

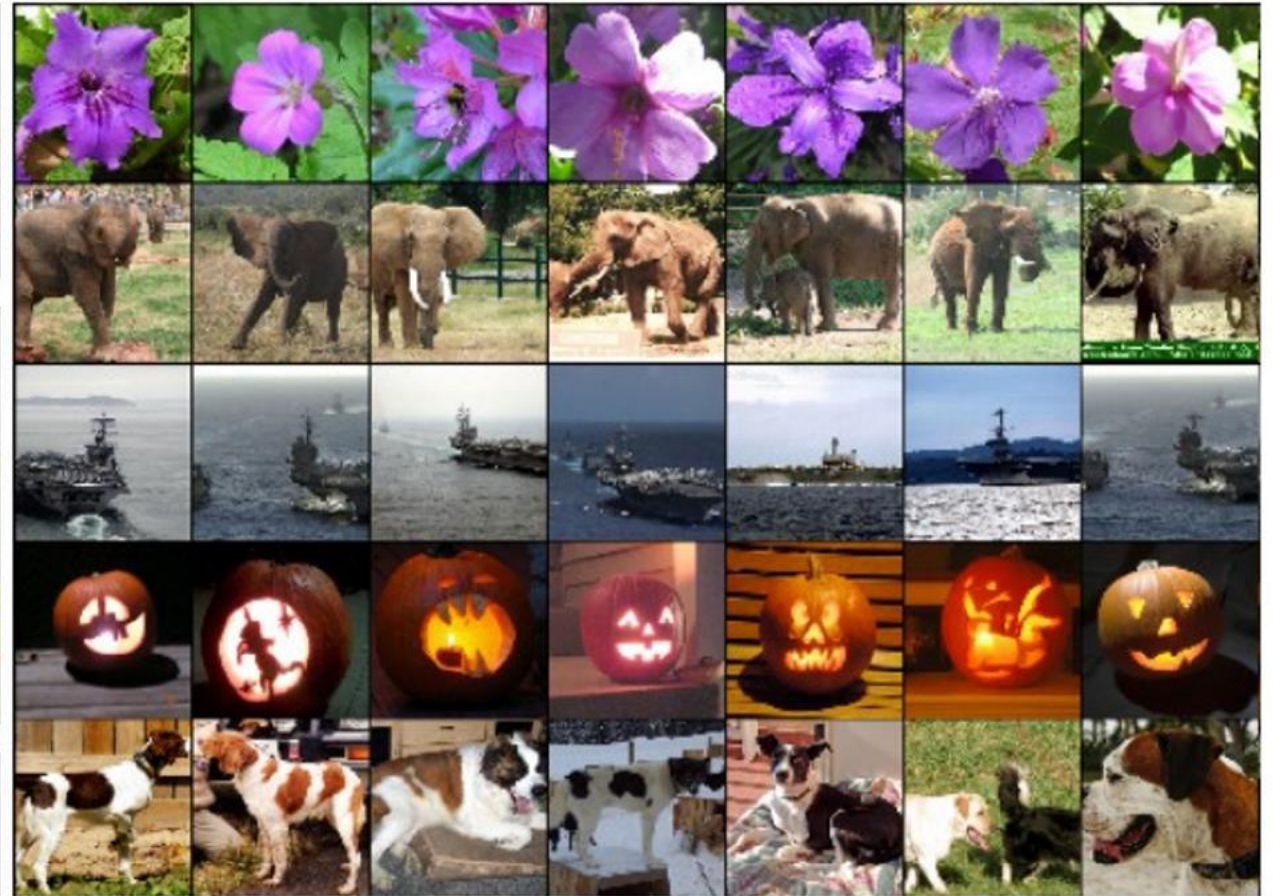
Технологии искусственного интеллекта

Машинное обучение

С 2012 по сейчас: глубокое обучение везде

Классификация изображений

Поиск изображений



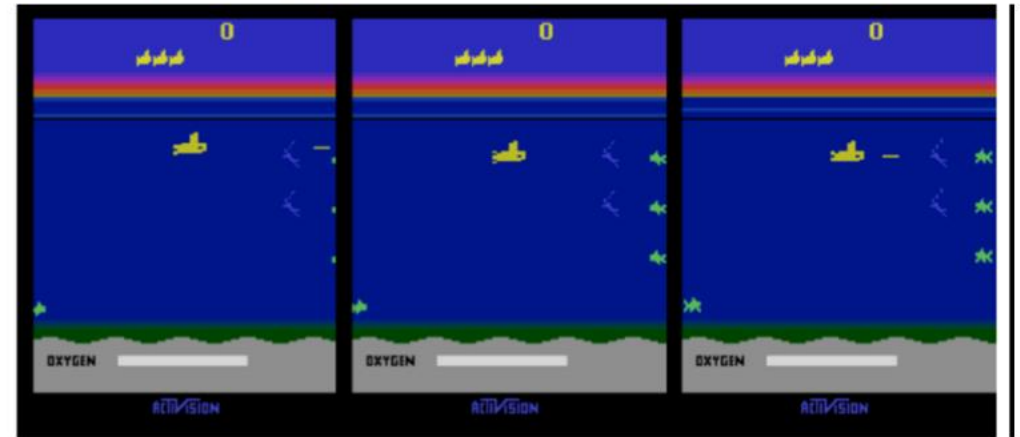
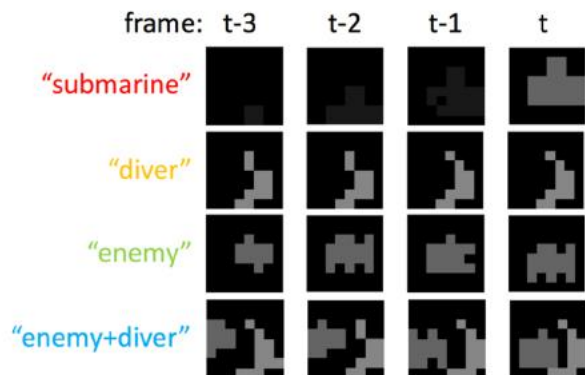
Figures copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

С 2012 по сейчас: глубокое обучение везде

Распознавание позы (Toshev and Szegedy, 2014)

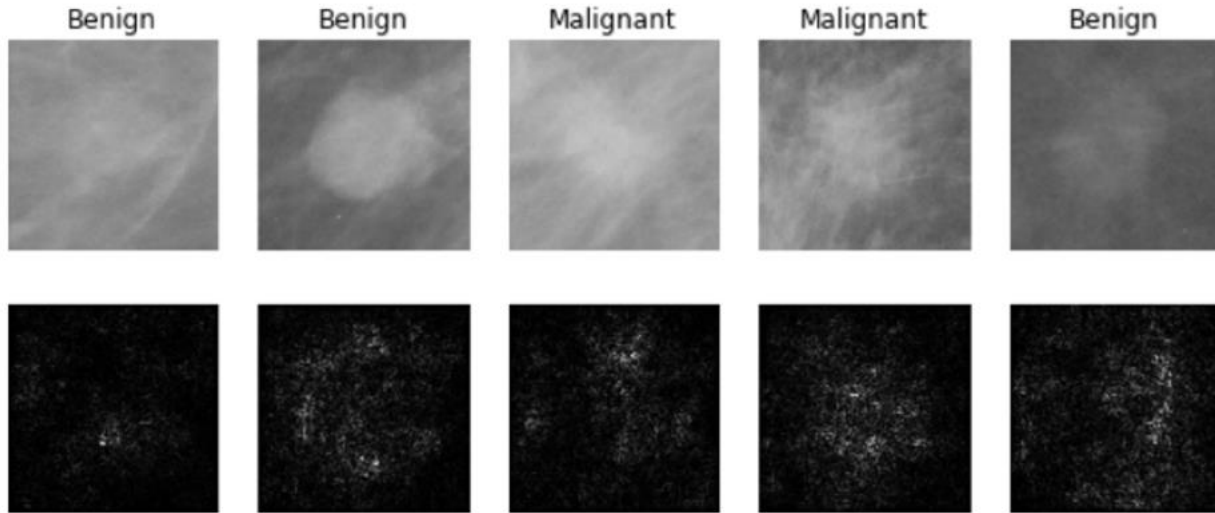


Обучение Atari-играм (Guo et al, 2014)



С 2012 по сейчас: глубокое обучение везде

Распознавание в медицине



Levy et al, 2016 Figure reproduced with permission

Классификация галактик



Dieleman et al, 2014

From left to right: [public domain by NASA](#), [usage permitted by ESA/Hubble](#), [public domain by NASA](#), and [public domain](#).

Распознавание китов



[Kaggle Challenge](#)

[This image](#) by Christin Khan is in the public domain and originally came from the U.S. NOAA.

