

Цифровая обработка изображения

9. Генеративные конкурирующие сети (GAN)

Super resolution

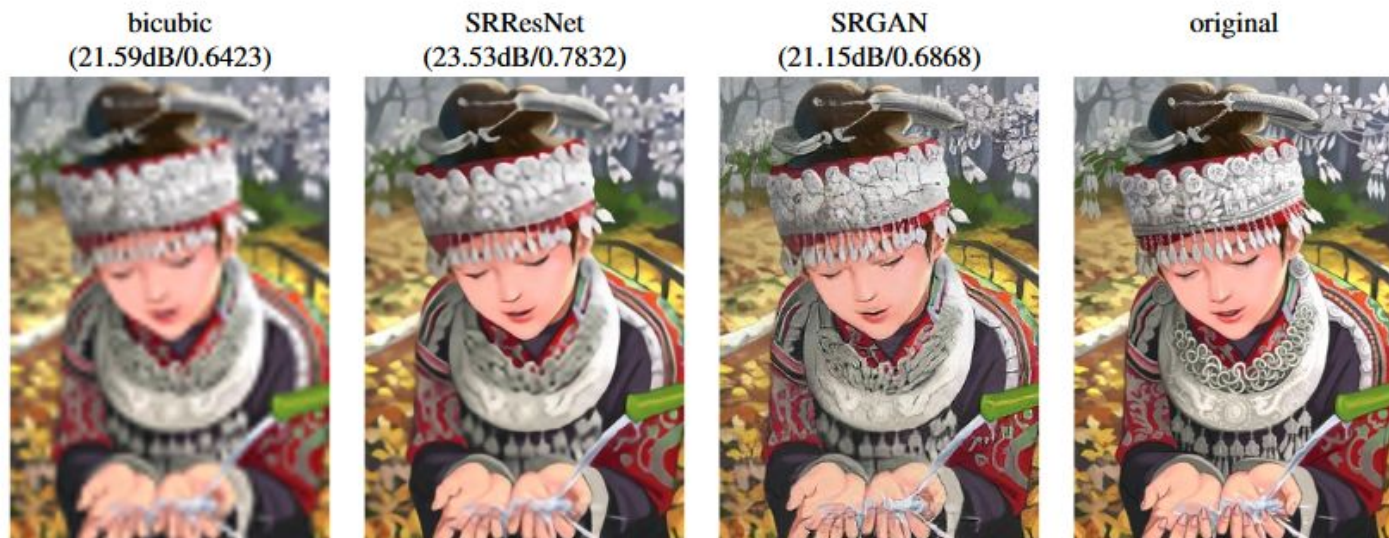


Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

Image-to-image

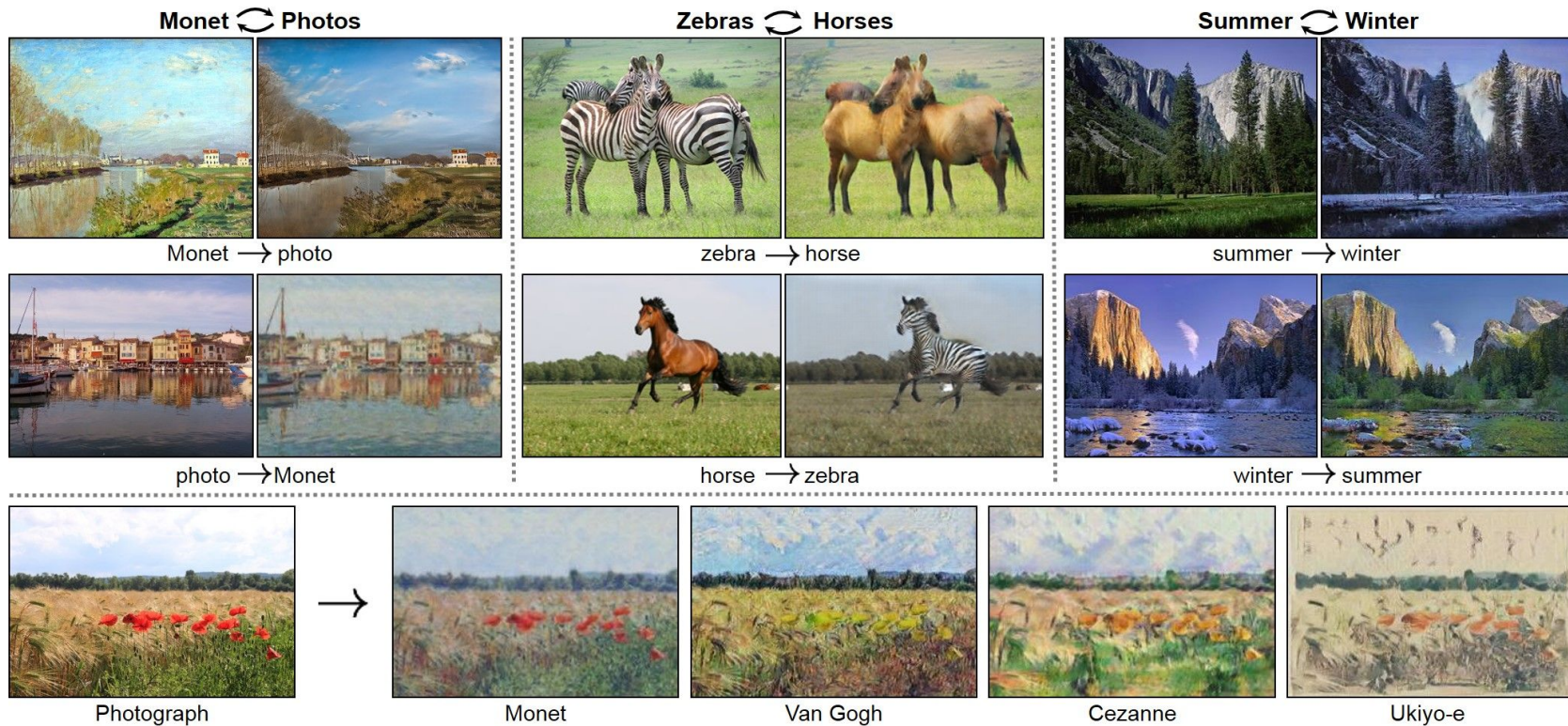
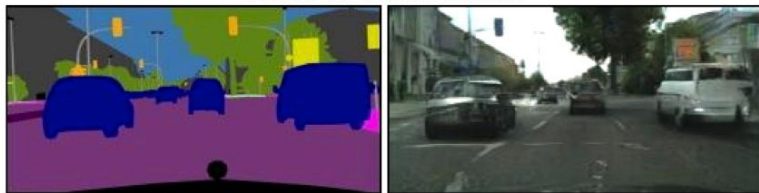


