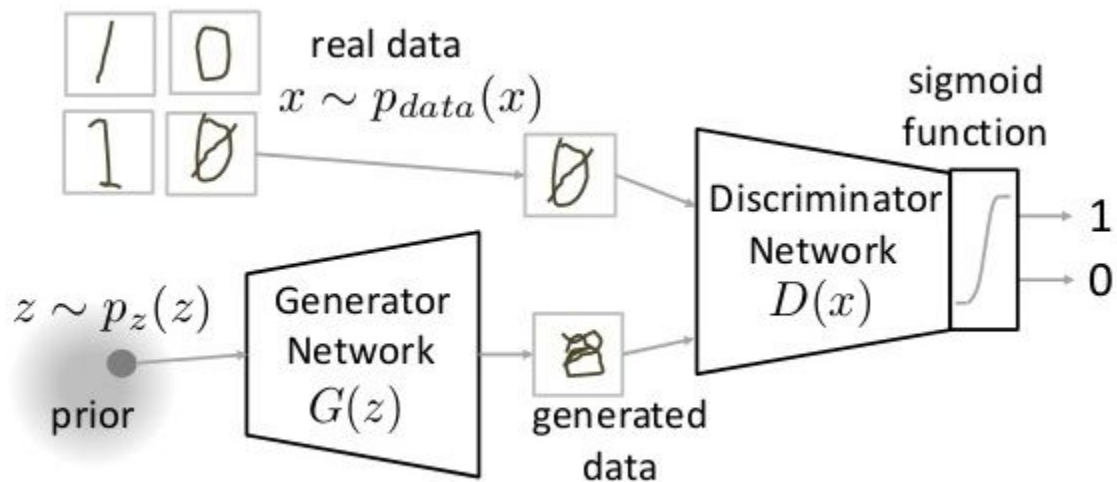


KL-divergence + reconstruction



Generative Adversarial Networks (GAN)

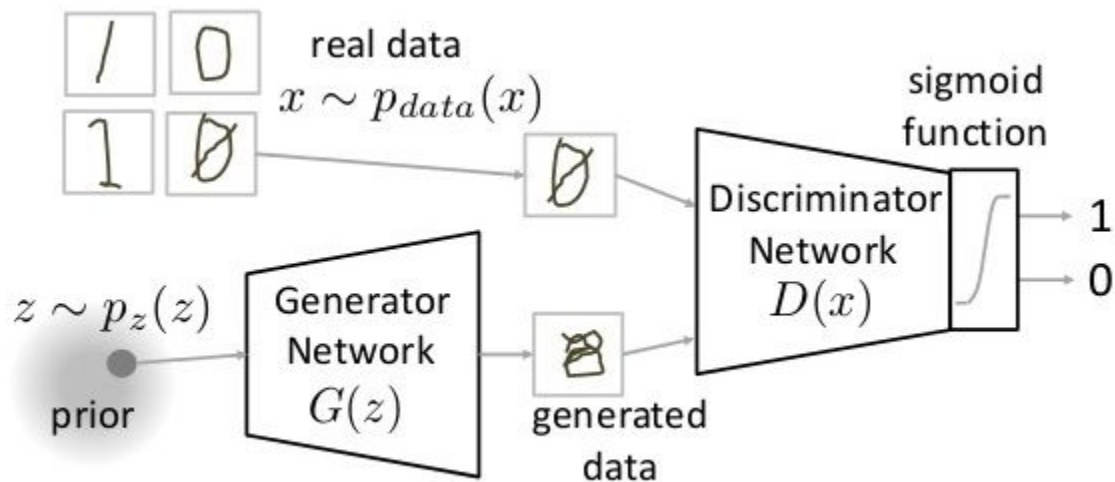
GAN



GAN

$$\min_G \max_D V(D, G)$$

$$V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$



GAN

$$\min_G \max_D V(D, G)$$

$$V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

V - функция потерь

x - наблюдение данных (изображение)

