

Генеративные модели в машинальном обучении

Лекция 3
Генеративно-состязательные сети

Михаил Гущин

mhushchyn@hse.ru

НИУ ВШЭ, 2024



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

План

- ▶ Приложения
- ▶ Генеративно-состязательные сети (GANы)
- ▶ Проблемы GANов
- ▶ Условный GAN

Приложения



This X Does Not Exist



This Foot Does Not Exist

Note that this is an SMS chatbot. You can text it to get pictures of feet. The pictures are animated. The feet are nonexistent. Why would you want to do this? Who knows.

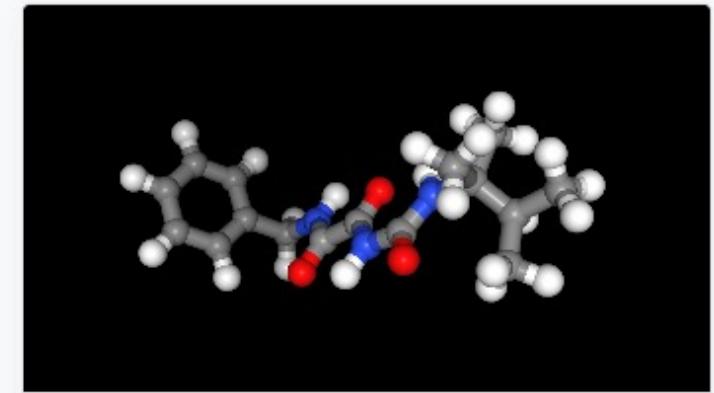
Created by MSCHF.



This Artwork Does Not Exist

Be inspired by minimalism, realism, post-modernism, pre-modernism, modernism, and ancientism (not actually a thing). No matter your art preferences, you can find it here with enough refreshing.

Created by Michael Friesen.



This Chemical Does Not Exist

Who said drug discovery is hard? Just refresh until you find the right chemical. In all seriousness, the fact that this renders a 3D model with the correct bond pairings is impressive.

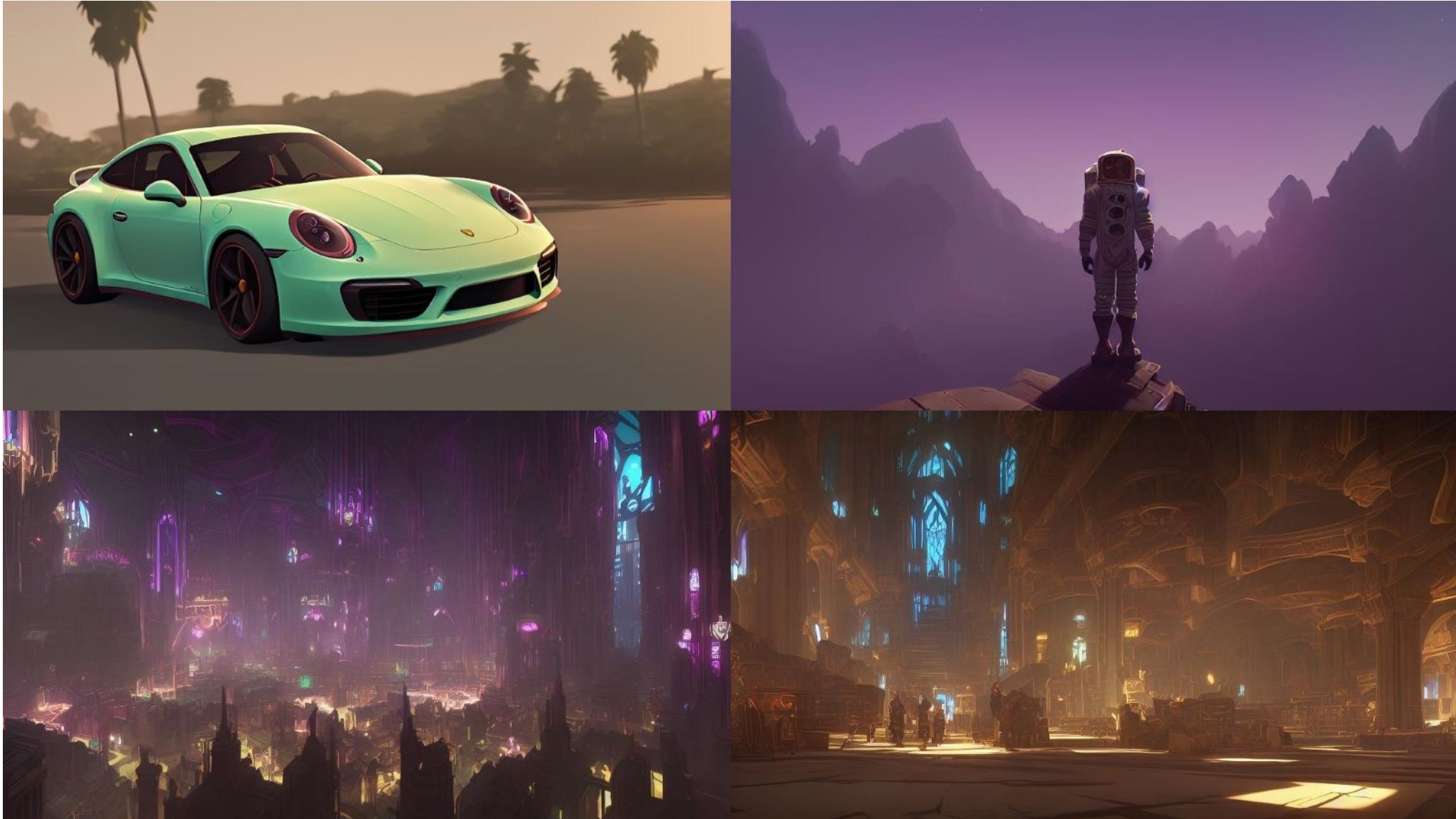
<https://thisxdoesnotexist.com>

Генерация фотографий



Source: Goodfellow et al., 2014; Radford et al., 2016; Liu & Tuzel, 2016;
Karras et al., 2018; Karras et al., 2019; Goodfellow, 2019; Karras et al.,
2020; AI Index, 2021; Vahdat et al., 2021

Stable Diffusion (StabilityAI)



Source: <https://learnopencv.com/image-generation-using-diffusion-models/>

Kandinsky 2.2



Кот фиксит баг в обучении
нейронной сети



Розовый фламинго стоит на
одной ноге в воде

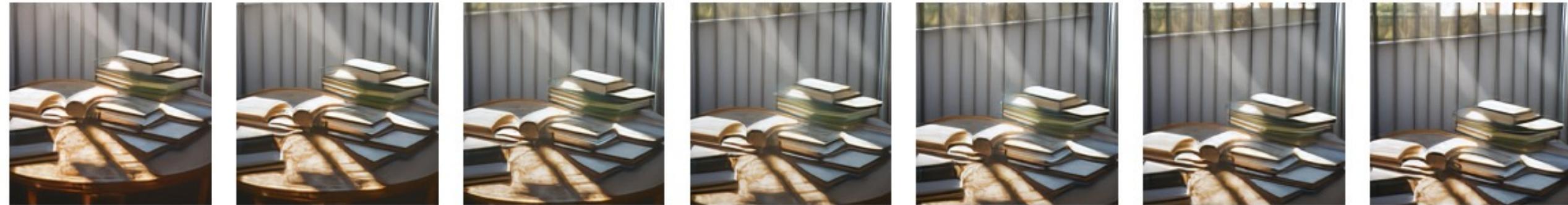
Ссылка: <https://www.sberbank.com/promo/kandinsky>

Text-to-Video

Make-a-Video from Meta



(a) A dog wearing a superhero outfit with red cape flying through the sky.



(b) There is a table by a window with sunlight streaming through illuminating a pile of books.

Source: stateof.ai 2022

Deepfake



John Wick 4: Dangerous Cleanup
7,4 млн просмотров



This is you right now 😂
1,2 млн просмотров



Katana wins
18 млн просмотров



Thank you, my followers ❤️
1,9 млн просмотров



10 minutes of eternity...
1,9 млн просмотров



Keanu Reeves lives with his girlfriend
19 млн просмотров



Did you recognize everyone?
2,1 млн просмотров



When I came back from filming
1,2 млн просмотров



You spin USB right 'round, baby, right 'round... #keanu...
1,5 млн просмотров



Who is the best dancer? 😎
#keanu #reeves #dance #fyp
7,9 млн просмотров



Mirror trick 😜 #keanu #reeves #mirror #tenet
1,1 млн просмотров



Dressing up like a cool guy.
#reeves #keanu #dressup
1,1 млн просмотров

Deepfake



Unreal Keanu Reeves

@unreal_keanu

617 тыс. подписчиков

ГЛАВНАЯ

SHORTS

ПЛЕЙЛИСТЫ

СООБЩЕСТВО

КАНАЛЫ

О КАНАЛЕ

Описание

This is the official deepfake channel of actor Keanu Reeves.
Like, share, subscribe!