С 2012 по сейчас: глубокое обучение везде

Аннотирование изображений

Vinyals et al, 2015
Karpathy and Fei-Fei, 2015



A white teddy bear sitting in the grass



A man in a baseball uniform throwing a ball



A woman is holding a cat in her hand



A man riding a wave on top of a surfboard



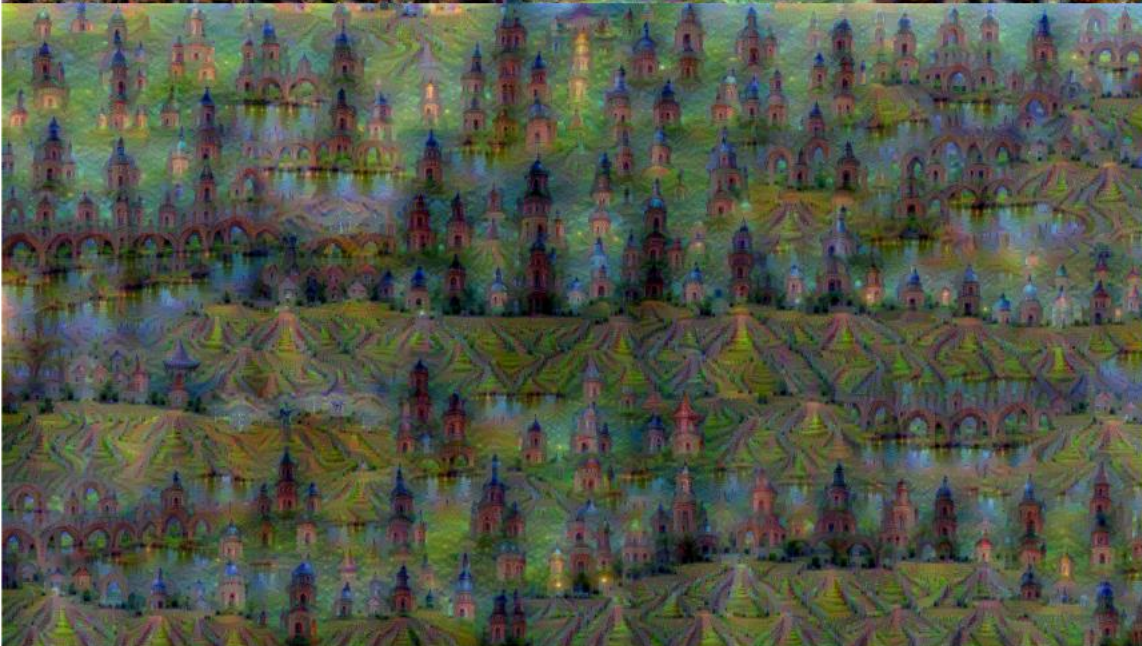
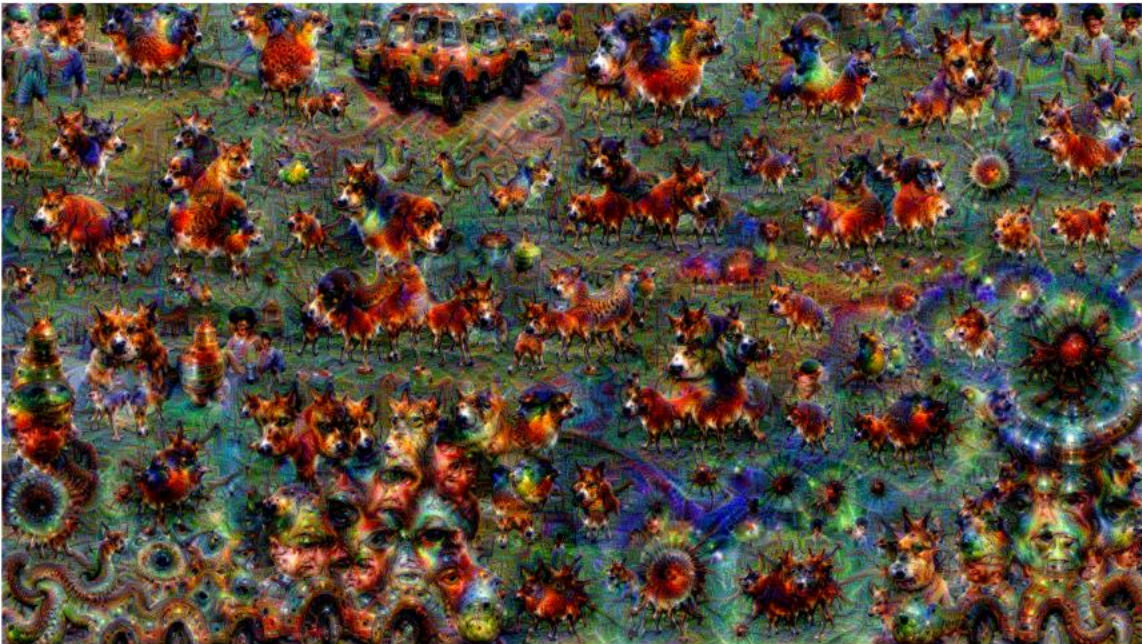
A cat sitting on a suitcase on the floor



A woman standing on a beach holding a surfboard

All images are CC0 Public domain:
<https://pixabay.com/en/luggage-antique-cat-1643010/>
<https://pixabay.com/en/teddy-plush-bears-cute-teddy-bear-1623436/>
<https://pixabay.com/en/surf-wave-summer-sport-litoral-1668716/>
<https://pixabay.com/en/woman-female-model-portrait-adult-983967/>
<https://pixabay.com/en/handstand-lake-meditation-496008/>
<https://pixabay.com/en/baseball-player-shortstop-infield-1045263/>

Captions generated by Justin Johnson using [NeuralTalk2](#).



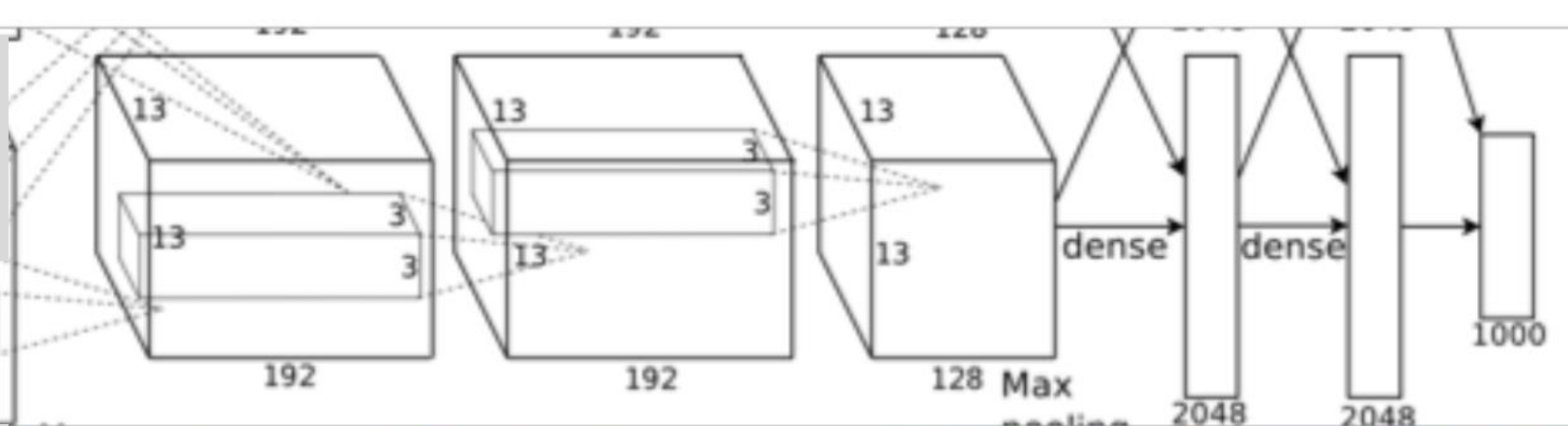
Original image is CC0 public domain
 Starry Night and Tree Roots by Van Gogh are in the public domain
 Bokeh image is in the public domain
 Stylized images copyright Justin Johnson, 2017;
 reproduced with permission



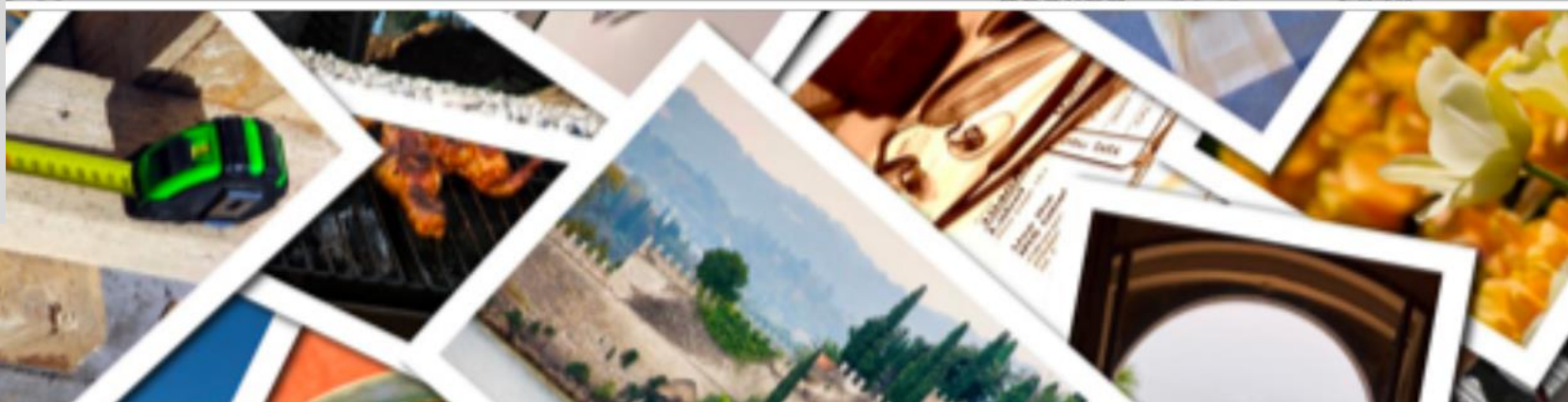
Mordvinsev et al, 2015
 Gatys et al, 2016

Figures copyright Justin Johnson, 2015. Reproduced with permission. Generated using the Inceptionism approach from a [blog post](#) by Google Research.