z - объект генерирующих данных

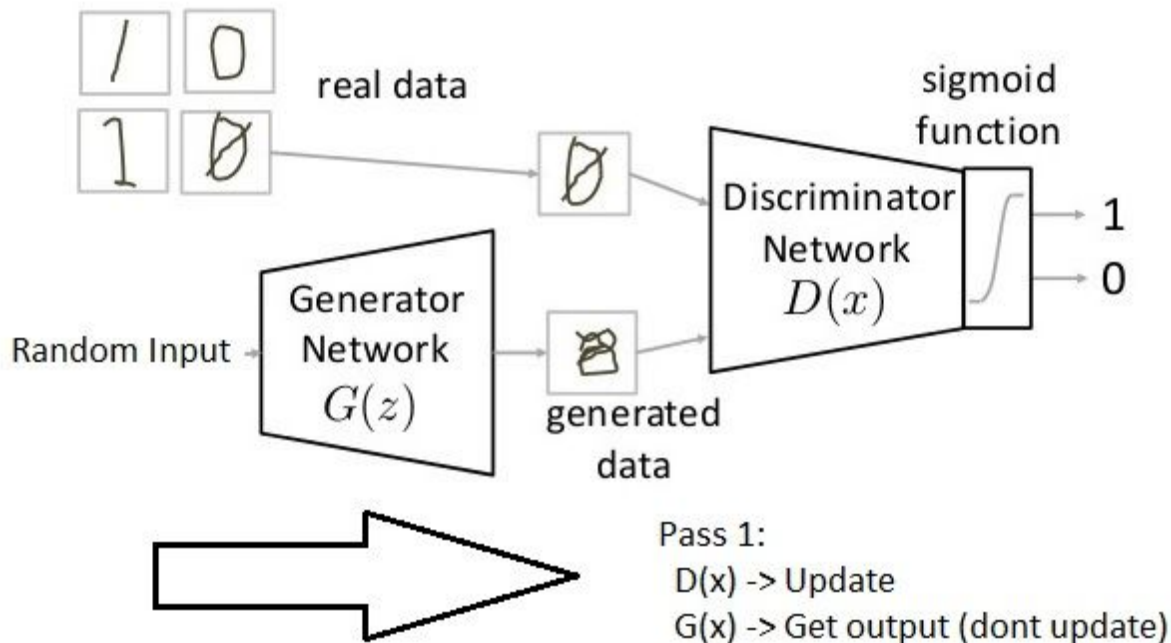
$D(x)$ - классификатор дискриминатора

$G(z)$ - выход генератора (изображение)

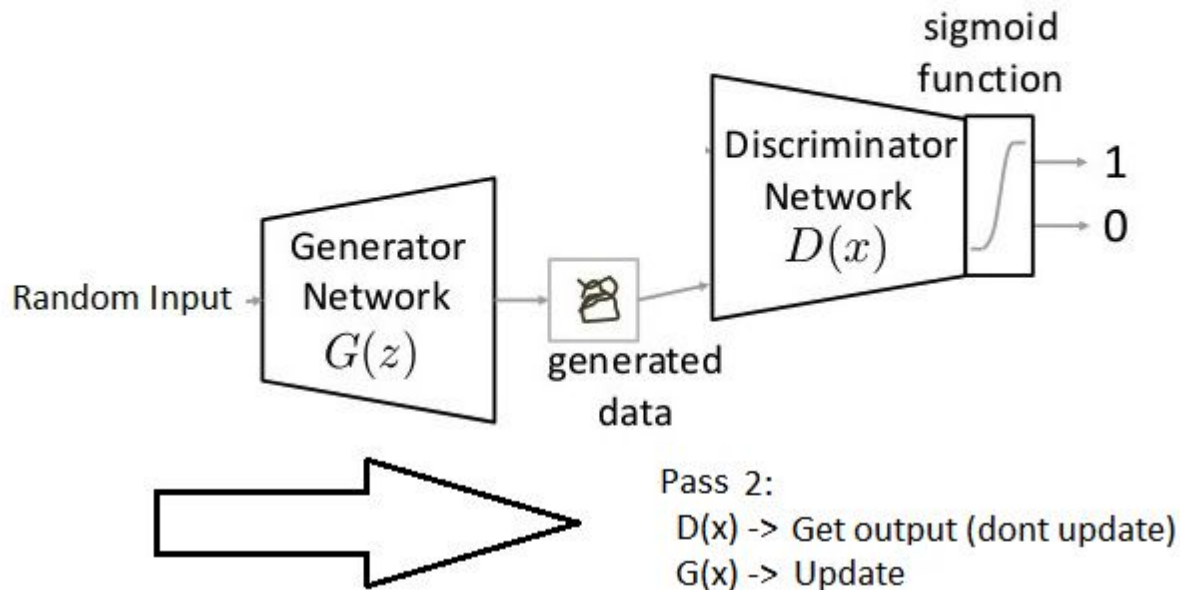
$p_{data}(x)$ - распределение исходных данных