Дополнительно

Для коммерческих запросов:

[Показать адрес электронной почты](#)

Страна:

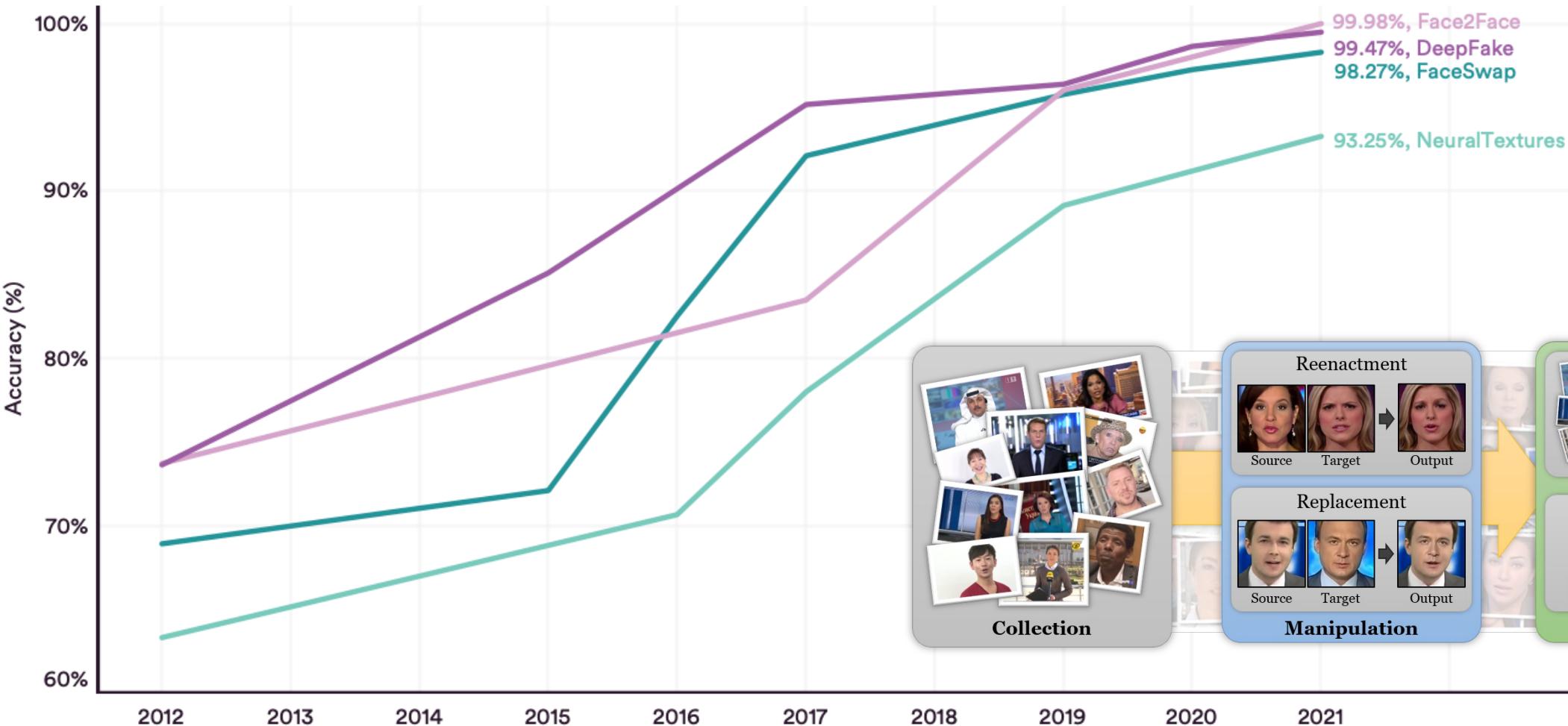
Грузия

Source: https://www.youtube.com/@unreal_keanu/about

Deepfake detection

FACEFORENSICS++: ACCURACY

Source: arXiv, 2021 | Chart: 2022 AI Index Report



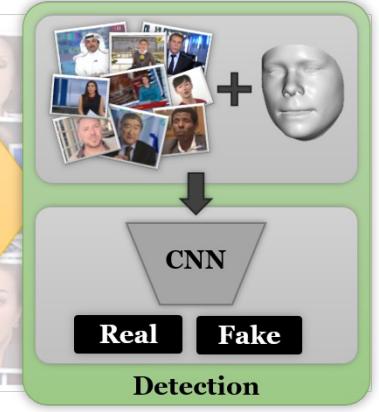
Collection



Manipulation

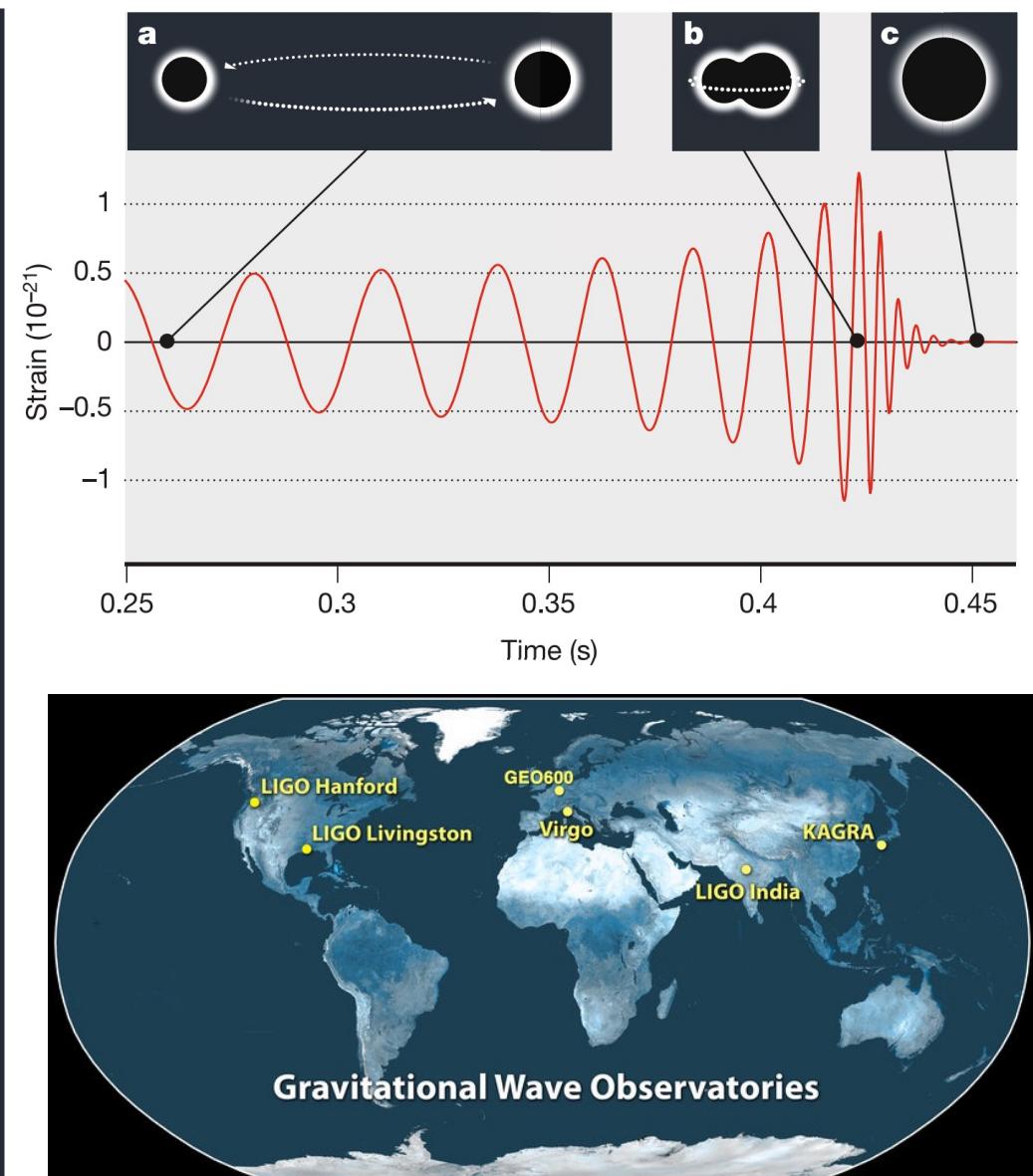
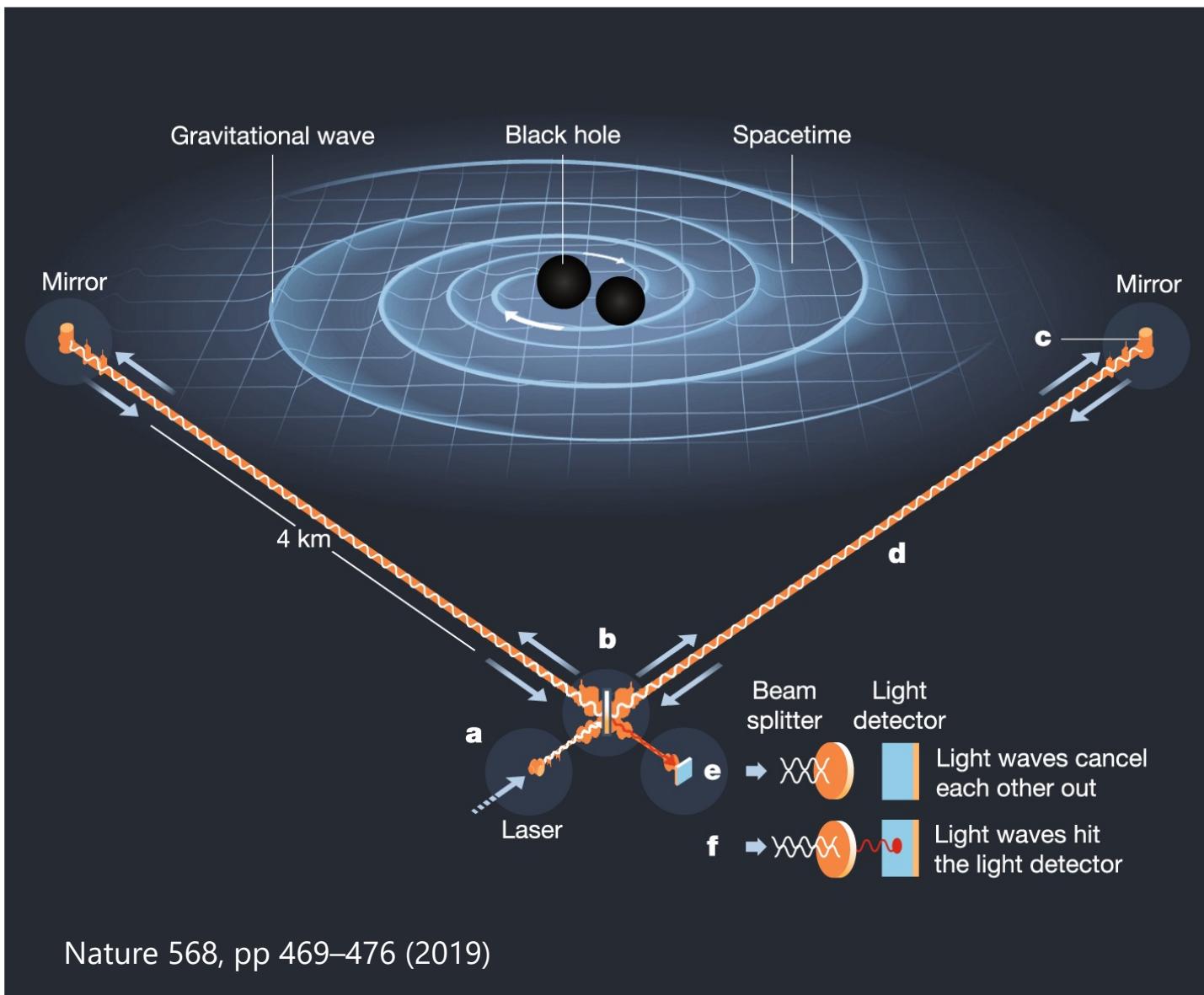
Reenactment
Source Target Output