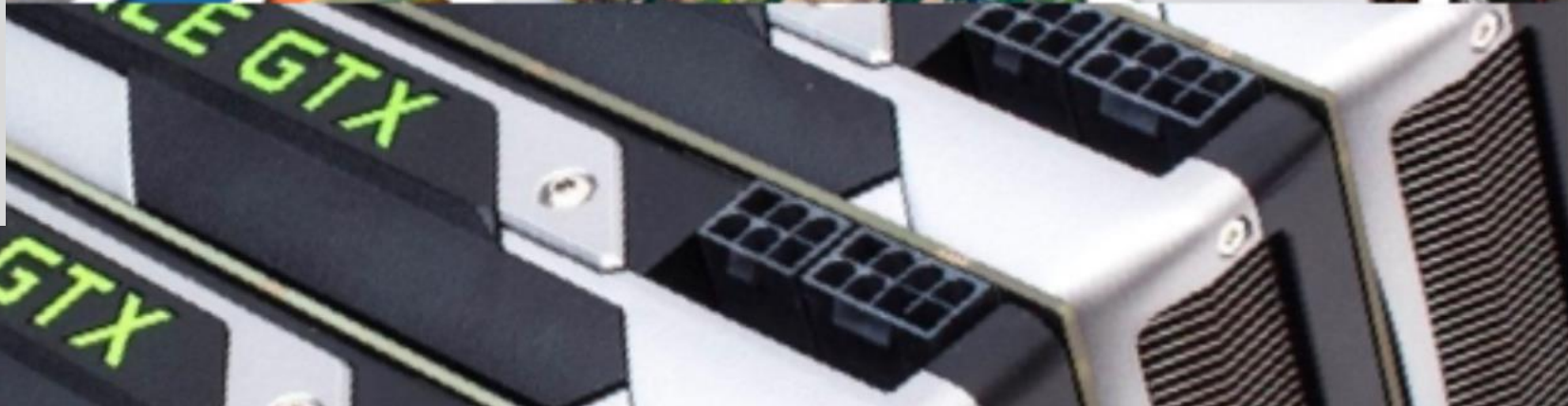
Алгоритмы



Данные



Вычисления



С 2012 по сейчас: глубокое обучение везде

TEXT PROMPT

an armchair in the shape of an avocado. an armchair imitating an avocado.

AI-GENERATED IMAGES

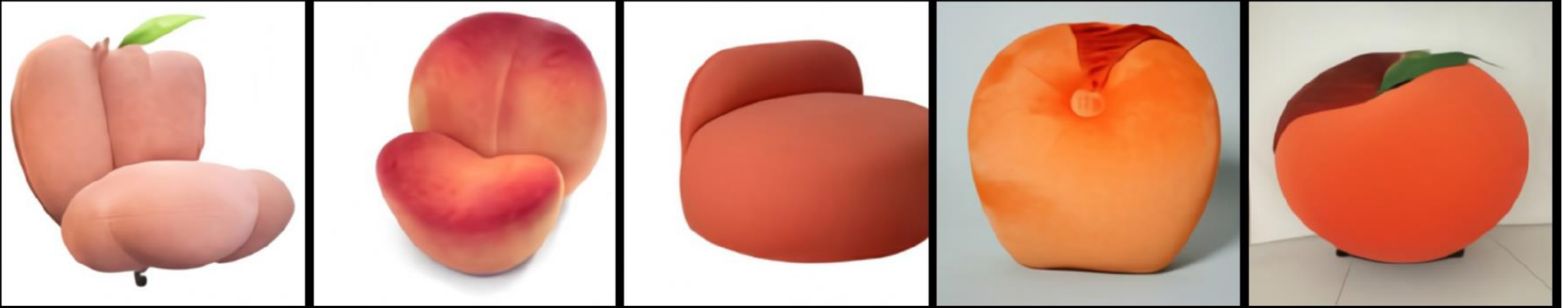


С 2012 по сейчас: глубокое обучение везде

TEXT PROMPT

an armchair in the shape of a peach. an armchair imitating a peach.

AI-GENERATED IMAGES



GPT-2: ответы на вопросы

GPT2 in the Economist

Q: How worried do you think we humans should be that machines will take our jobs?

A: It depends what role machine intelligence will play. Machine intelligence in some cases will be useful for solving problems, such as translation. But in other cases, such as in finance or medicine, it will replace people.

Q: Do fake news stories, generated using AI, pose a threat to democracy? Are we likely to see this tactic being used in the 2020 American presidential elections?

A: Yes, we'll probably see them. It's just a matter of when. Fake news stories are generally generated by political campaigns, and have the potential to have a huge impact on the outcome of an election. This is because they spread misinformation, often with no basis in fact.

GPT-3: перенос литературного стиля

Topic: humor: literary parodies and homages, pastiches, style parodies.

Parodies of the fantasy novel series *Harry Potter* in the style of various famous authors:

By Ernest Hemingway:

"It was a cold day on Privet Drive. A child cried. Harry felt nothing.

GPT-3: перенос литературного стиля

Topic: humor: literary parodies and homages, pastiches, style parodies.

Parodies of the fantasy novel series *Harry Potter* in the style of various famous authors:

By Ernest Hemingway:

"It was a cold day on Privet Drive. A child cried. Harry felt nothing. He was dryer than dust. He had been silent too long. He had not felt love. He had scarcely felt hate. Yet the Dementor's Kiss killed nothing. Death didn't leave him less dead than he had been a second before. It wasn't about living or dying really. It wasn't about death. It was about the Dark Lord, his creator, his engenderer. He was not resurrected. He was created again. He was whole again in the courtyard of Malfoy Manor."

GPT-3: перенос литературного стиля

Topic: humor: literary parodies and homages, pastiches, style parodies.

Parodies of the fantasy novel series *Harry Potter* in the style of various famous authors:

By Ernest Hemingway:

"It was a cold day on Privet Drive. A child cried. Harry felt nothing. He was dryer than dust. He had been silent too long. He had not felt love. He had scarcely felt hate. Yet the Dementor's Kiss killed nothing. Death didn't leave him less dead than he had been a second before. It wasn't about living or dying really. It wasn't about death. It was about the Dark Lord, his creator, his engenderer. He was not resurrected. He was created again. He was whole again in the courtyard of Malfoy Manor."

By Jane Austen:

"It is a truth universally acknowledged, that a broken Harry is in want of a book. This he knows to his cost; pressing that wretched nose against the window of a bookstore in Diagon Alley the last thing before he goes into hiding for several years, where upon, he goes straight to Azkaban."

<https://www.gwern.net/GPT-3#successes>

GPT-3: перенос литературного стиля

Topic: humor: literary parodies and homages, pastiches, style parodies.

Parodies of the fantasy novel series *Harry Potter* in the style of various famous authors:

By Ernest Hemingway:

"It was a cold day on Privet Drive. A child cried. Harry felt nothing. He was dryer than dust. He had been silent too long. He had not felt love. He had scarcely felt hate. Yet the Dementor's Kiss killed nothing. Death didn't leave him less dead than he had been a second before. It wasn't about living or dying really. It wasn't about death. It was about the Dark Lord, his creator, his engenderer. He was not resurrected. He was created again. He was whole again in the courtyard of Malfoy Manor."

By Jane Austen:

"It is a truth universally acknowledged, that a broken Harry is in want of a book. This he knows to his cost; pressing that wretched nose against the window of a bookstore in Diagon Alley the last thing before he goes into hiding for several years, where upon, he goes straight to Azkaban."

By Arthur Conan Doyle:

"Harry pushed at the swinging doors of the bookshop hard, and nearly knocked himself unconscious. He staggered in with his ungainly package, his cheeks scarlet with cold and the shame of having chosen the wrong month to go Christmas shopping. The proprietor of the store, however, didn't cast even a cursory glance at him, being far more interested in an enormous hunk of wood lying in the middle of the floor, which certainly looked like a gallows. Yes, the proprietor said to a reedy old man wearing a bowler hat and a forlorn expression that made Harry want to kick him, I can rent you such a gallows for a small fee."

<https://www.gwern.net/GPT-3#successes>

GPT-3: генерация историй

GPT3

You arrive at a small wooden store with sheet music in the window. The door is locked so you knock. The music suddenly stops and you hear footsteps coming to the door.

"Hello?" A voice says behind the door.

"I'm looking for the book of essence," You say.

"What do you want it for?" The voice answers, a bit of an edge to it.



> You tell them you are a humble seeker of knowledge.

"I am a humble seeker of knowledge," You answer.

"Ah, I see. You seek to learn about the essence of reality, and the universe." The voice says. "I know of this book. I am the guard of this library. The only way you can get to it, is if you get past me. Are you up for the challenge?"

> You ask them what the challenge is.

"The challenge is, you have to win a game of chess against me. If you win, then you may try to take the book. I will get the board." The door unlocks and opens.