$p_z(z)$ - распределение генерирующих данных

GAN - обучение модели



GAN - обучение модели



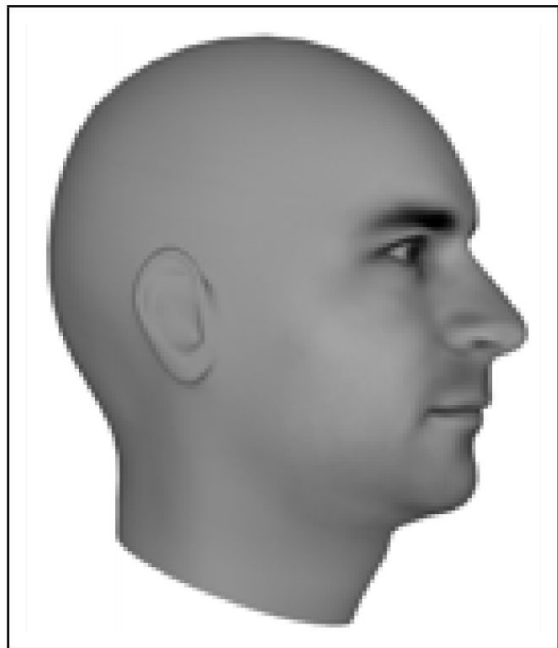
GAN - обучение модели

- выбираем данные для обучения (изображения, текст, аудио, видео)
- подбираем архитектуру модели в соответствии с данными
- фиксируем генератор, обучаем дискриминатор отличать исходные данные от сгенерированных
- фиксируем дискриминатор, обучаем генератор создавать данные похожие на исходные
- повторяем итерацию с переключением генератора и дискриминатора

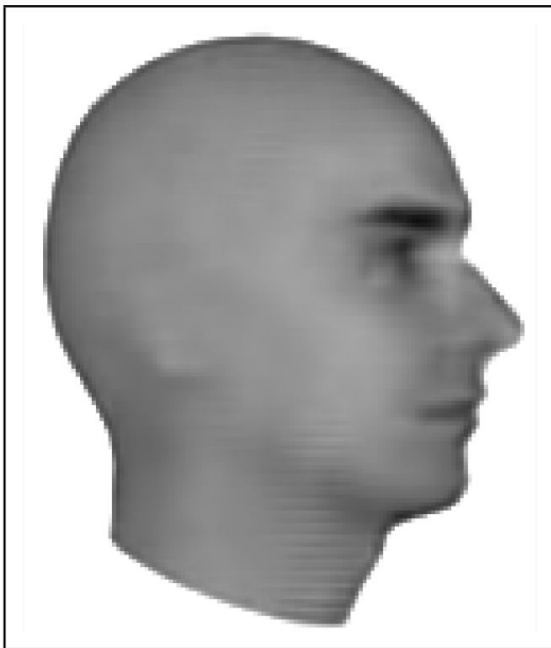
[Improved Techniques for Training GANs](#)