Image-to-image

Labels to Street Scene



input

output

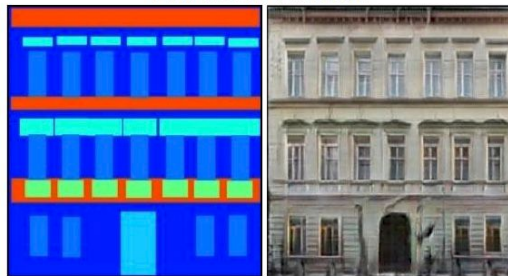
Aerial to Map



input

output

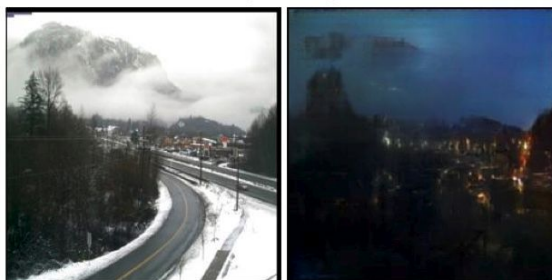
Labels to Facade



input

output

Day to Night



input

output

BW to Color



input

output

Edges to Photo



input

output

Image-to-image

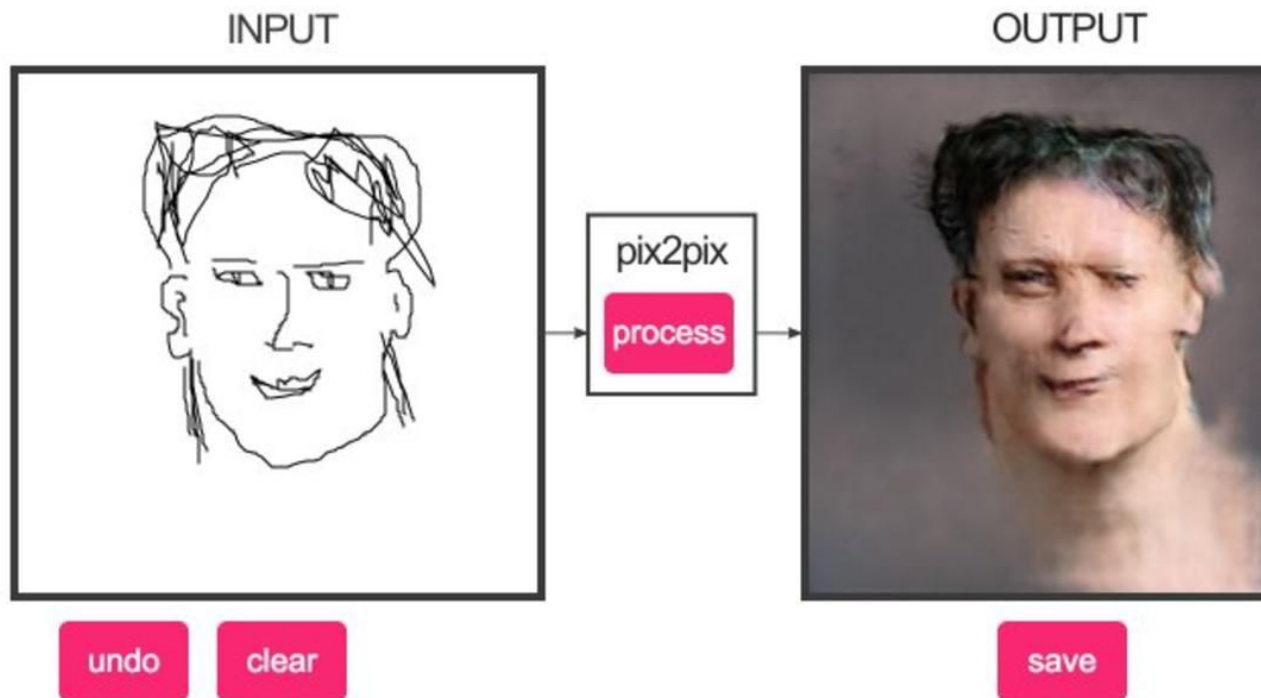
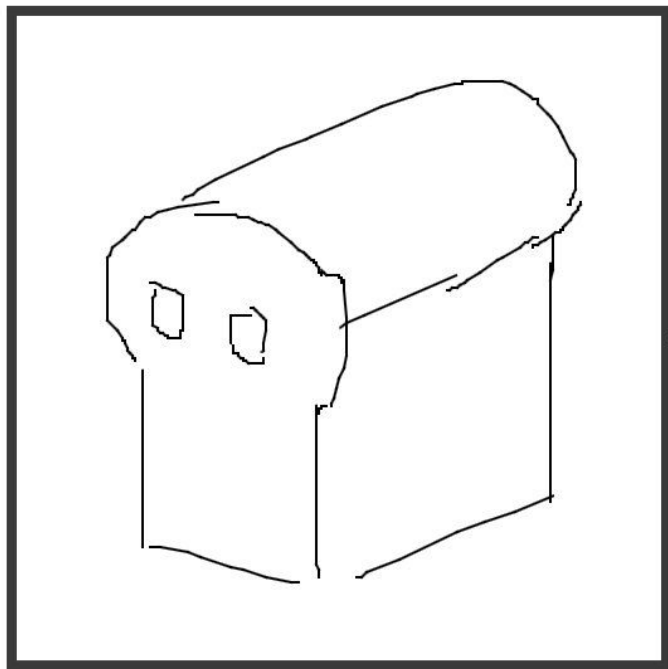