GPT-3: copilot.github.com



The image shows a code editor interface with a dark background. At the top, there are four tabs: 'sentiments.ts' (selected), 'write_sql.go', 'parse_expenses.py', and 'addresses.rb'. The main area displays the following code:

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5
6
7
8
9
10
11
12
13
14
15
16
17
```

<https://github.com/features/copilot/>

Комбинируем VAE + авторегрессию: VQ-VAE2

Обусловленные классом сэмплы 256 x 256, модель обучена на ImageNet



Razavi et al, "Generating Diverse High-Fidelity Images with VQ-VAE-2", NeurIPS 2019

Комбинируем VAE + авторегрессию: VQ-VAE2

Обусловленные классом сэмплы 256 x 256, модель обучена на ImageNet



Комбинируем VAE + авторегрессию: VQ-VAE2

Сгенерированные лица 1024 x 1024, модель обучена на FFHQ



Razavi et al, "Generating Diverse High-Fidelity Images with VQ-VAE-2", NeurIPS 2019

Комбинируем VAE + авторегрессию: VQ-VAE2

Сгенерированные лица 1024 x 1024, модель обучена на FFHQ



Razavi et al, "Generating Diverse High-Fidelity Images with VQ-VAE-2", NeurIPS 2019

StyleGAN2



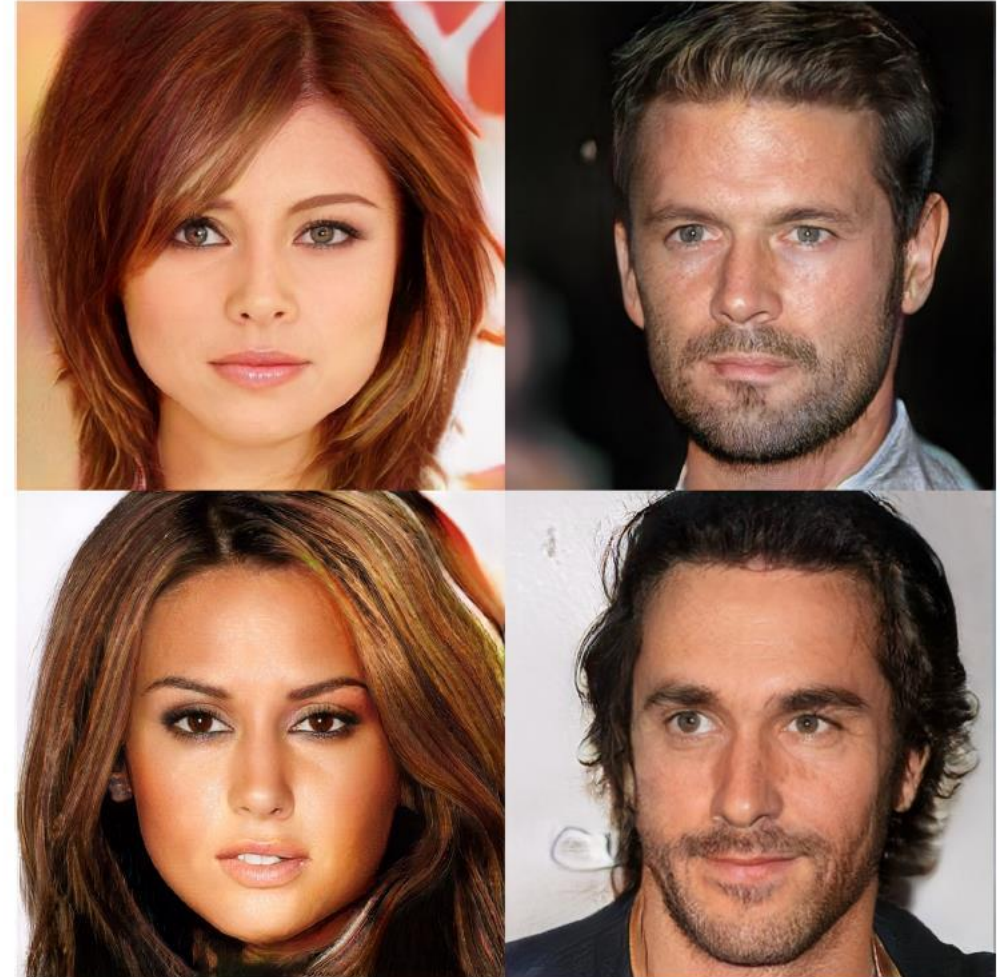
Karras et al, "Analyzing and Improving the Image Quality of StyleGAN", CVPR 2020

Улучшения GAN: увеличено разрешение