GAN - предсказание следующего кадра на видео

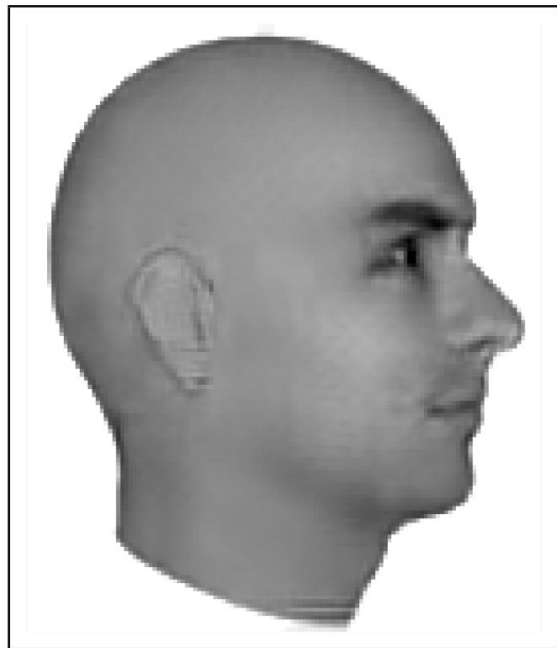
Ground Truth



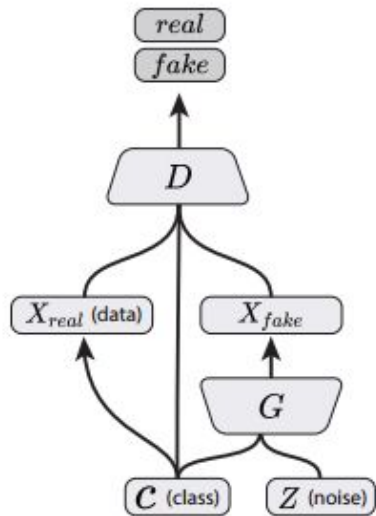
MSE



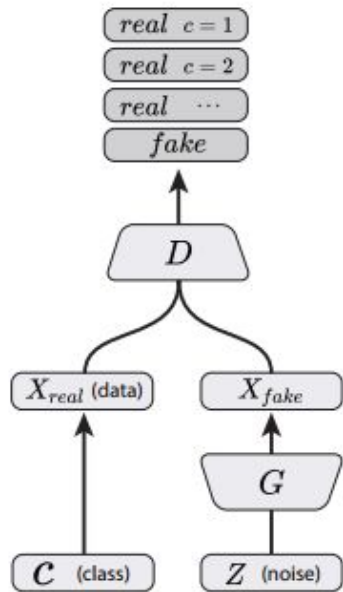
Adversarial