Replacement
Source Target Output

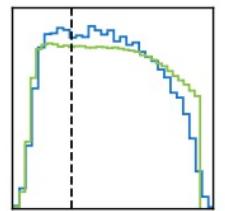


Detection

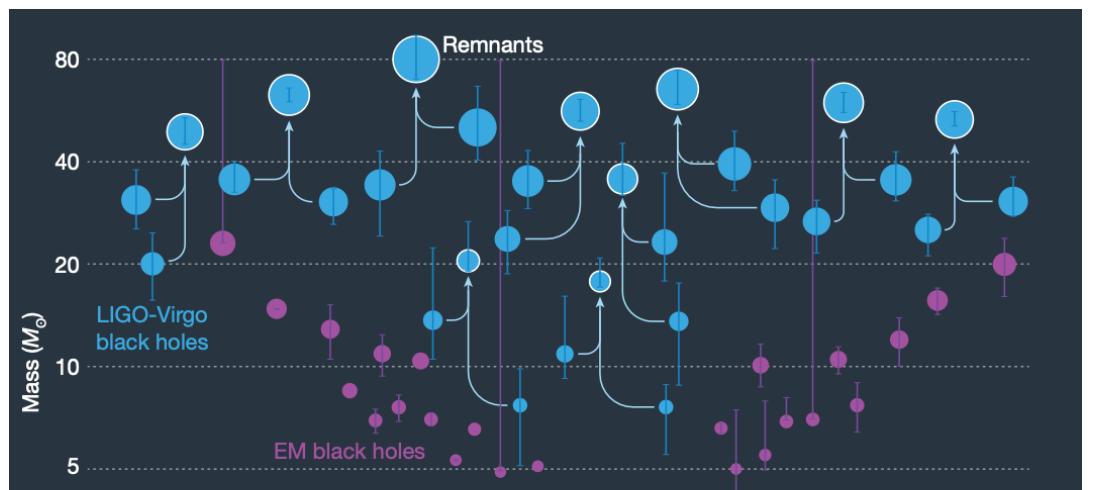
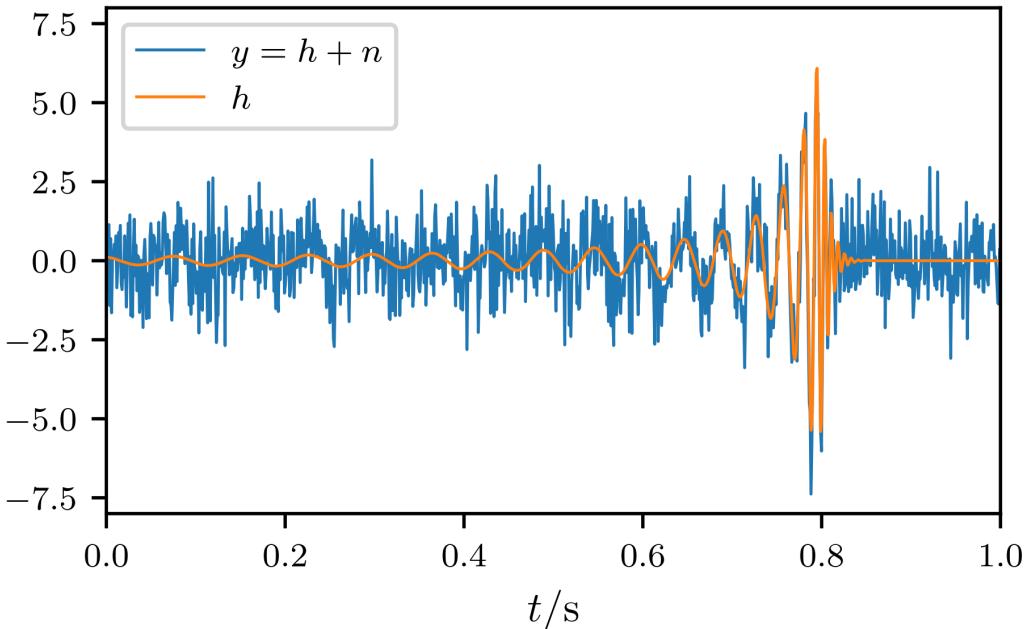
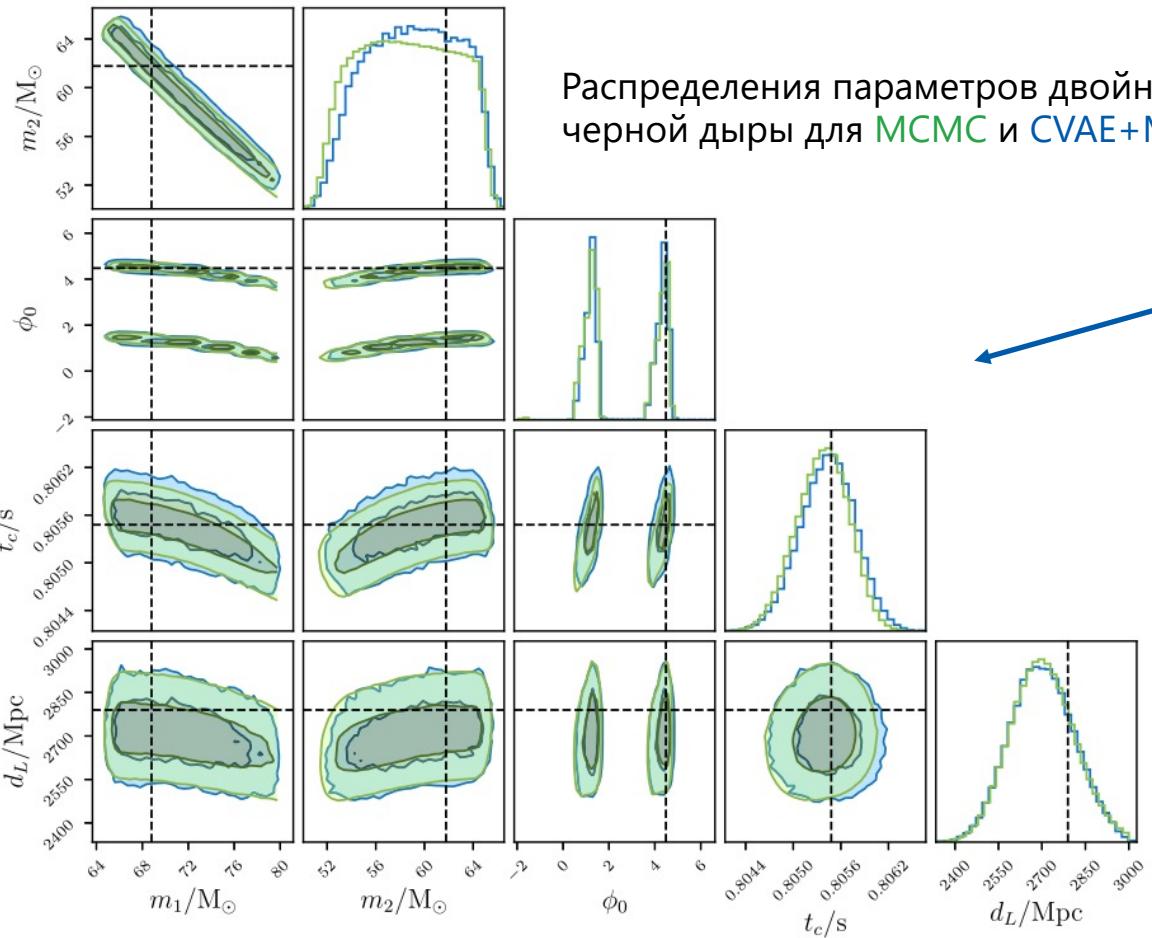
Гравитационные волны



Гравитационные волны



Gravitational-wave parameter estimation with autoregressive neural network flows,
Phys. Rev. D **102**, 104057