Image-to-image

INPUT

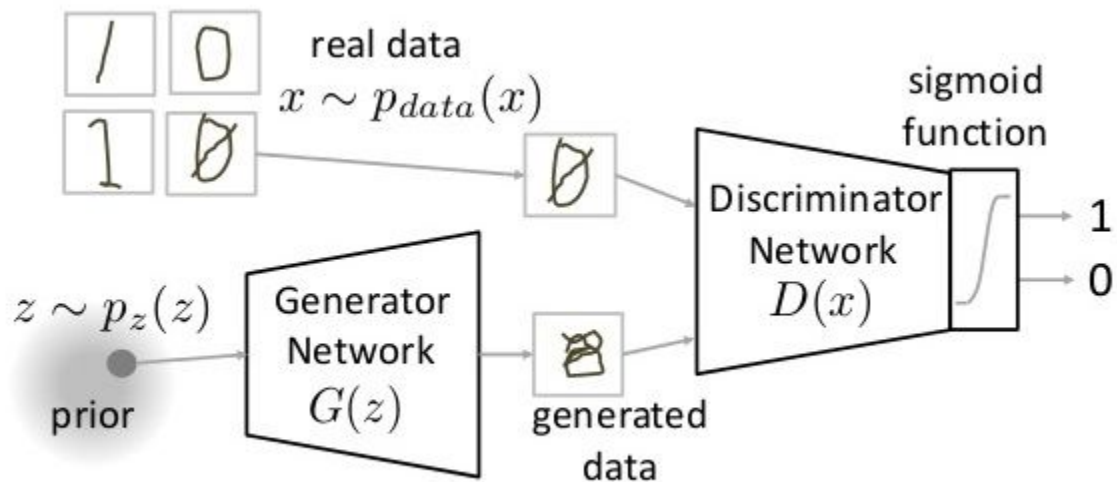


OUTPUT



Generative Adversarial Networks (GAN)