спальни, 256x256

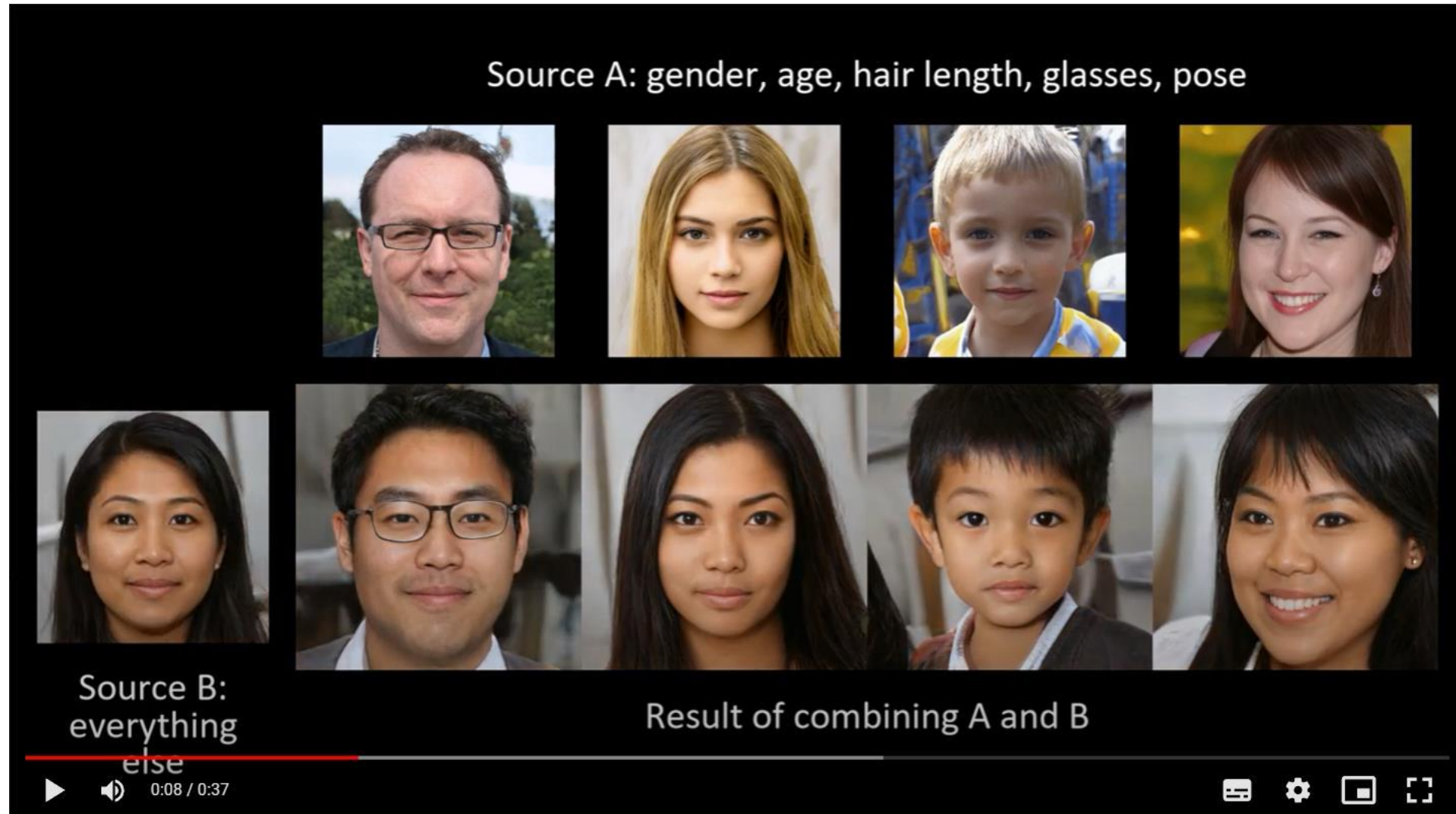


лица, 1024x1024



Karras et al, "Progressive Growing of GANs for Improved Quality, Stability, and Variation", ICLR 2018

StyleGAN2

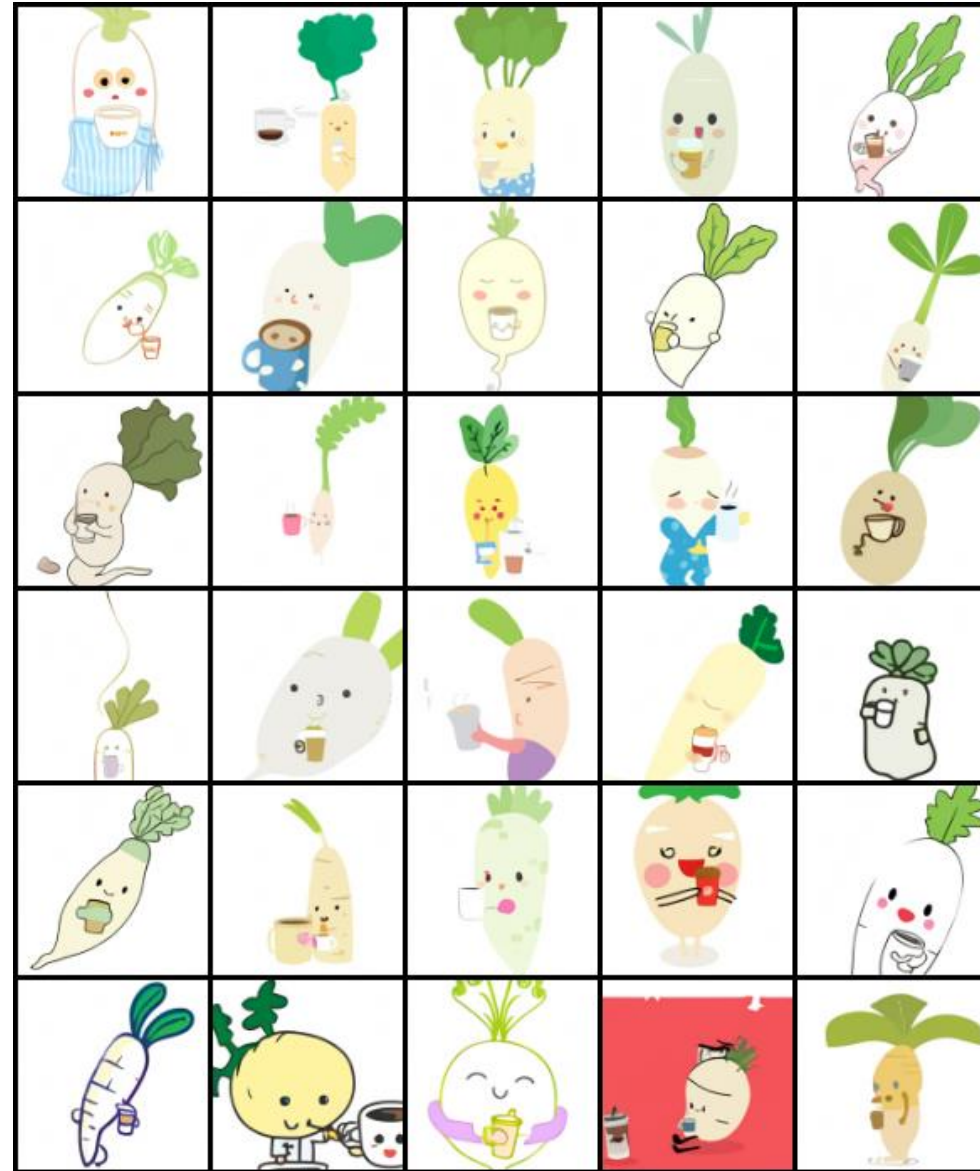


https://drive.google.com/drive/folders/1NFO7_vH0t98J13ckJYFd7kuaTkYeRJ86

DALL·E

an illustration of a baby daikon radish in pajamas sipping a latte

“ We find that DALL·E is sometimes able to transfer some human activities and articles of clothing to animals and inanimate objects, such as food items. ”



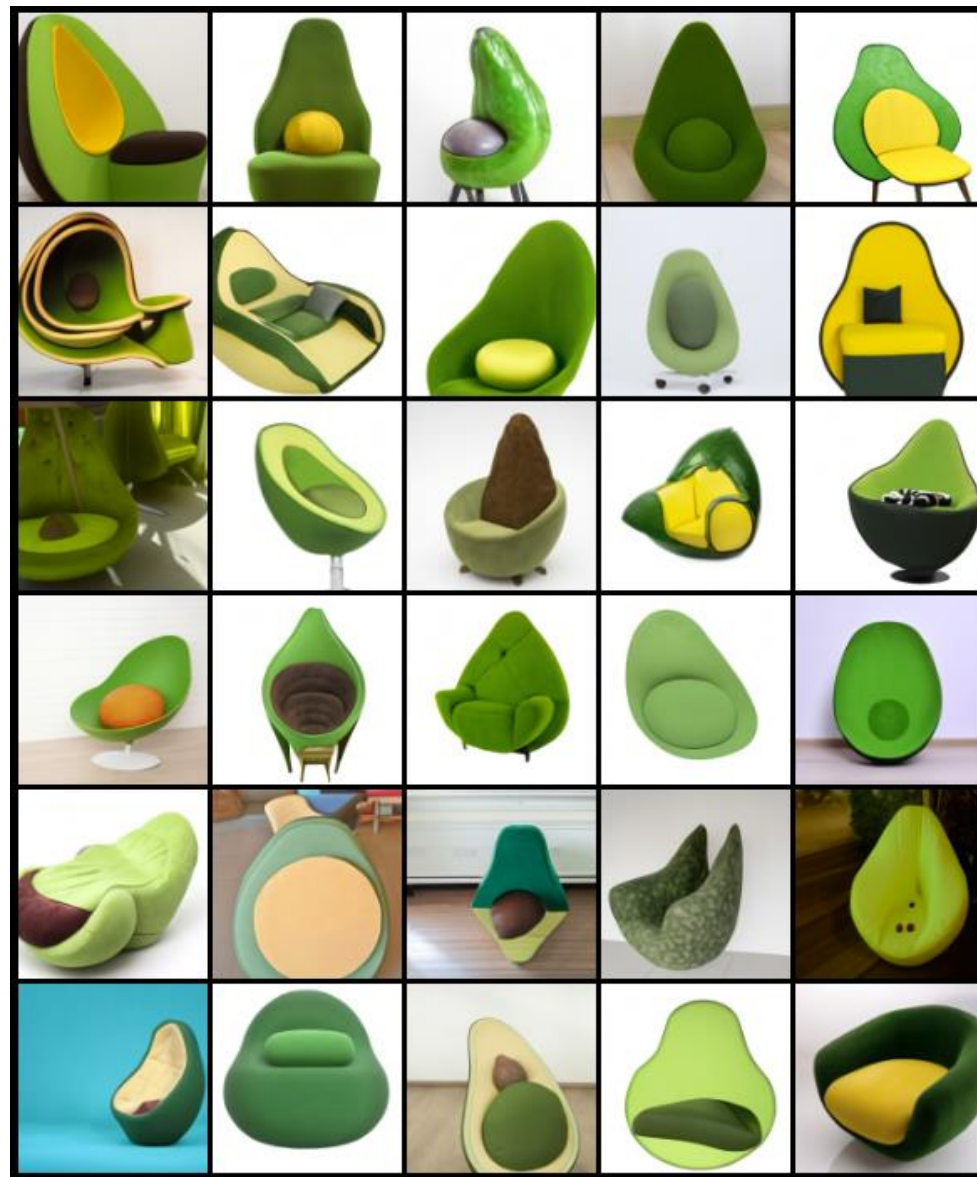
DALL·E

an armchair in the shape of an avocado,
an armchair imitating an avocado

“ Here, we explore its ability to take inspiration from an unrelated idea while respecting the form of the thing being designed, ideally producing an object that appears to be practically functional.

When generating some of these objects, such as “an armchair in the shape of an avocado”, DALL·E appears to relate the shape of a half avocado to the back of the chair, and the pit of

the avocado to the cushion. ”



DALL·E 2: создание новых изображений

TEXT DESCRIPTION

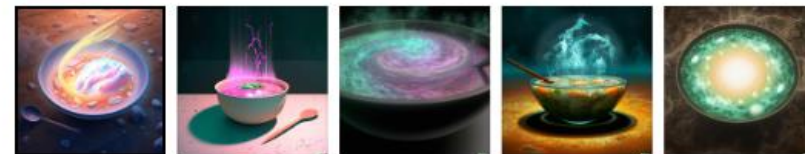
An astronaut Teddy bears A bowl
of soup