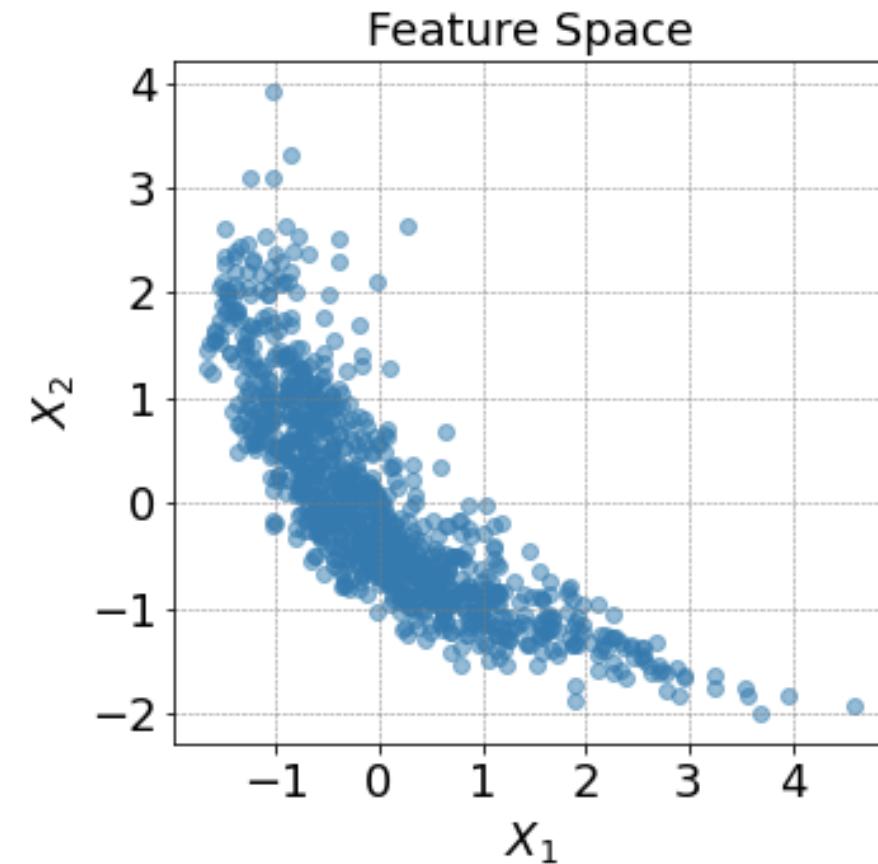


Генеративно-состязательные сети



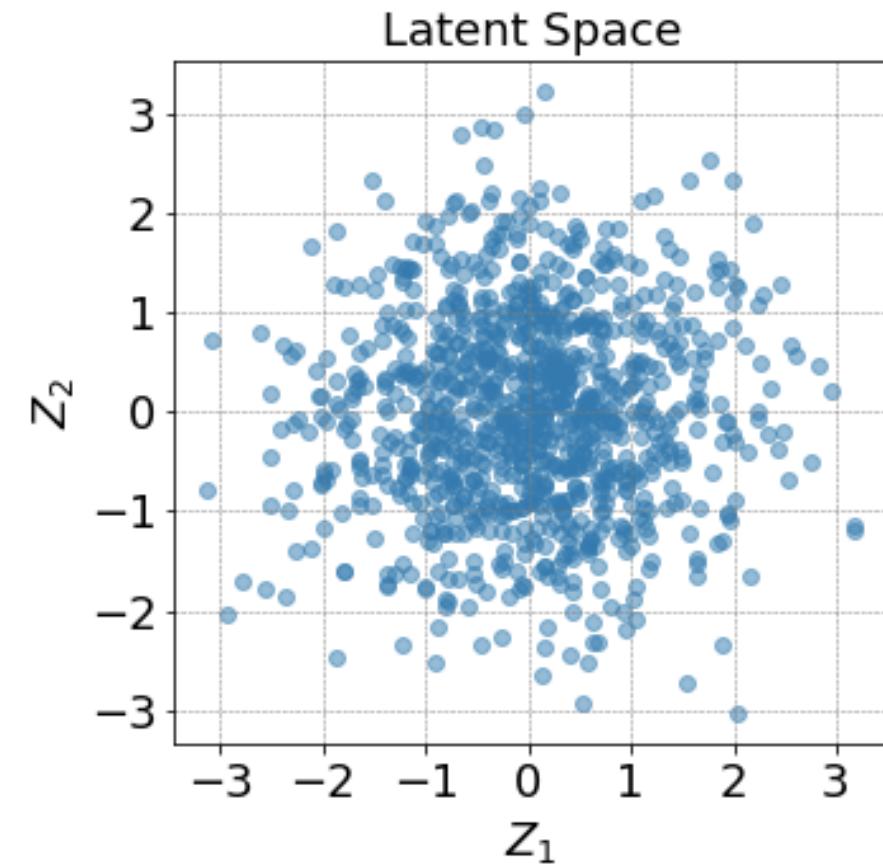
Постановка задачи

- ▶ **Дано:** выборка объектов $x_i \sim p_x(x)$,
где распределение $p_x(x)$ неизвестно.
- ▶ **Задача:** сгенерировать новые объекты
 $\hat{x}_j \sim p_x(x)$ с таким же распределением,
т.е. похожими на x_i .

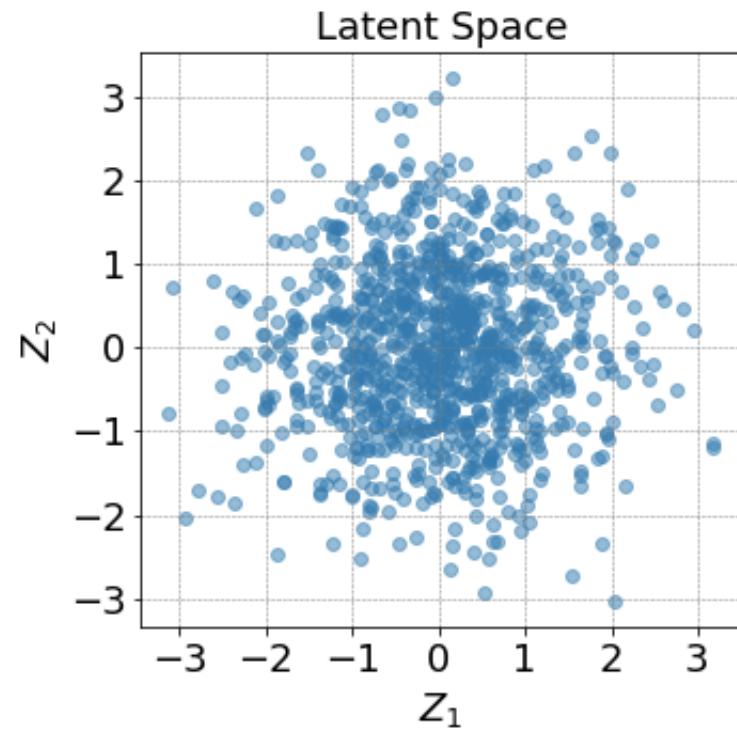


Известные распределения

- ▶ Мы умеем генерировать объекты из известных распределений.
- ▶ Генерируем выборку объектов $\mathbf{z}_i \sim p_z(\mathbf{z})$, где $p_z(z)$ - двумерное нормальное распределение.
- ▶ Пространство \mathbf{Z} называется **скрытым**.



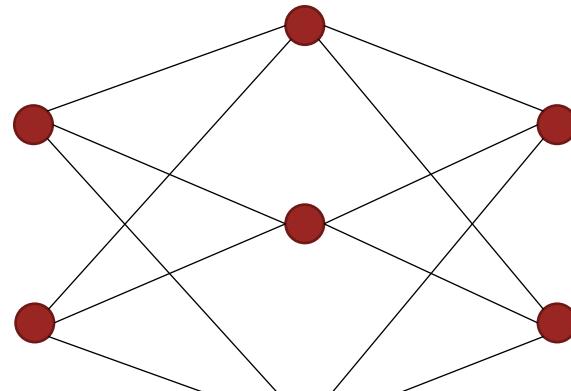
Генератор



Z

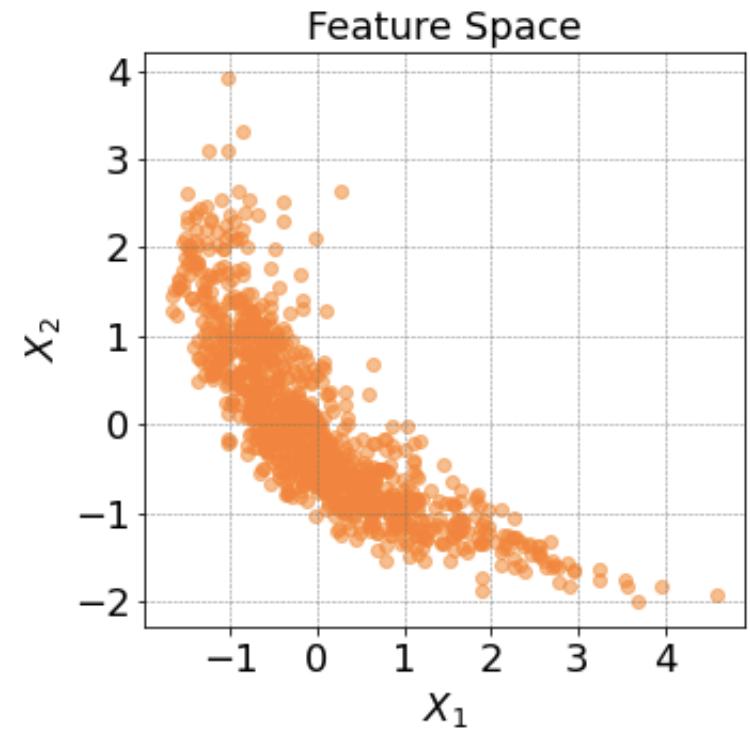
Случайный шум

Нейронная
сеть



$G(Z)$

Генератор



$\hat{X} = G(Z)$

Сгенерированные
объекты

Генератор

- ▶ X – матрица реальных объектов.
- ▶ $\hat{X} = G(Z)$ – матрица сгенерированных объектов.
- ▶ Хотим, чтобы объекты из \hat{X} были похожи на объекты из X .
- ▶ Как посчитать сходство между \hat{X} и X ?
- ▶ Как обучить генератор $G(Z)$?