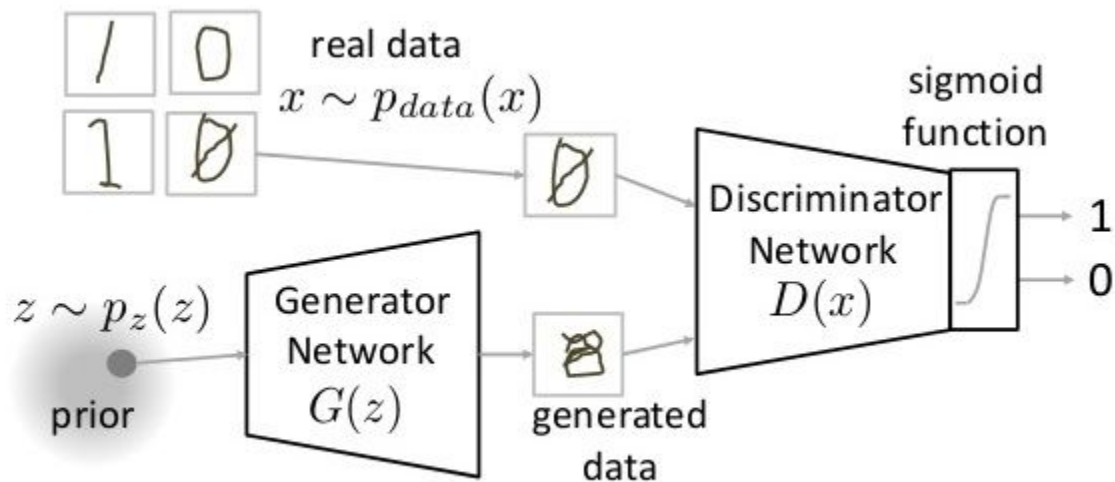
GAN



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

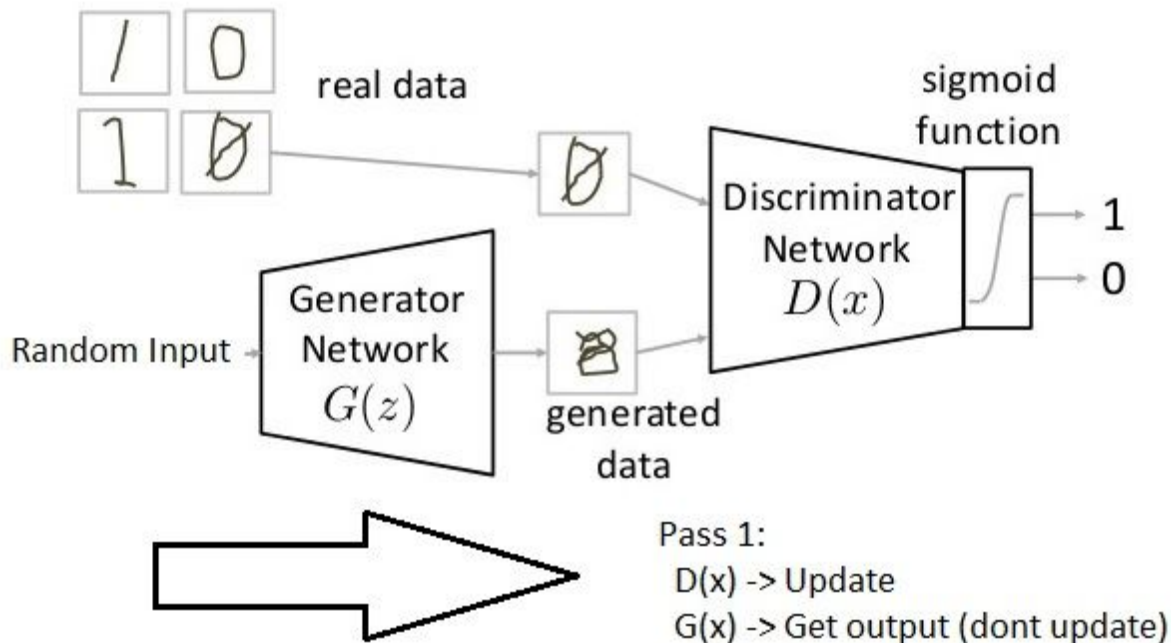
$D(x)$ - классификатор дискриминатора

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

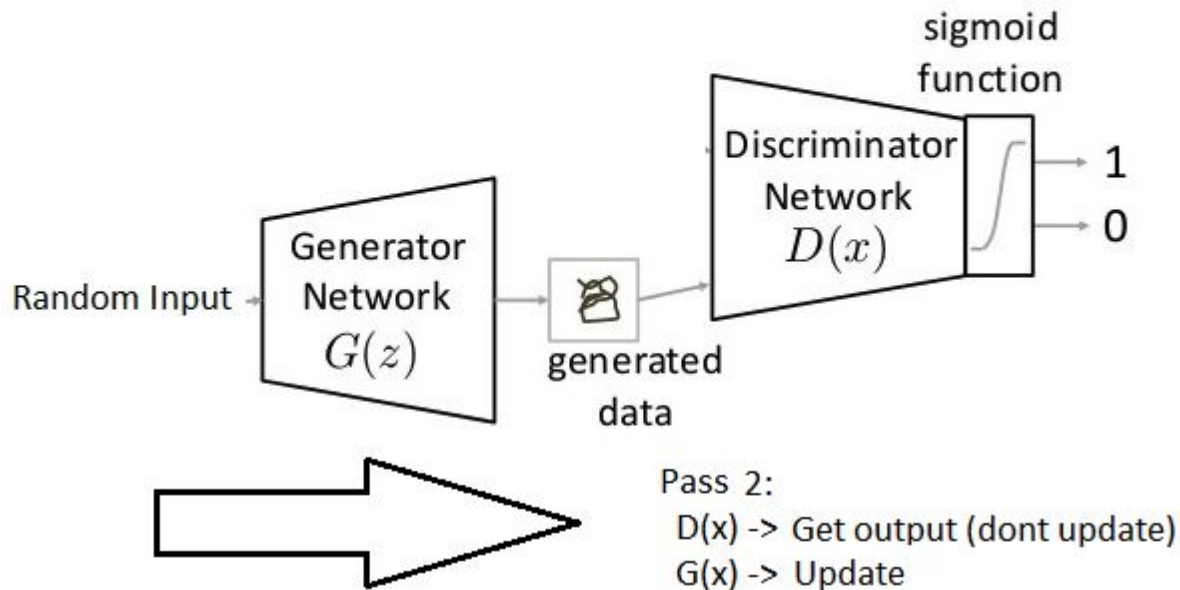
$p_{data}(x)$ - распределение исходных данных

$p_z(z)$ - распределение генерирующих данных

GAN - обучение модели



GAN - обучение модели

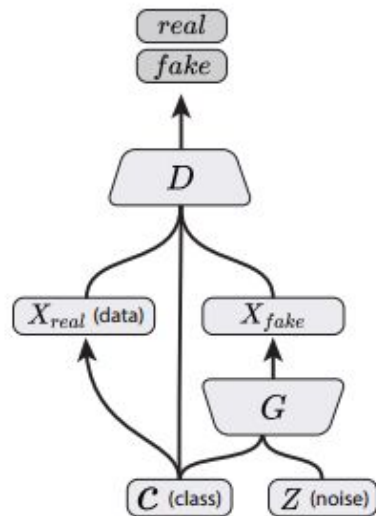


GAN - обучение модели

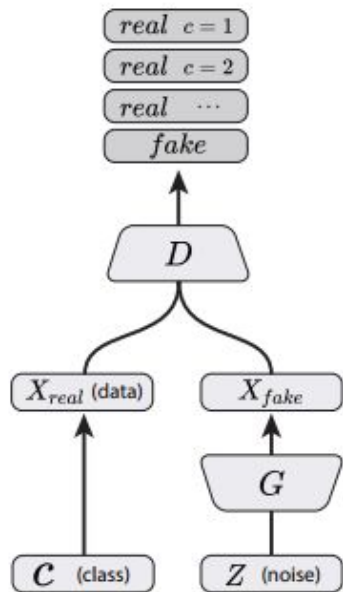
- выбираем данные для обучения (изображения, текст, аудио, видео)
- подбираем архитектуру модели в соответствии с данными
- фиксируем генератор, обучаем дискриминатор отличать исходные данные от сгенерированных
- фиксируем дискриминатор, обучаем генератор создавать данные похожие на исходные
- повторяем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

[Improved Techniques for Training GANs](#)