that is a portal to another
dimension that looks like a
monster as a planet in the
universe

as digital art in the style of
Basquiat drawn on a cave wall



DALL·E 2



<https://openai.com/dall-e-2/>

DALL·E 2: создание новых изображений

TEXT DESCRIPTION

An astronaut Teddy bears A bowl of soup

that is a portal to another dimension that looks like a monster as a planet in the universe

knitted out of wool spray-painted on a wall made out of plasticine



DALL·E 2



DALL·E 2: создание новых изображений

TEXT DESCRIPTION

An astronaut Teddy bears A bowl of soup

that is a portal to another dimension that looks like a monster as a planet in the universe

as a 1960s poster as mixed media with needlework as digital art



DALL·E 2



DALL·E 2: дополнение изображений

Оригинал



Girl with a Pearl Earring
Johannes Vermeer, 1665

DALL·E 2: дополнение изображений

Дополненное изображение

Оригинал

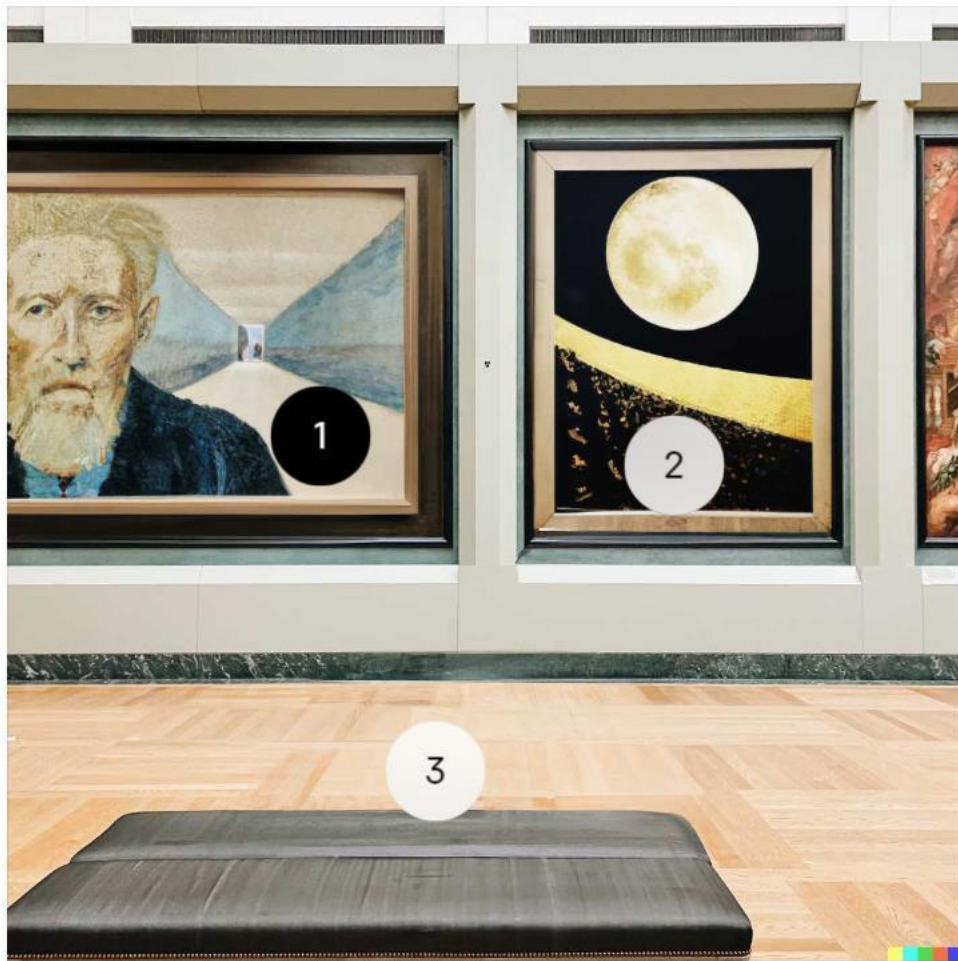


Girl with a Pearl Earring
Johannes Vermeer, 1665



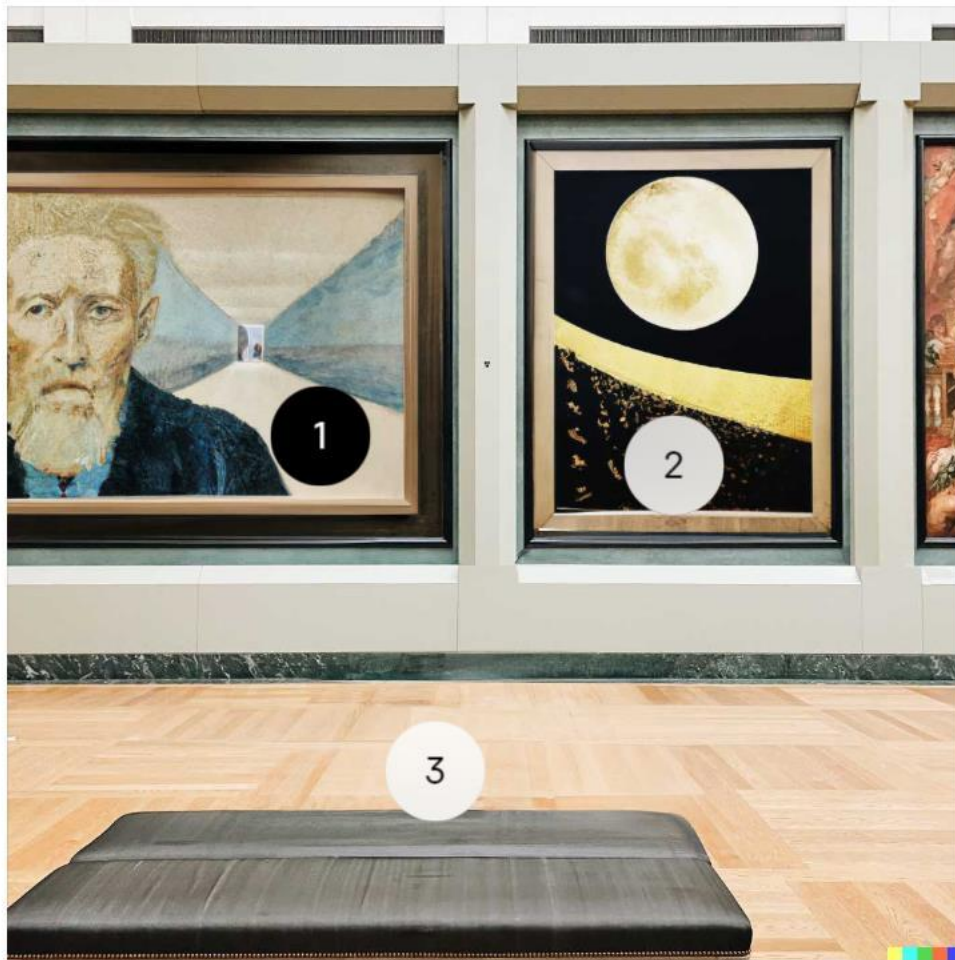
DALL·E 2: редактирование изображений

ORIGINAL IMAGE

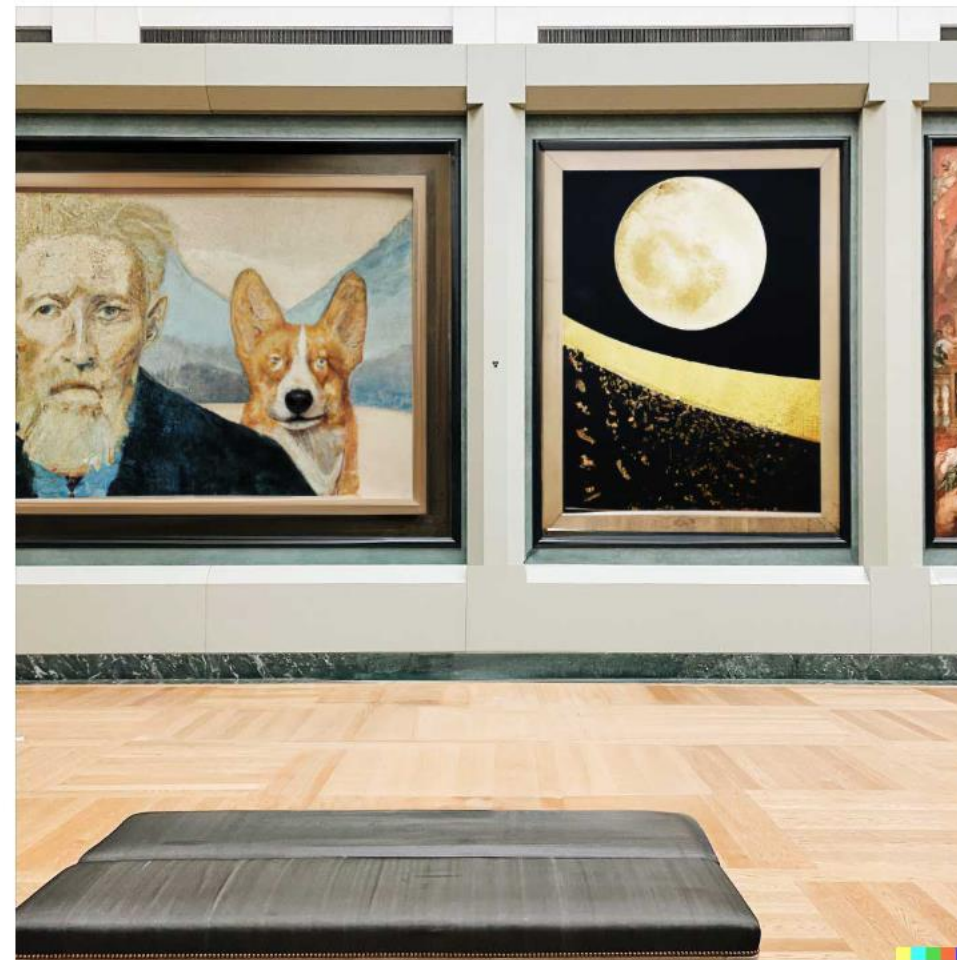


DALL·E 2: редактирование изображений

ORIGINAL IMAGE

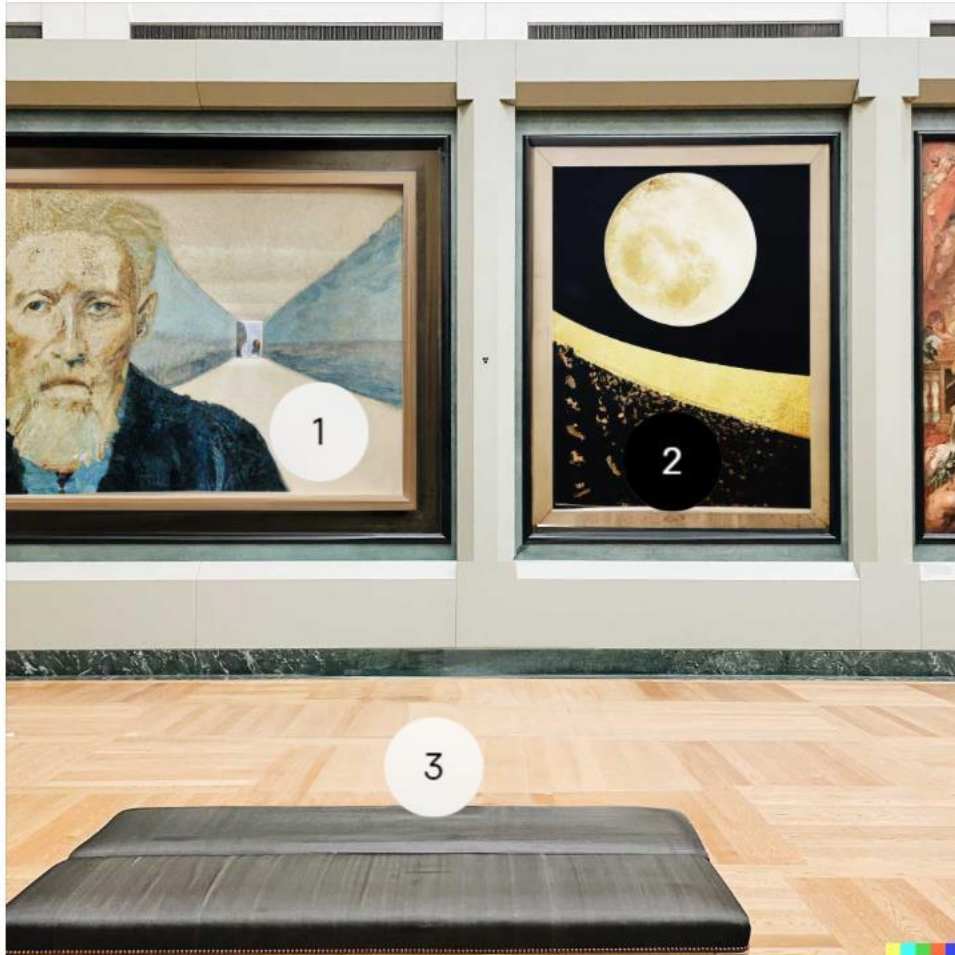


DALL·E 2 EDITS