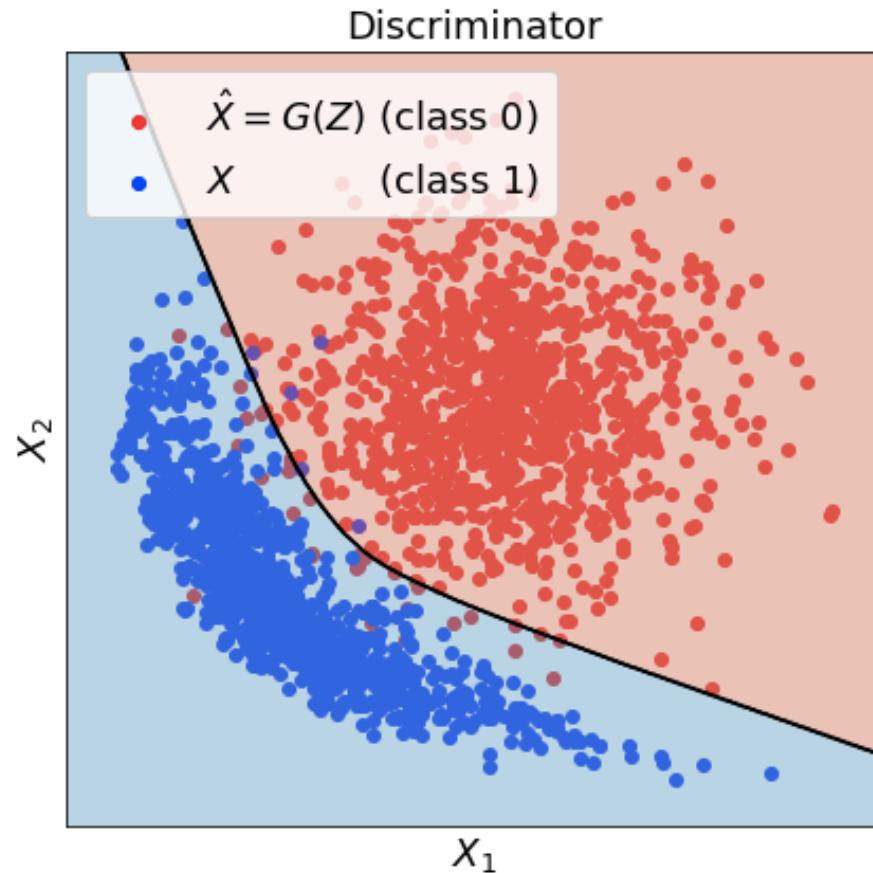
Дискриминатор

Давайте рассмотрим задачу **бинарной классификации**:

- ▶ пусть X – класс 1,
- ▶ пусть \hat{X} – класс 0.

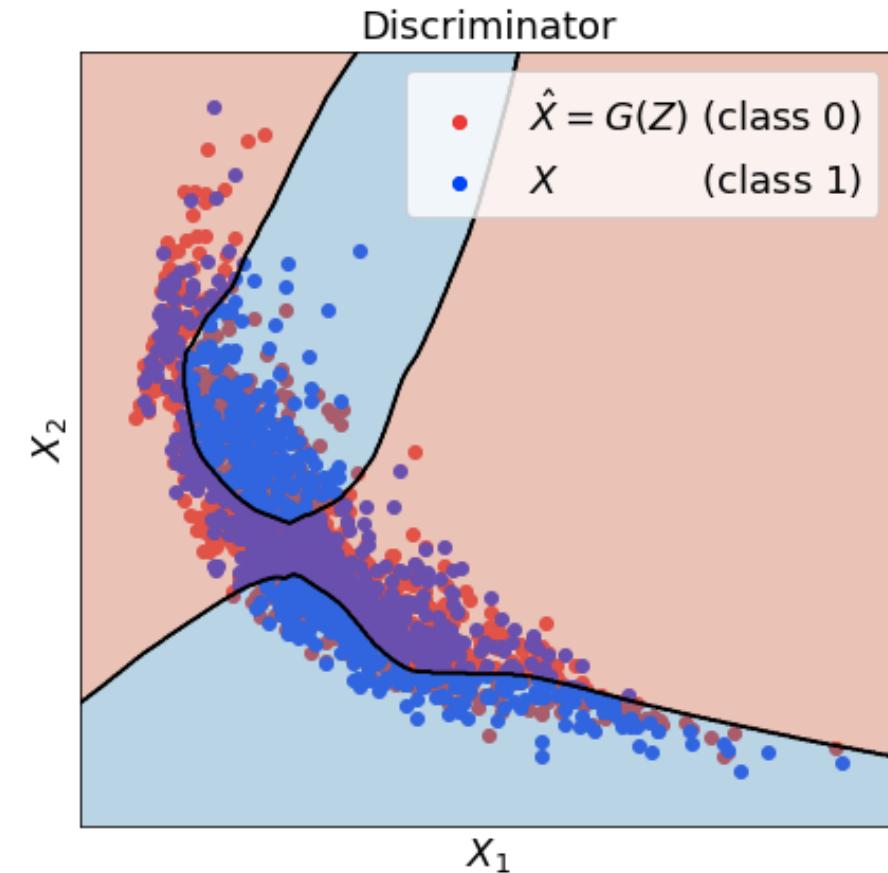
Задача классификатора – разделить два класса. Будем называть его **дискриминатором**.

Дискриминатор



ROC AUC = 0.995

Log Loss = 0.135



ROC AUC = 0.554

Log Loss = 0.689

Дискриминатор

В классификации используется **функция потерь L** :

$$L = -\frac{1}{n} \sum_{i=1}^n y_i \log D(x_i) + (1 - y_i) \log(1 - D(x_i))$$

где $D(x_i)$ – выход дискриминатора, $y_i \in \{0, 1\}$ – метка объекта.

Будем использовать значение этой функции как меру сходства объектов двух классов.

Дискриминатор

Перепишем функцию потерь:

$$L = -\frac{1}{n} \sum_{i:y_i=1} \log D(x_i) - \frac{1}{n} \sum_{i:y_i=0} \log(1 - D(x_i))$$

Здесь каждая сумма берется по своему классу.

Дискриминатор

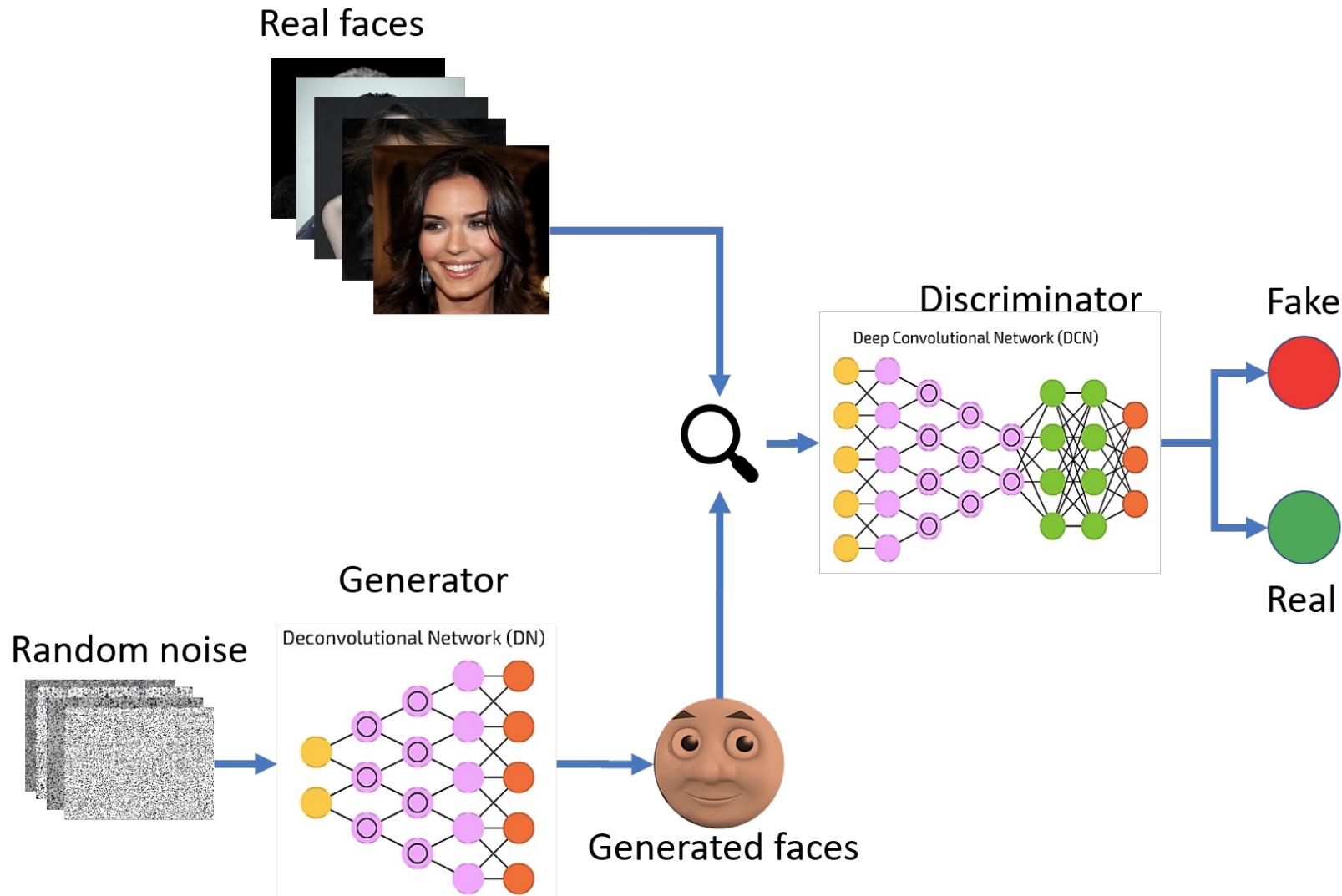
- ▶ Мы договорились, что X – класс 1, а $\hat{X} = G(Z)$ – класс 0.
- ▶ Тогда функция потерь имеет вид:

$$L(\mathbf{D}, \mathbf{G}) = -\frac{1}{n} \sum_{x_i \in X} \log \mathbf{D}(x_i) - \frac{1}{n} \sum_{z_i \in Z} \log(1 - \mathbf{D}(\mathbf{G}(z_i)))$$