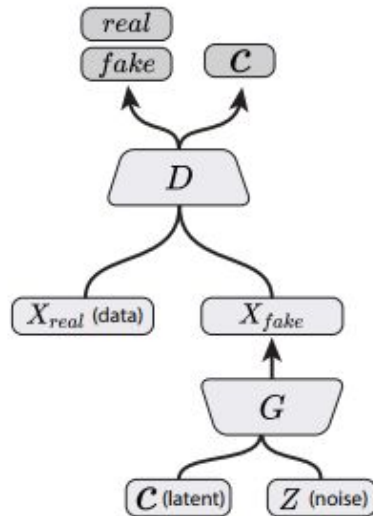
GAN



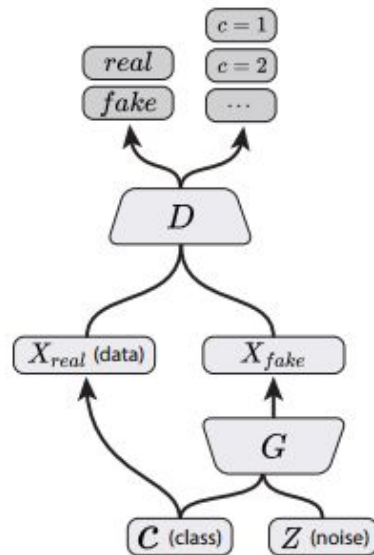
Conditional GAN
(Mirza & Osindero, 2014)



Semi-Supervised GAN
(Odena, 2016; Salimans, et al., 2016)



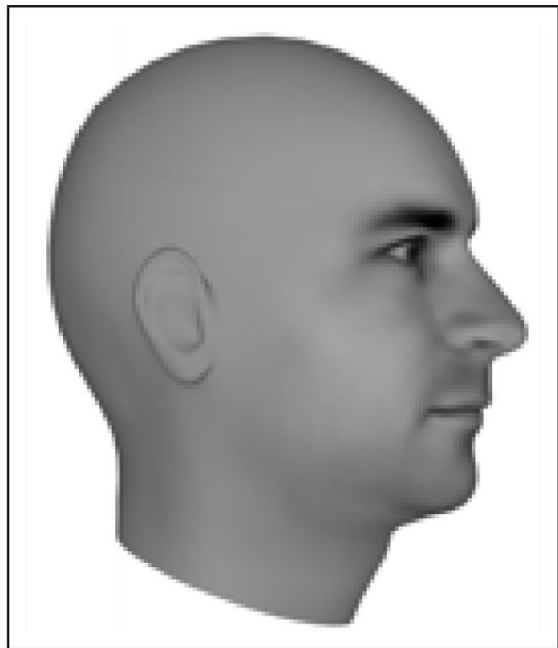
InfoGAN
(Chen, et al., 2016)



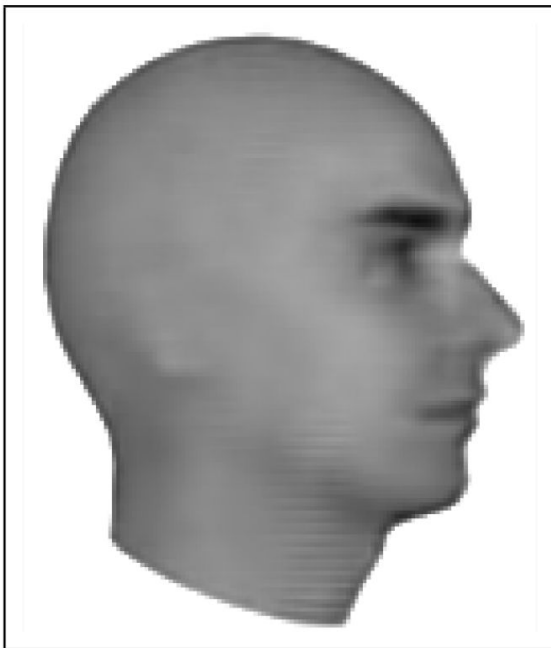
AC-GAN
(Present Work)