DALL·E 2: редактирование изображений

ORIGINAL IMAGE

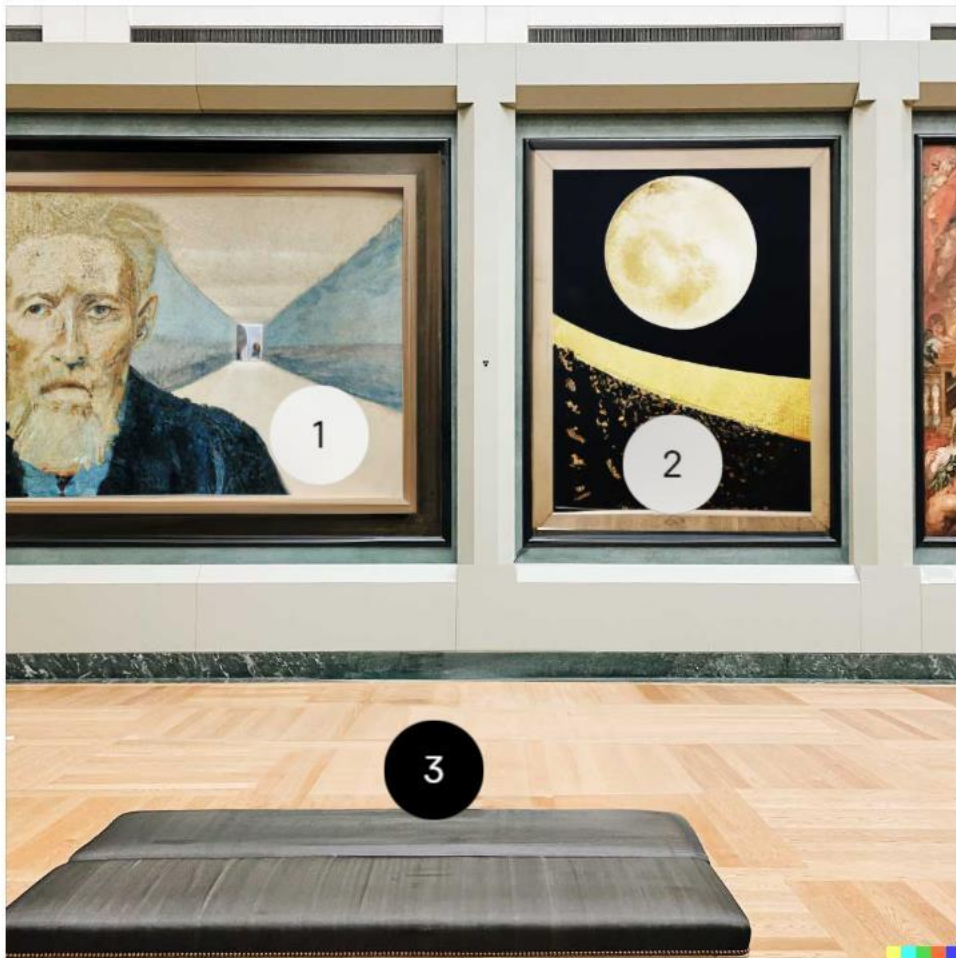


DALL·E 2 EDITS



DALL·E 2: редактирование изображений

ORIGINAL IMAGE



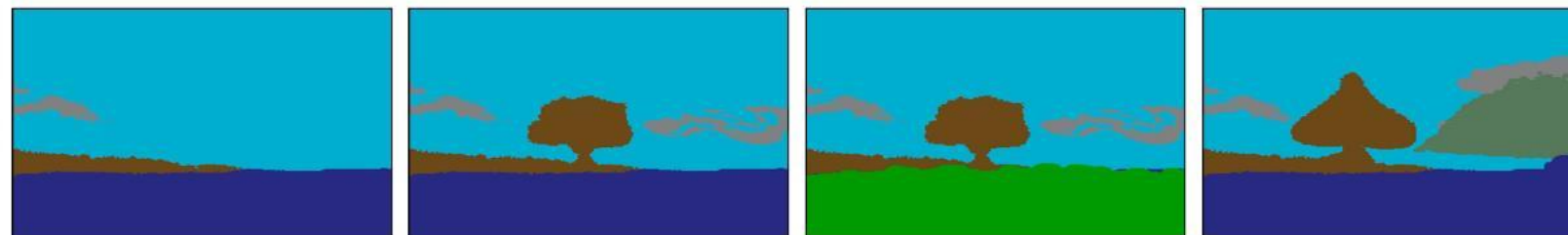
DALL·E 2 EDITS



Сегментные карты

Вход: сегментная карта

cloud	sky
tree	mountain
sea	grass



Semantic Manipulation Using Segmentation Map →

Вход:
стилевое
изображение



Stylization using Guide Images ↓



Методы супер-разрешения: Low-Res в High-Res

bicubic
(21.59dB/0.6423)



SRResNet
(23.53dB/0.7832)



SRGAN
(21.15dB/0.6868)



original

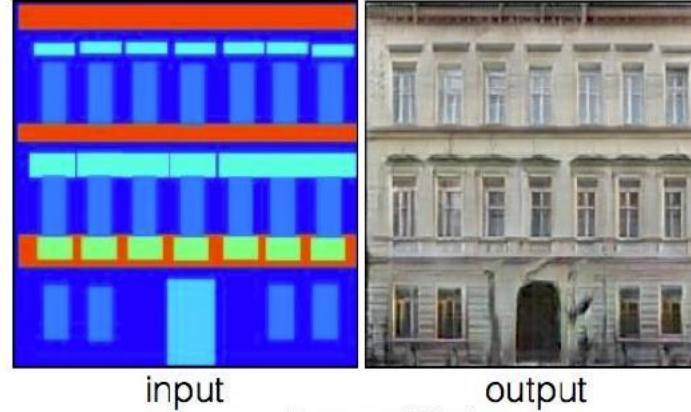


Image-to-Image трансляция: Pix2Pix

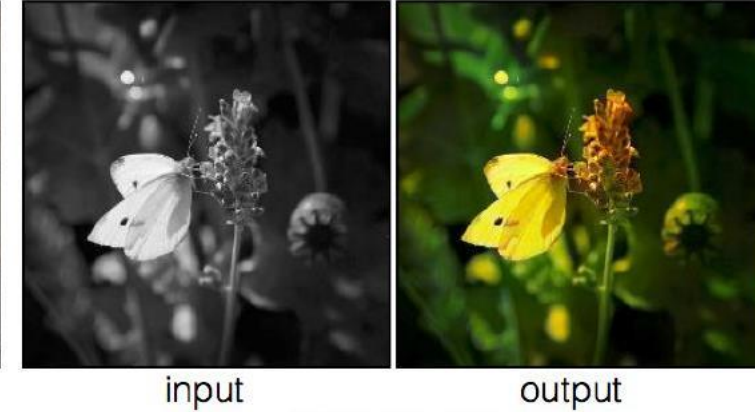
Labels to Street Scene



Labels to Facade



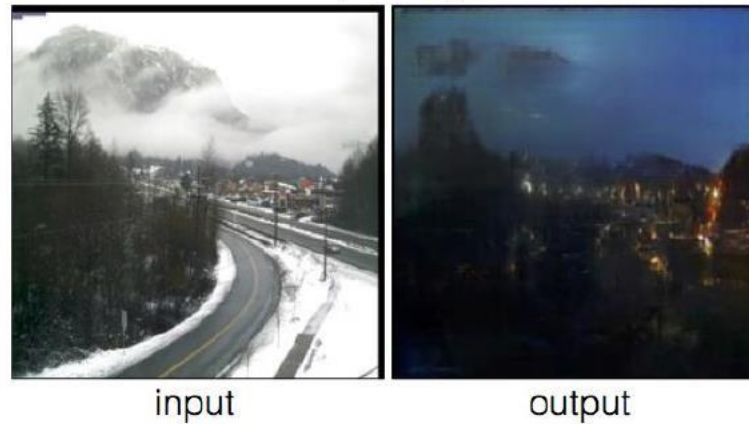
BW to Color



Aerial to Map



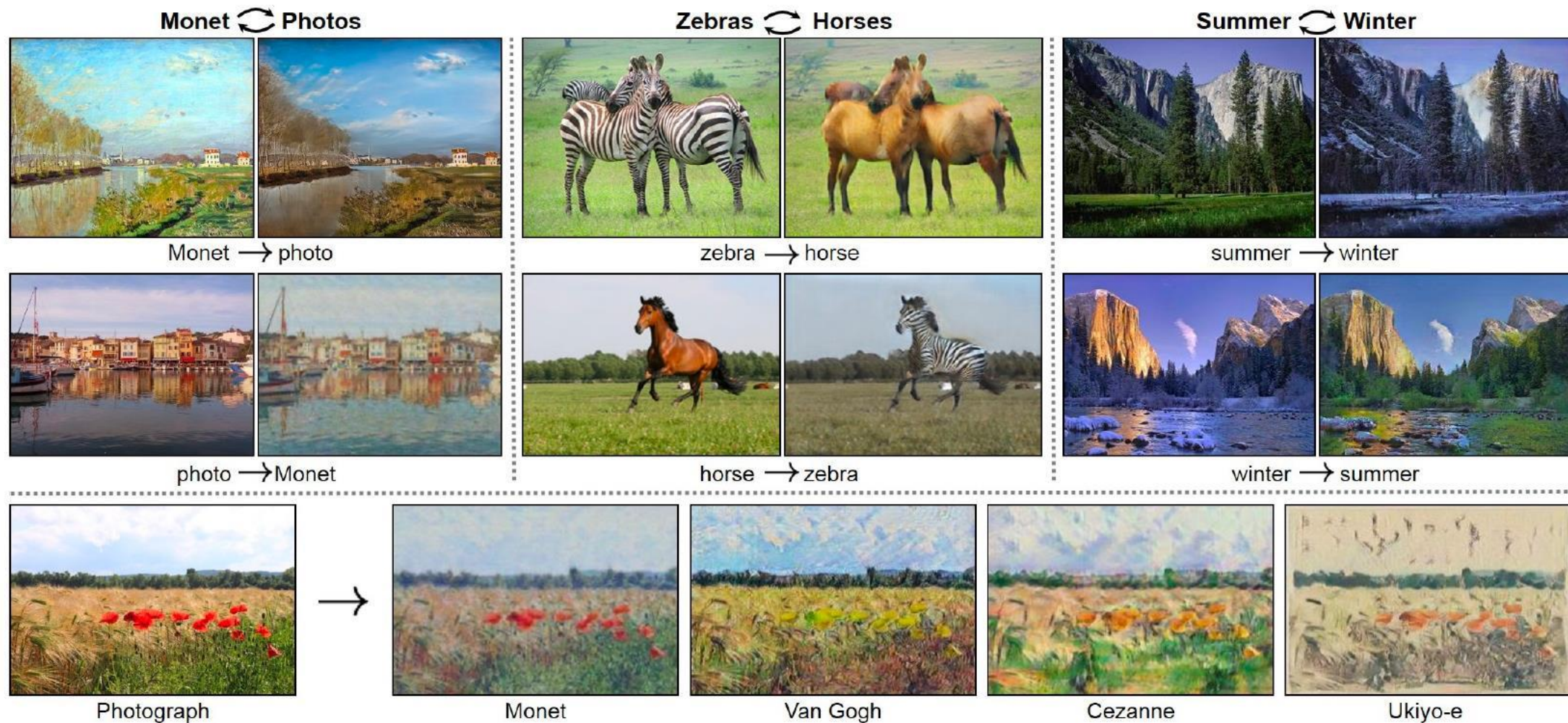
Day to Night



Edges to Photo



Непарная трансляция картинки в картинку: CycleGAN



Вариационные автоэнкодеры

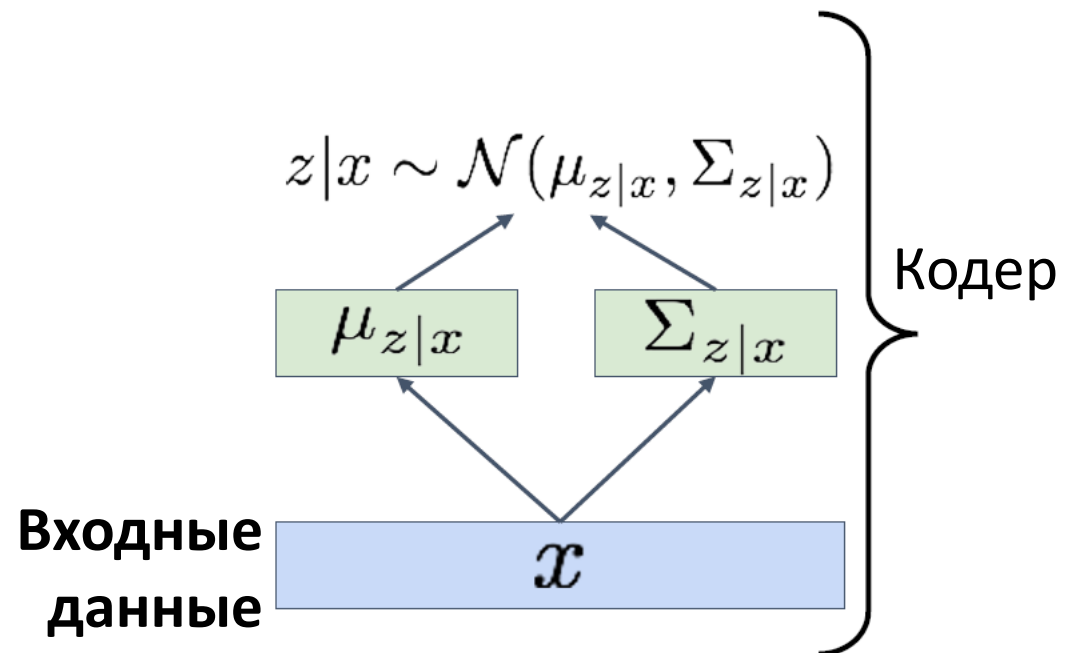
Обучаем максимизацией **нижней вариационной границы**

$$E_{z \sim q_\phi(z|x)} [\log p_\theta(x|z)] - D_{KL}(q_\phi(z|x), p(z))$$

1. Пропускаем входные данные через **кодер**, чтобы получить распределение по скрытым кодам.