- ▶ Обучение дискриминатора и генератора:

$$\max_{\mathbf{G}} \min_{\mathbf{D}} L(\mathbf{D}, \mathbf{G})$$

Алгоритм



Link: <https://www.spindox.it/en/blog/generative-adversarial-neural-networks/>

Алгоритм

- ▶ k раз делаем:
 - Генерируем шум $\{z_1, z_2, \dots, z_m\}$
 - Берем m реальных объектов $\{x_1, x_2, \dots, x_m\}$
 - Обновляем веса дискриминатора W_D :

$$W_D = W_D - \alpha \nabla L(\mathbf{D}, \mathbf{G})$$

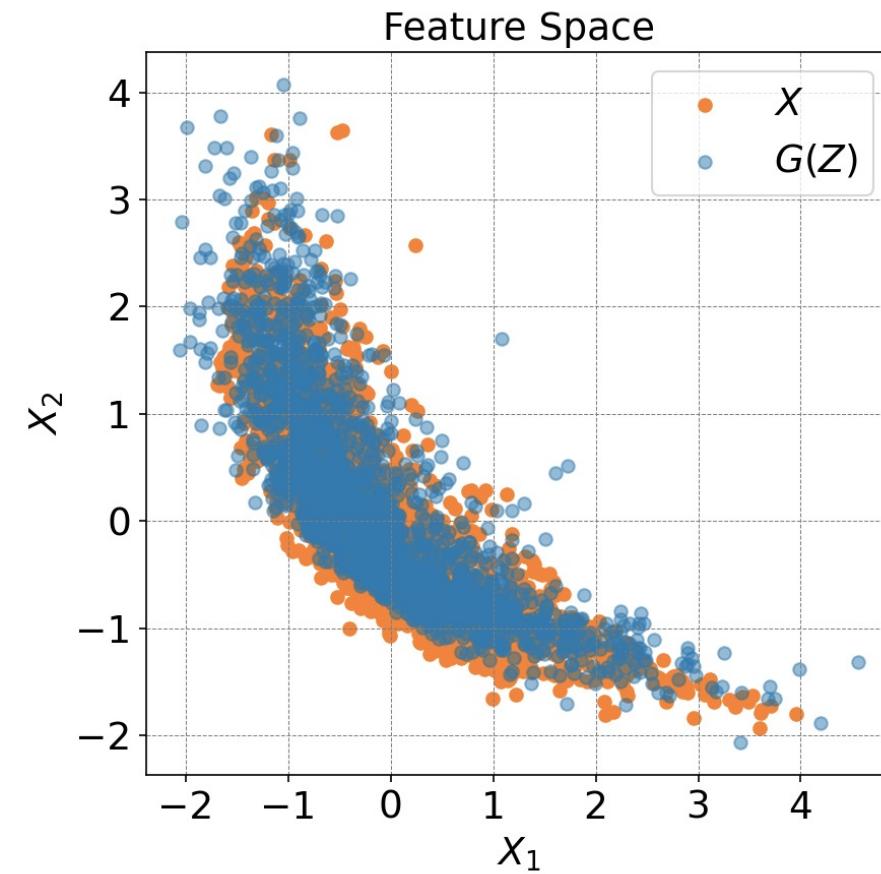
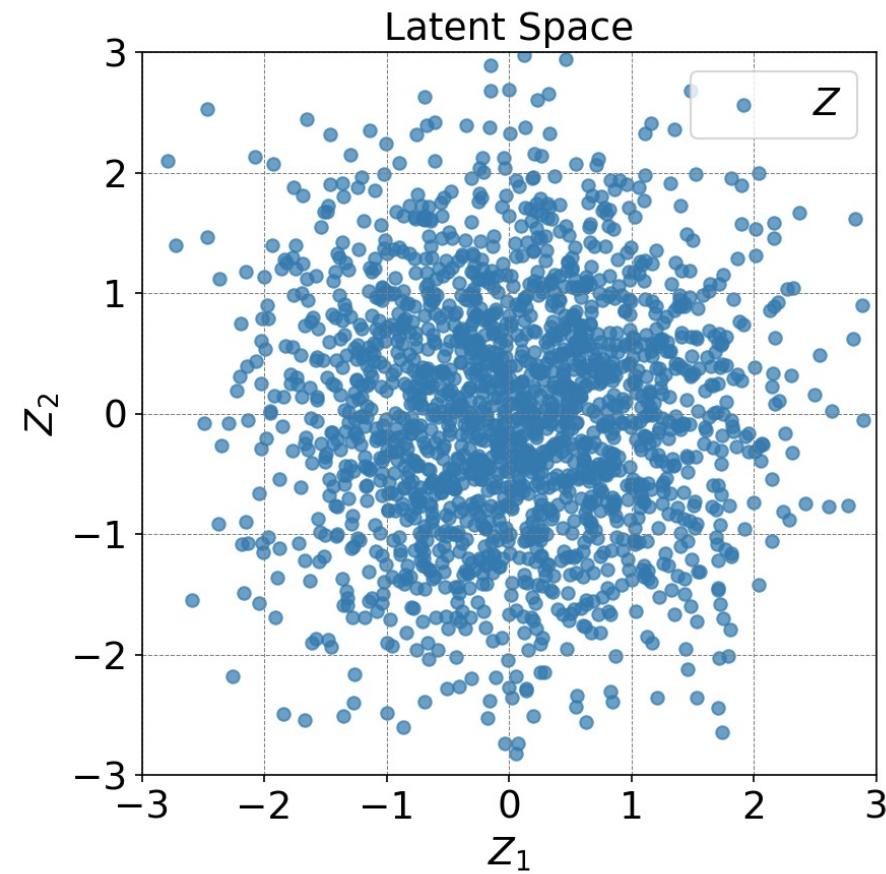
- ▶ Генерируем шум $\{z'_1, z'_2, \dots, z'_m\}$
- ▶ Обновляем веса генератора W_G :