GAN - предсказание следующего кадра на видео

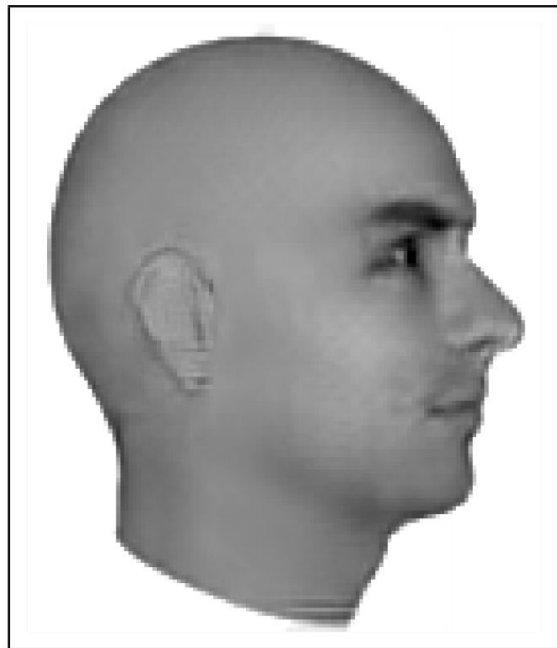
Ground Truth



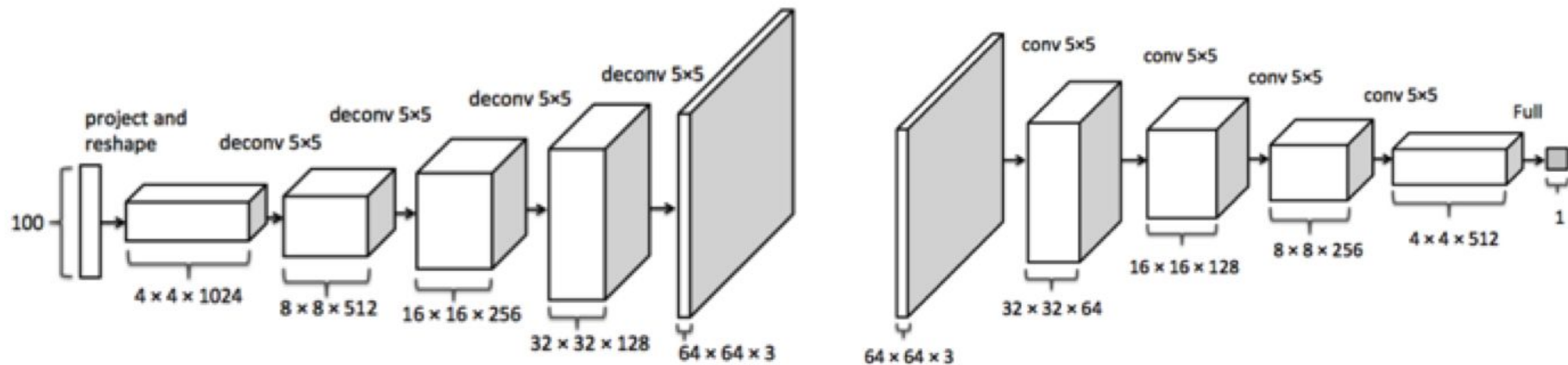
MSE



Adversarial

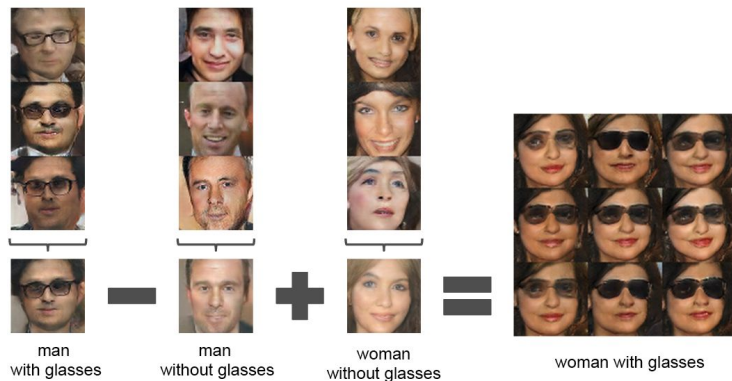
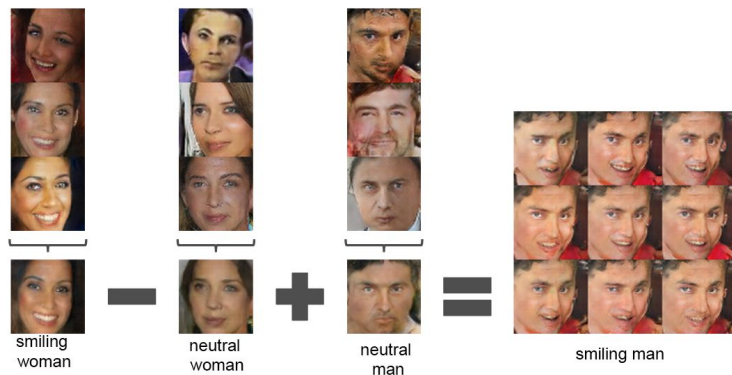


DCGAN - векторное представление изображений



UNSUPERVISED REPRESENTATION LEARNING WITH DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS

DCGAN - векторное представление изображения

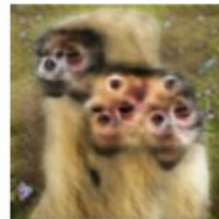


GAN - сложности

- 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

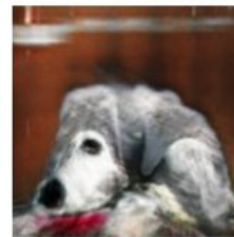
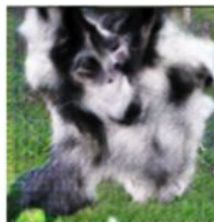
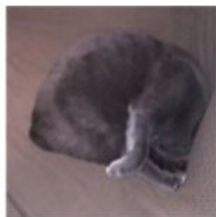
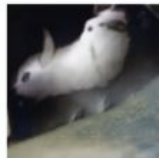
Problems with Counting



(Goodfellow 2016)

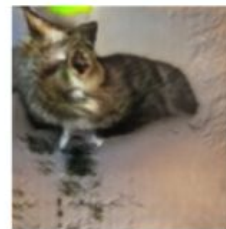
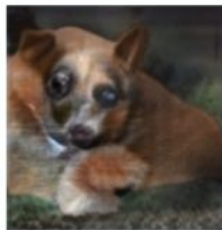
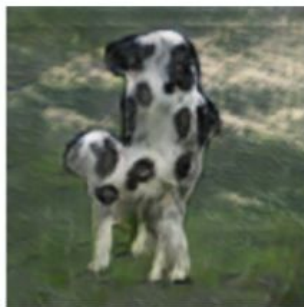
GAN - сложности