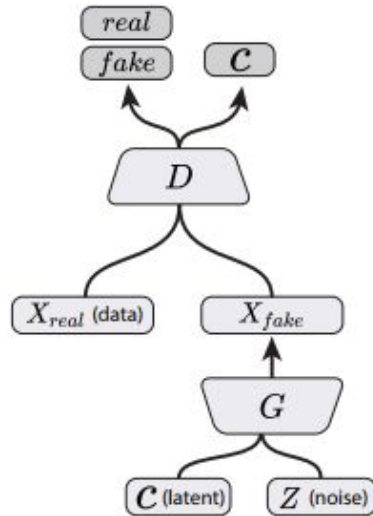
Архитектуры GAN



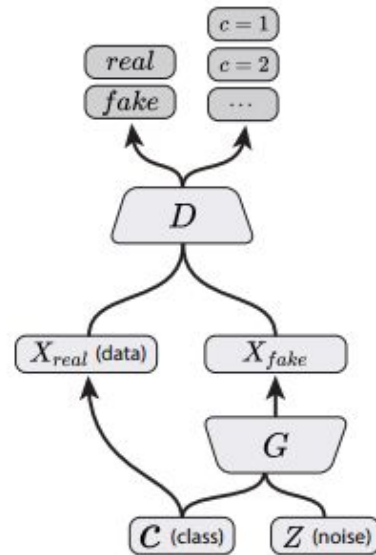
Conditional GAN



Semi-Supervised GAN



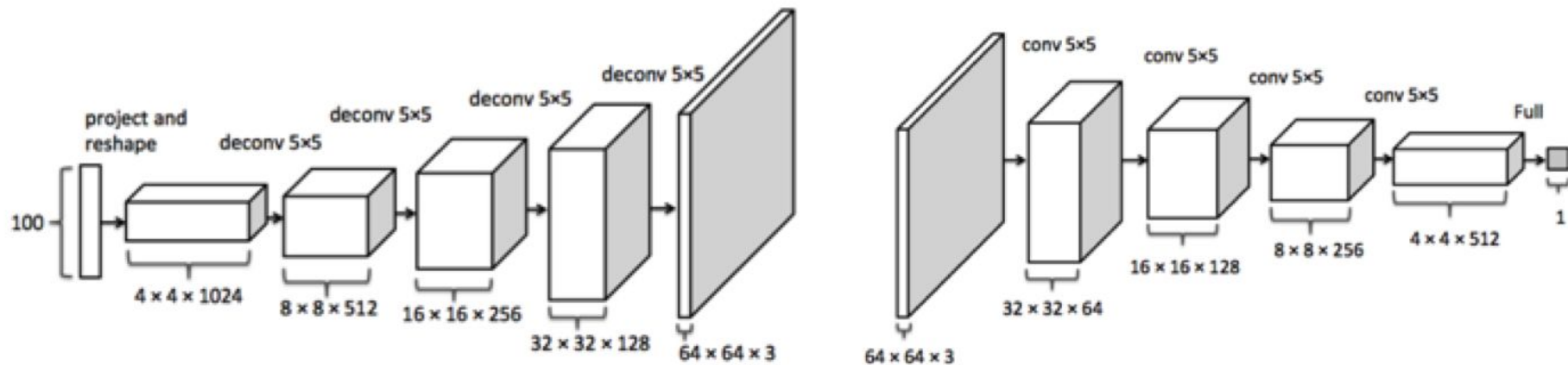
InfoGAN



AC-GAN