$$W_G = W_G + \alpha \nabla L(\mathbf{D}, \mathbf{G})$$

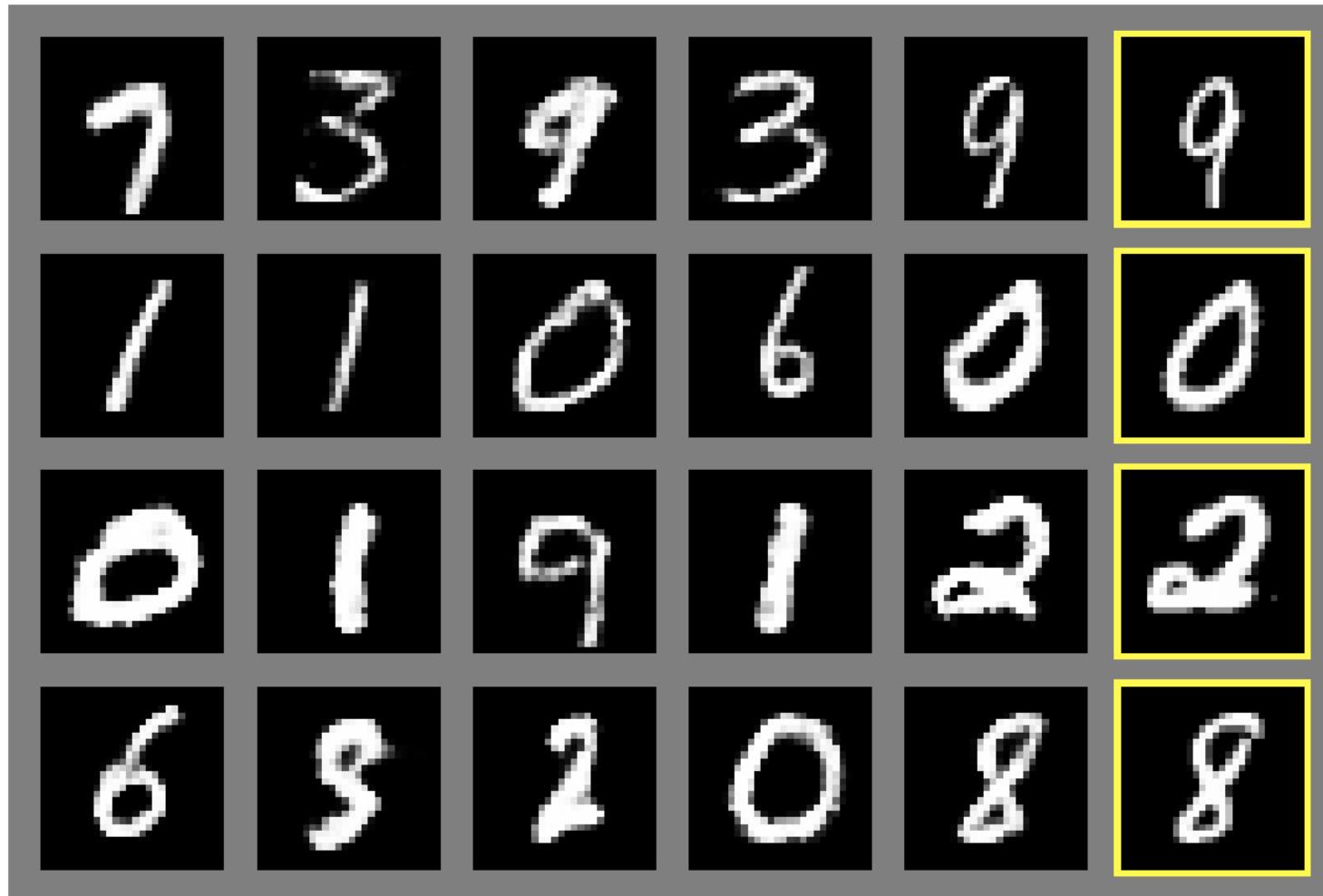
- ▶ Повторить шаги



Пример



Пример



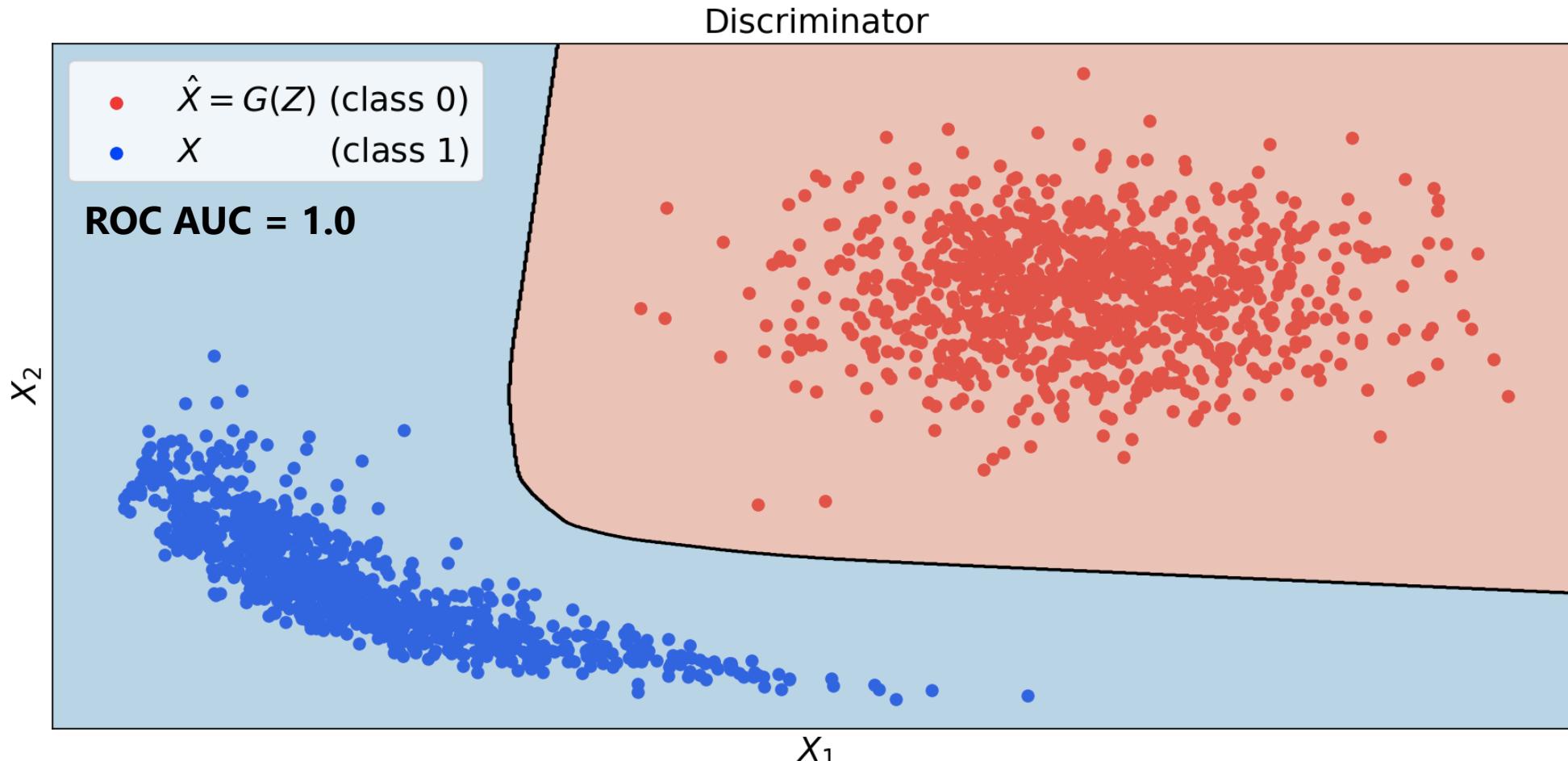
Пример



Проблемы GANов

Затухание градиентов

Какая-то итерация обучения
GAN:



Затухание градиентов

- ▶ Классы идеально разделяются дискриминатором =>
 - ▶ $\mathbf{D}(x_i) = \mathbf{1}$ для любого x_i в X ;
 - ▶ $\mathbf{D}(\hat{x}_j) = \mathbf{D}(\mathbf{G}(z_j)) = \mathbf{0}$ для любого z_j в Z .
-
- ▶ Подставим эти значения в функцию потерь:

$$L(\mathbf{D}, \mathbf{G}) = -\frac{1}{n} \sum_{x_i \in X} \log \mathbf{D}(x_i) - \frac{1}{n} \sum_{z_i \in Z} \log(1 - \mathbf{D}(\mathbf{G}(z_i)))$$

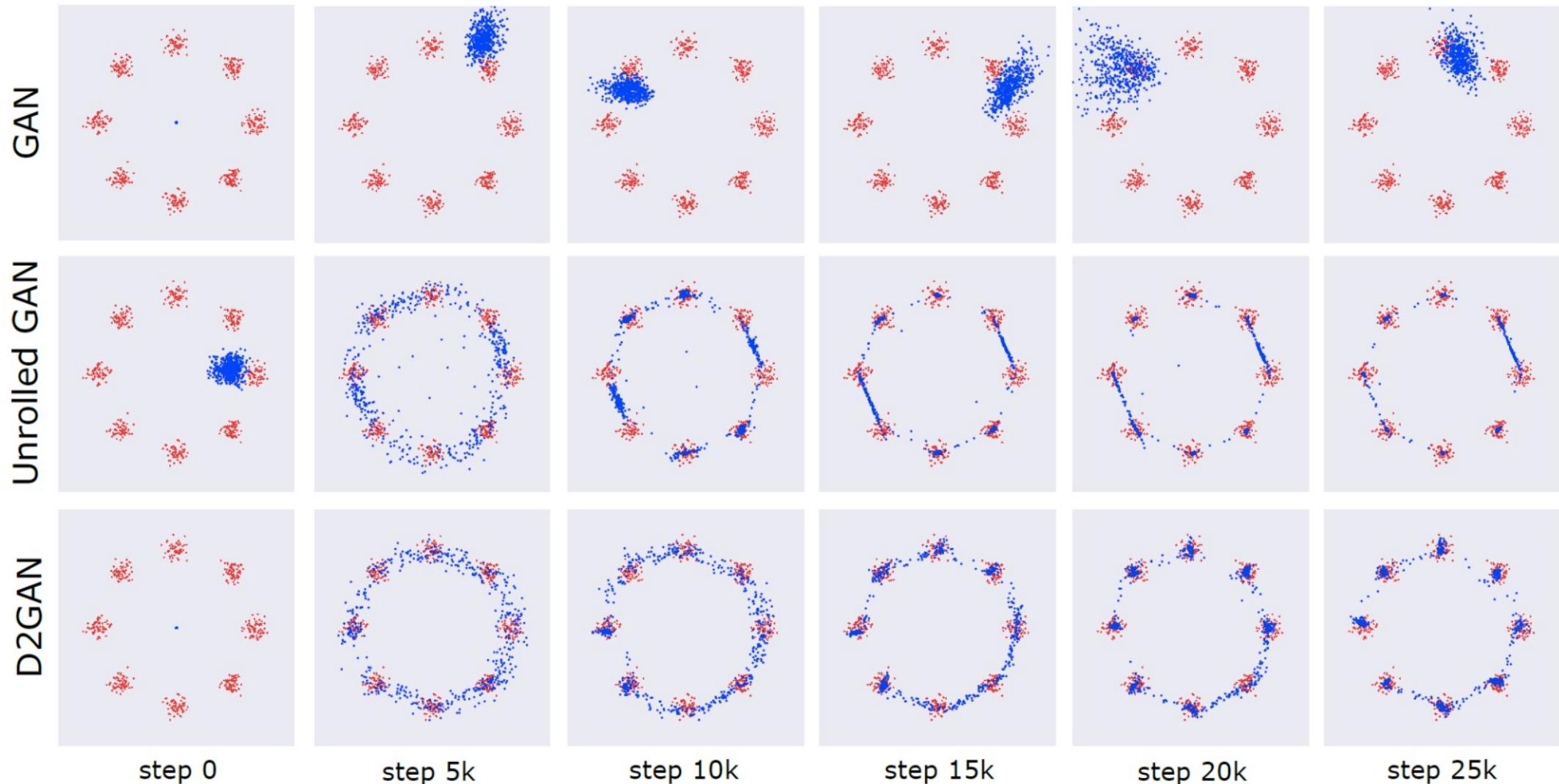
Затухание градиентов

- ▶ В результате получаем:

$$L(\mathbf{D}, \mathbf{G}) = -\frac{1}{n} \sum_{x_i \in X} \log \mathbf{1} - \frac{1}{n} \sum_{z_i \in Z} \log(1 - \mathbf{0}) = 0 = \text{const}$$