Problems with Perspective



GAN - сложности

Problems with Global Structure



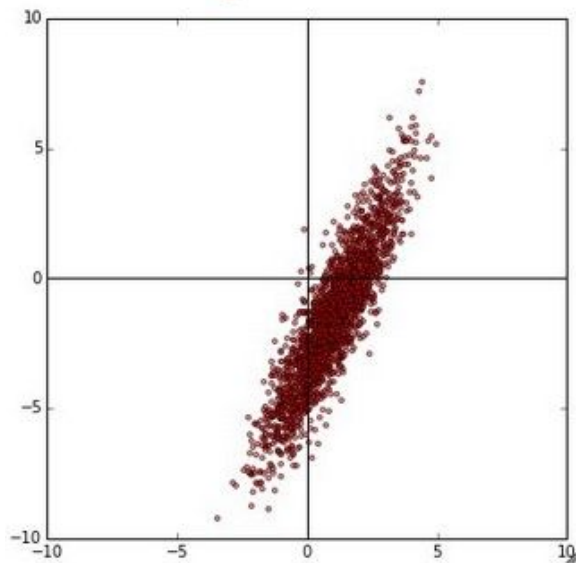
GAN - сложности



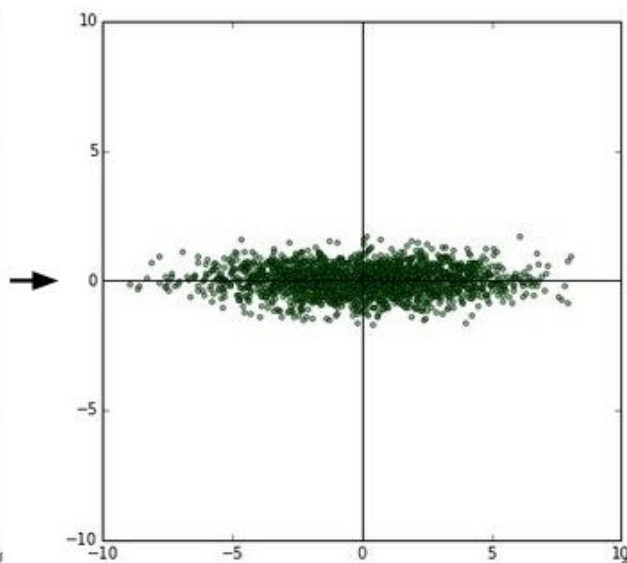
PCA

PCA

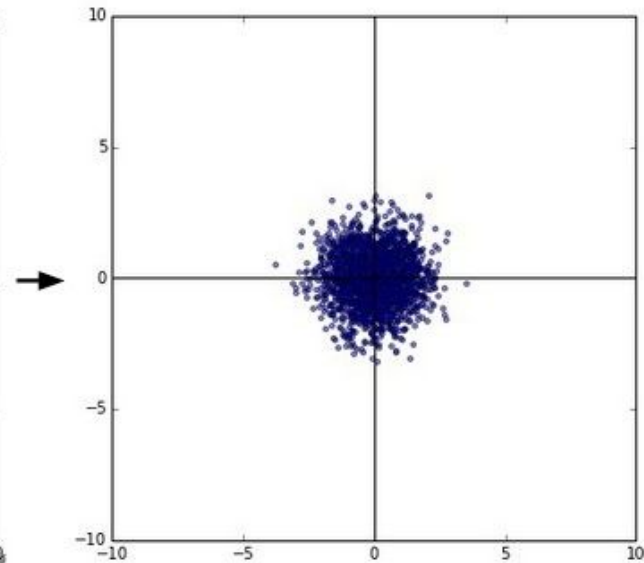
original data



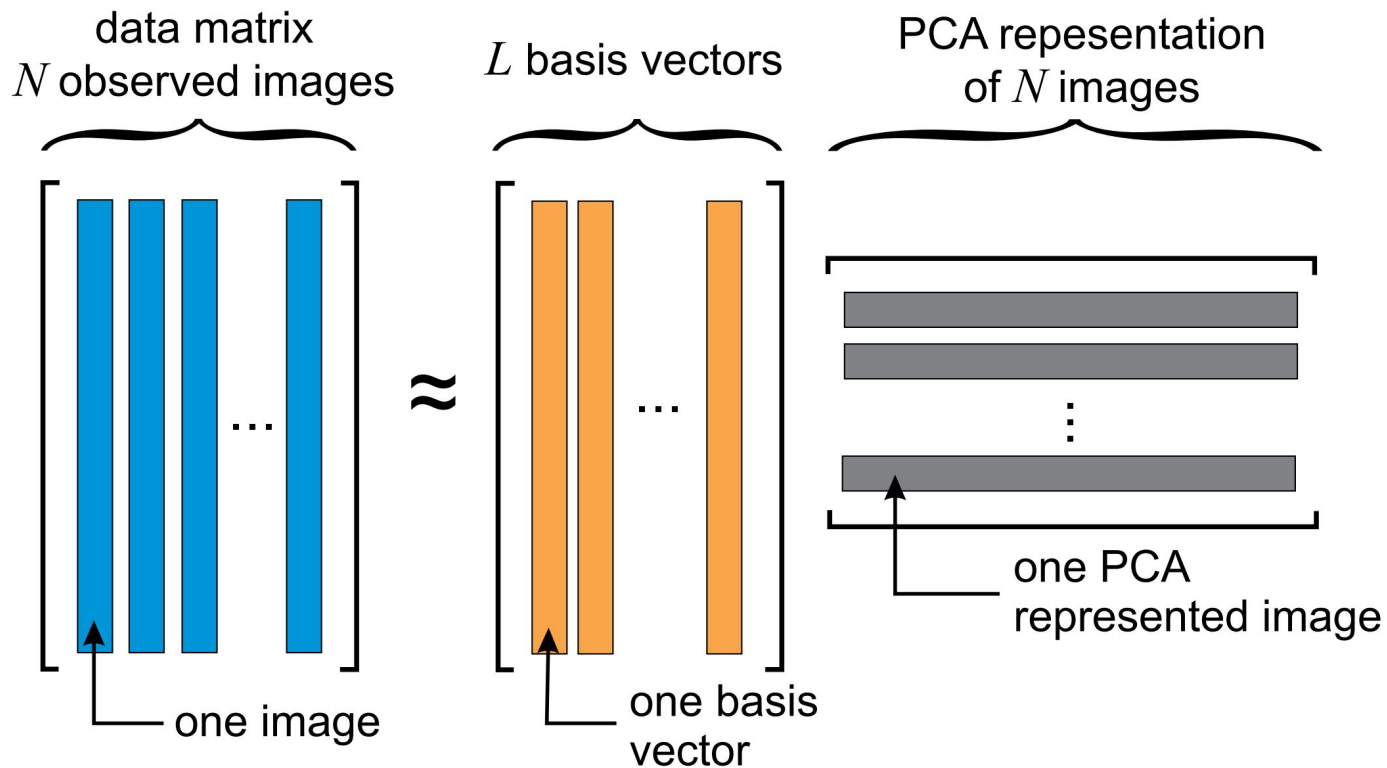
decorrelated data



whitened data



PCA



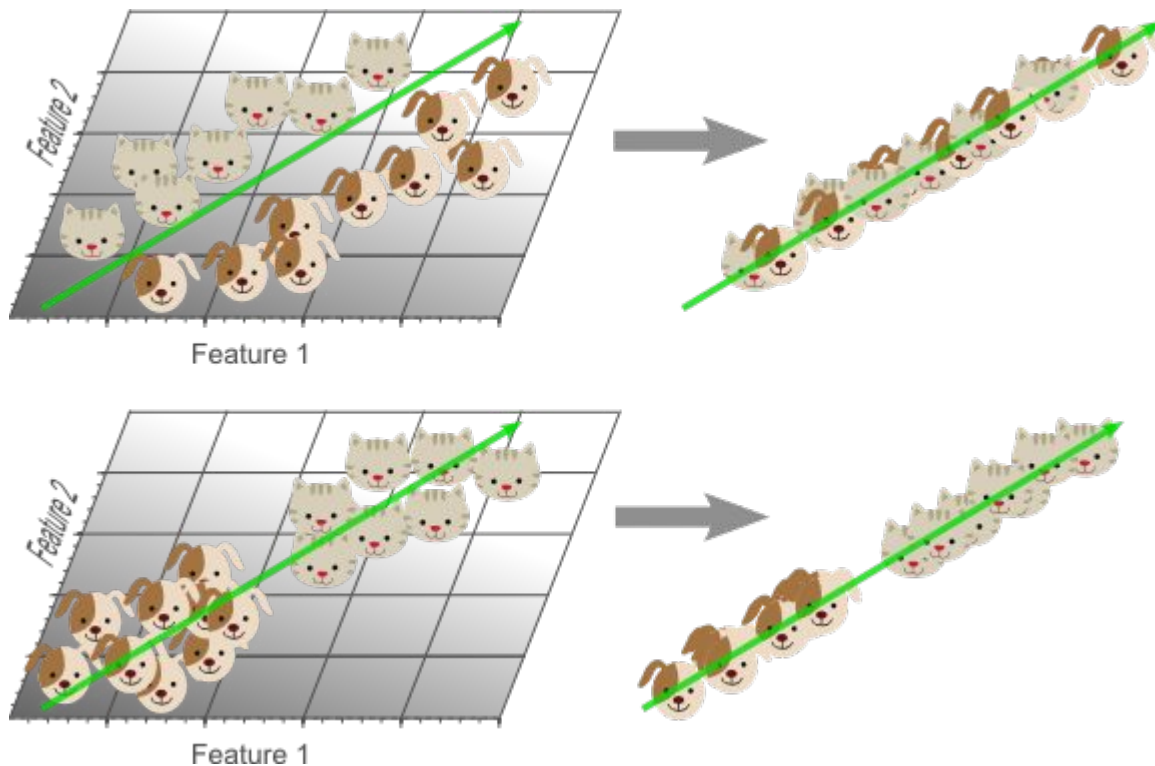
PCA



PCA

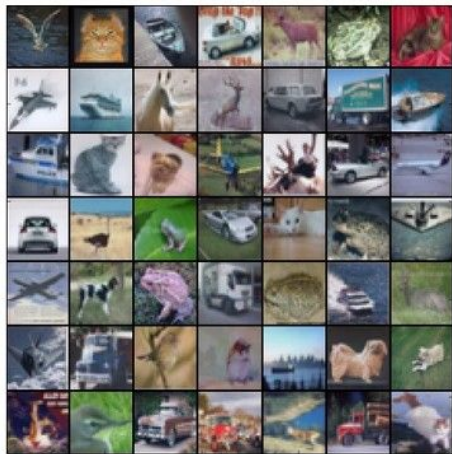

$$\text{Original Face} = q_1 + q_2 + q_3 + q_4$$

Сложности РСА