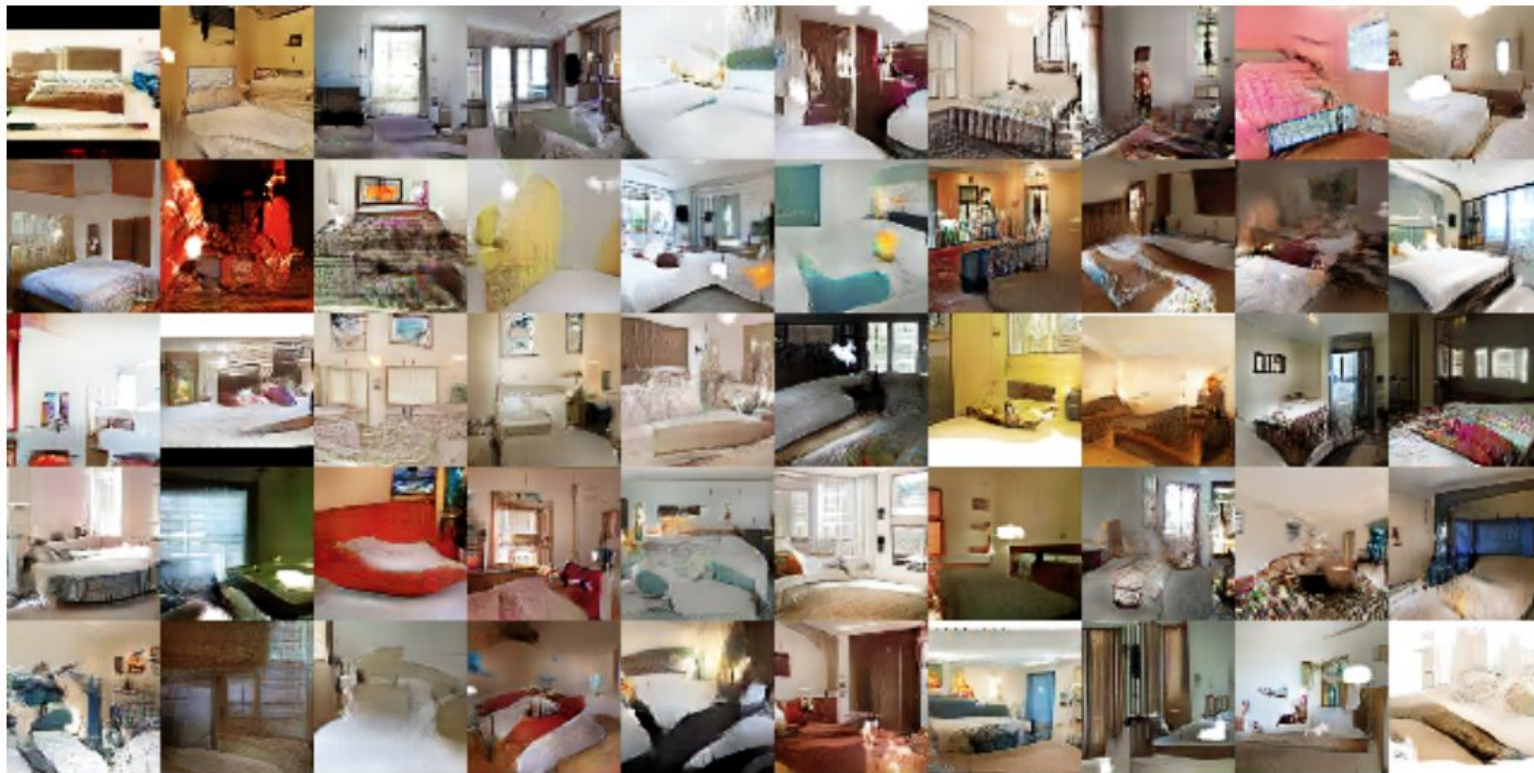
DCGAN - Deep Convolutional Generative Adversarial Network

DCGAN - векторное представление изображений

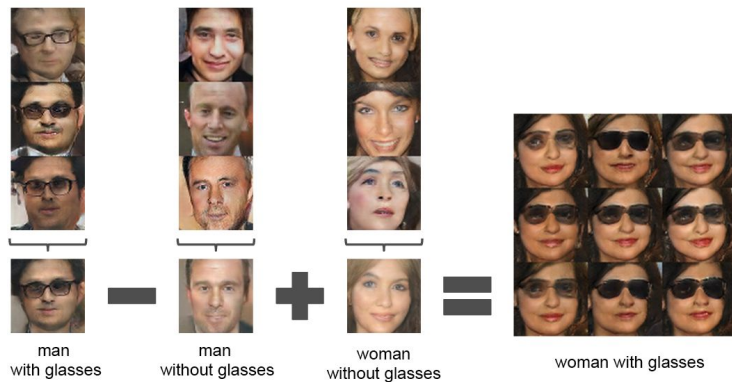
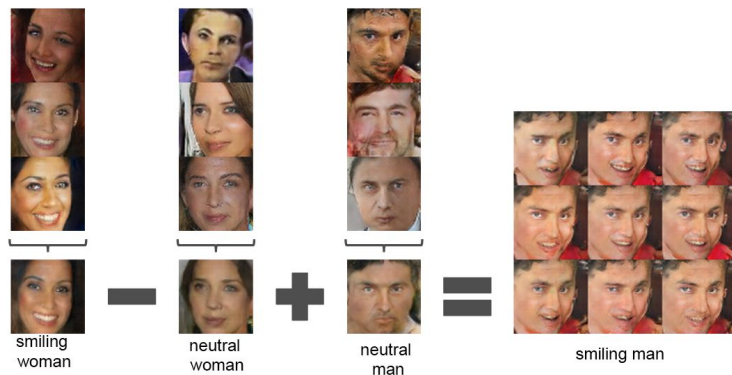