- ▶ $\nabla L(\mathbf{D}, \mathbf{G}) = 0 \Rightarrow$
- ▶ Дискриминатор и генератор больше **не обучаются!**

Схлопывание мод



<https://arxiv.org/abs/1709.03831>

Пример

1	1	1	0	0	0	0	0
0	0	1	1	0	0	0	1
1	1	1	0	1	0	0	1
0	0	1	1	1	1	0	1
0	0	1	1	1	1	1	0
0	0	0	0	0	0	1	1
0	1	0	1	0	0	0	1
1	0	1	1	1	1	0	1

Схлопывание мод

Причины:

- ▶ $D(x_i) = 1$ для x_i из некоторых мод =>
- ▶ $\nabla L(D, G) = 0$ в окрестности этих мод =>
- ▶ GAN не учится генерировать эти моды;
- ▶ Что-то еще, что все плохо понимают ☺

Условный GAN



Задача

Input

A yellow bird with brown wings and a pointed bill



This bird is blue and black in color, with a sharp beak and small eyes



A white bird with white feathers and gray wings



This small bird with a red belly, a pointed bill and red wings



StackGAN++



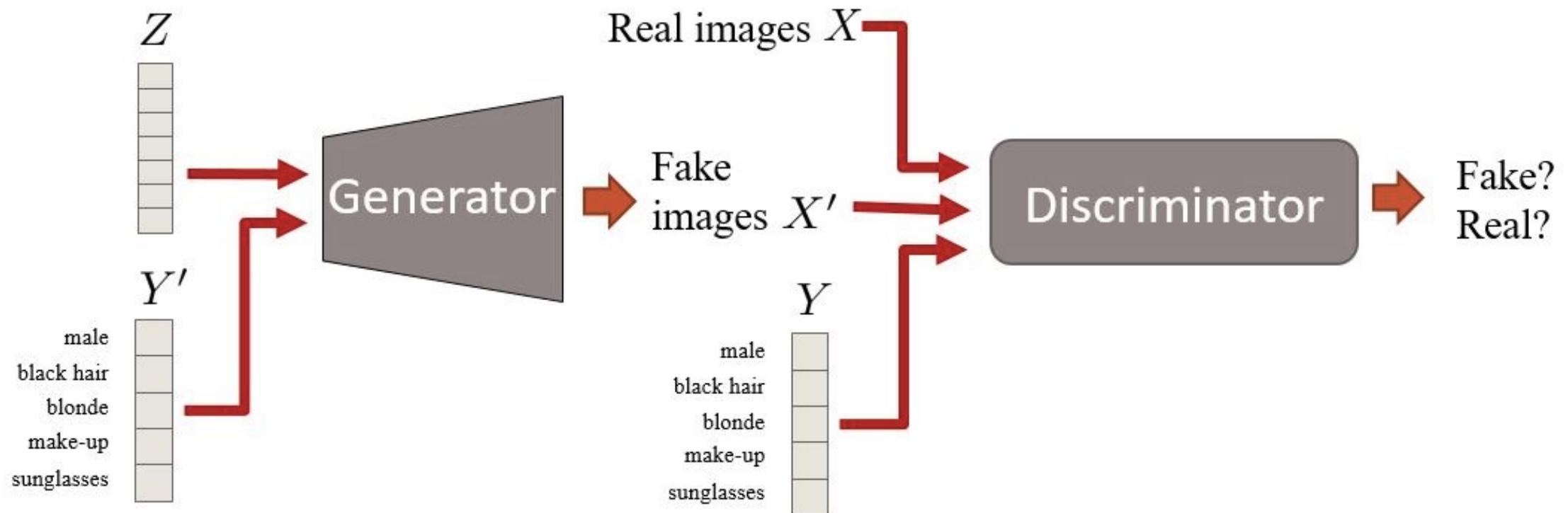
AttnGAN



ACGAN
(256 × 256)



Conditional GAN



<https://guimperarnau.com/blog/2017/03/Fantastic-GANs-and-where-to-find-them>

ФУНКЦИЯ ПОТЕРЬ

$$L(\mathbf{D}, \mathbf{G}) = -\frac{1}{n} \sum_{x_i \in X} \log \mathbf{D}(x_i, \mathbf{c}_i) - \frac{1}{n} \sum_{z_i \in Z} \log(1 - \mathbf{D}(\mathbf{G}(z_i, \mathbf{c}_i)))$$

где \mathbf{c}_i - вектор некоторого условия

Conditional GAN

Условием Y может являться:

- ▶ Метка класса
 - «кот», «собака», «машина» и т.д.
- ▶ Дополнительные признаки
 - возраст, рост, освещенность и т.д.
- ▶ Изображение
 - эскиз рисунка
- ▶ Текст
 - описание изображения для генерации

Пример



<https://arxiv.org/abs/1611.0700>

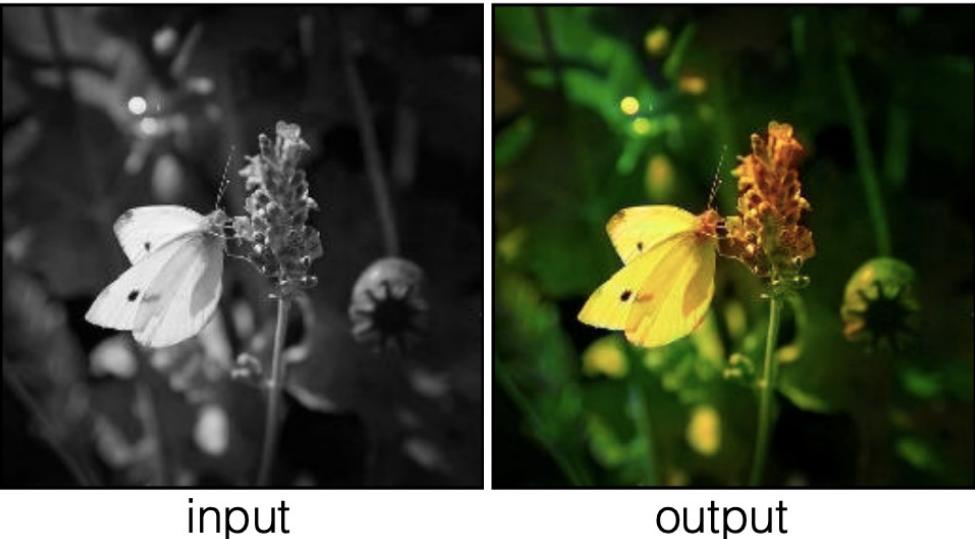
Пример

<https://proceedings.neurips.cc/paper/2017/file/dc6a6489640ca02b0d42dabeb8e46bb7-Paper.pdf>



Пример

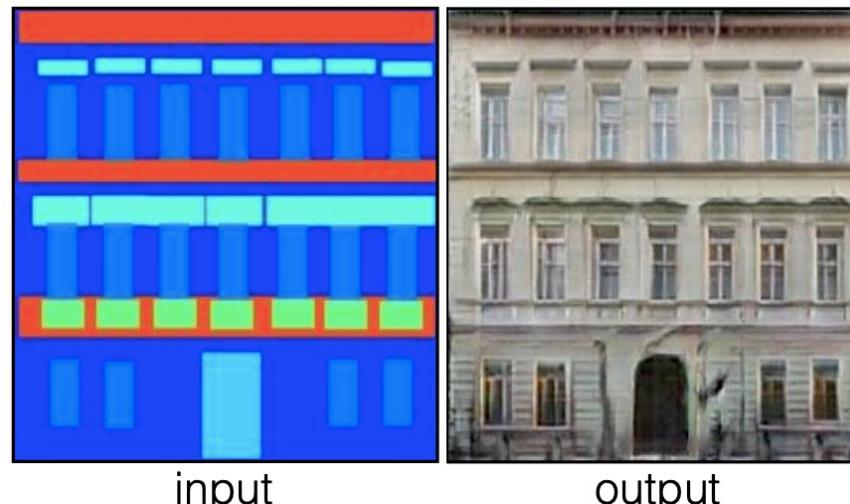
BW to Color



Aerial to Map



Labels to Facade



Другие модели

DCGAN	SGAN	AffGAN	RTT-GAN	SL-GAN
WGAN	SimGAN	TP-GAN	GANCS	Context-RNN-GAN
CGAN	VGAN	IcGAN	SSL-GAN	SketchGAN
LAPGAN	iGAN	ID-CGAN	MAD-GAN	GoGAN
SRGAN	3D-GAN	AnoGAN	PrGAN	RWGAN
CycleGAN	CoGAN	LS-GAN	AL-CGAN	MPM-GAN
WGAN-GP	CatGAN	Triple-GAN	ORGAN	MV-BiGAN
EBGAN	MGAN	TGAN	SD-GAN	
VAE-GAN	S ² GAN	BS-GAN	MedGAN	
BiGAN	LSGAN	MaIGAN	SGAN	