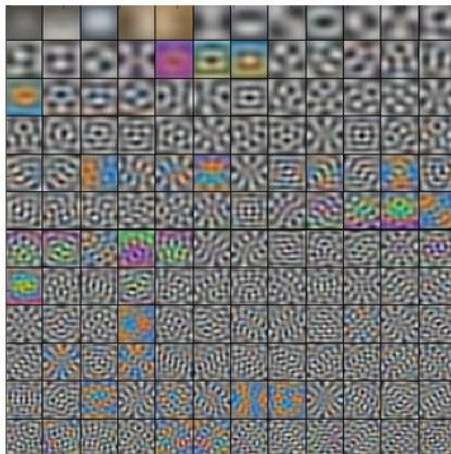


Предобработка данных PCA

original images



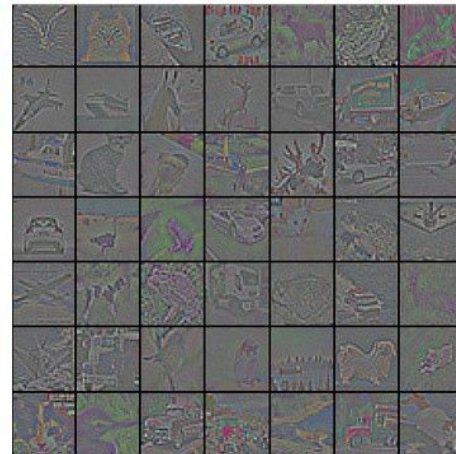
top 144 eigenvectors



reduced images



whitened images



Полезные материалы

- [Generative Adversarial Networks \(GANs\)](#)
- [Introductory guide to Generative Adversarial Networks \(GANs\)](#)
- [Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks](#)
- [Inceptionism: Going Deeper into Neural Networks](#)
- [Breaking Linear Classifiers on ImageNet](#)