UNSUPERVISED REPRESENTATION LEARNING WITH DEEP
CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS

DCGAN - векторное представление изображений

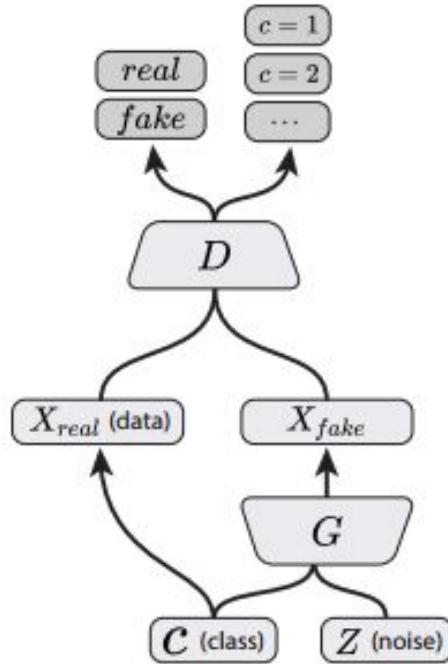


DCGAN - векторная арифметика



ACGAN - Auxiliary Classifier GANs

AC GAN - Auxiliary Classifier GANs



AC-GAN

AC GAN - Auxiliary Classifier GANs



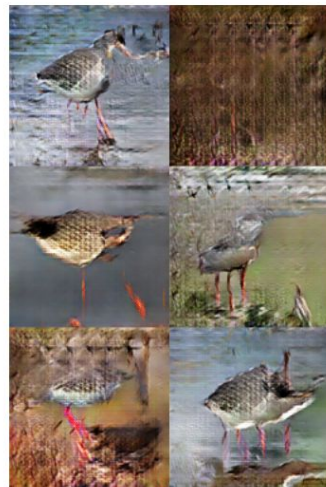
monarch butterfly



goldfinch



daisy



redshank



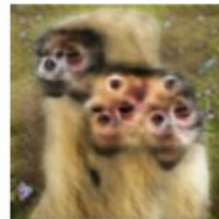
grey whale

GAN - сложности

- баланс сложности генератора/дискриминатора
- устойчивость процесса тренировки
- определение числа объектов для генерации
- ошибки перспективы - модель не получает информации о 3D форме объекта
- ошибки структуры - отсутствует информация о свойствах объекта
- возможны ошибки в случае, если пример существенно отличается от обучающей выборки

GAN - сложности

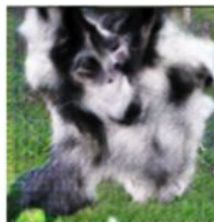
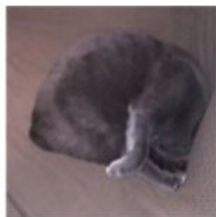
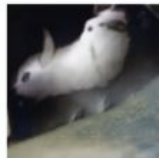
Problems with Counting



(Goodfellow 2016)

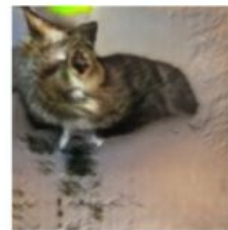
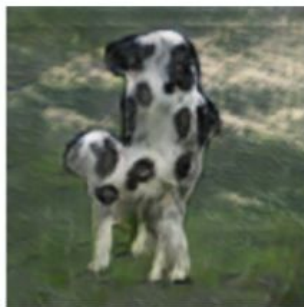
GAN - сложности

Problems with Perspective



GAN - сложности

Problems with Global Structure



GAN - сложности



Полезные материалы

- [Generative Adversarial Networks \(GANs\)](#)
- [Introductory guide to Generative Adversarial Networks \(GANs\)](#)
- [Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks](#)