2. **Выход кодера должен соответствовать $p(z)$!**

$$\begin{aligned} -D_{KL}(q_\phi(z|x), p(z)) &= \int_Z q_\phi(z|x) \log \frac{p(z)}{q_\phi(z|x)} dz \\ &= \int_Z N(z; \mu_{z|x}, \Sigma_{z|x}) \log \frac{N(z; 0, I)}{N(z; \mu_{z|x}, \Sigma_{z|x})} dz \\ &= \frac{1}{2} \sum_{j=1}^J \left(1 + \log \left((\Sigma_{z|x})_j^2 \right) - (\mu_{z|x})_j^2 - (\Sigma_{z|x})_j^2 \right) \end{aligned}$$

Есть решение в закрытой аналитической форме, когда q_ϕ является диагональной гауссианой, а p – единичной гауссианой (J – размер z)



Генеративные состязательные сети

Совместно обучаем генератор G и дискриминатор D с помощью **минимаксной игры**

Обучаем G и D , используя чередующиеся шаги обновления градиента

$$\min_G \max_D \left(E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p(z)} \left[\log \left(1 - D(G(z)) \right) \right] \right)$$

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

For t in $1, \dots, T$:

1. Обновляем D $D = D + \alpha_D \frac{\partial V}{\partial D}$

2. Обновляем G $G = G - \alpha_G \frac{\partial V}{\partial G}$

Трансформер

Послойная нормализация

(Layer Normalization):

Даны: h_1, \dots, h_N (форма D)

параметр масштаба: γ (форма D)

параметр сдвига: β (форма D)

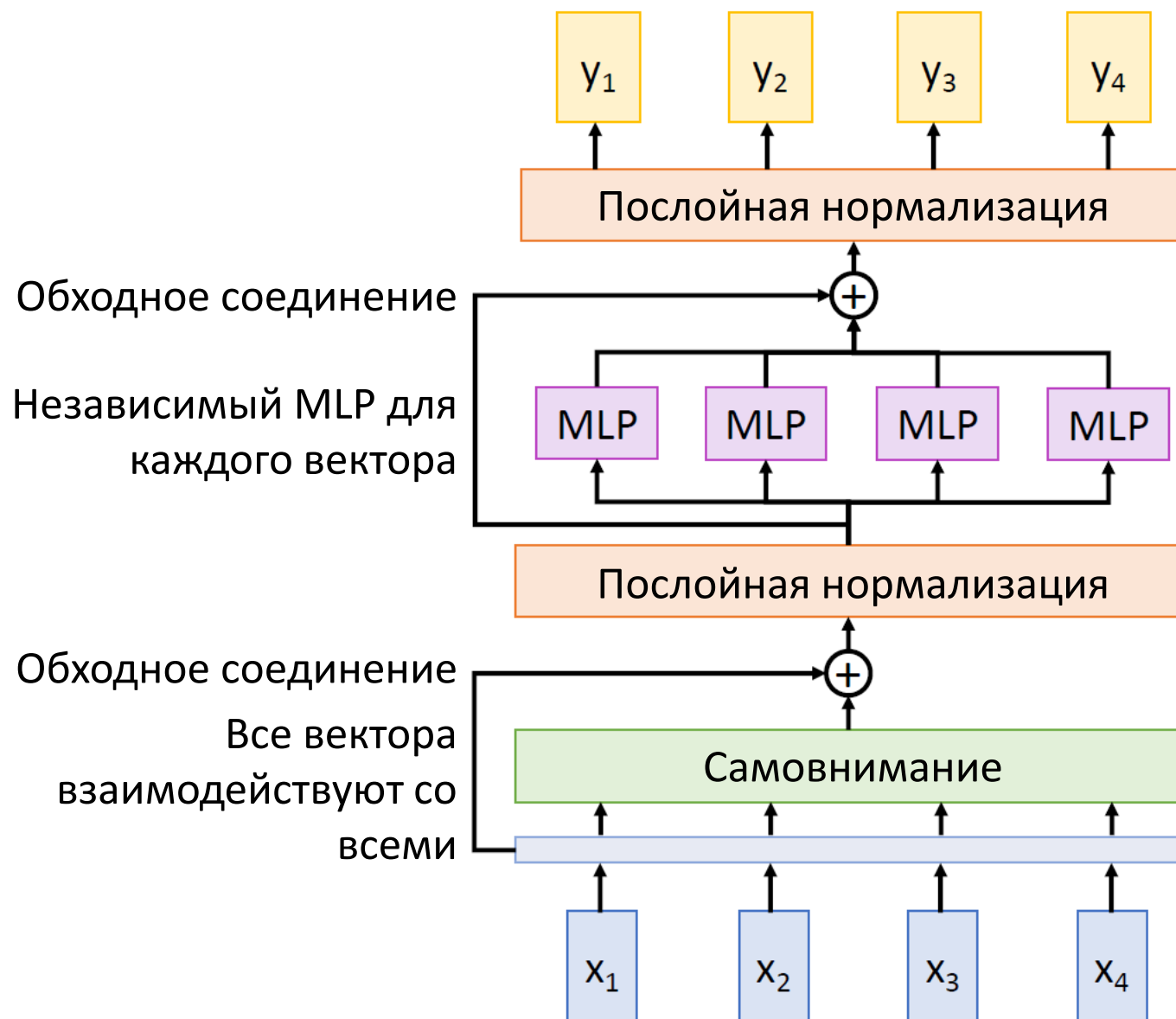
$$\mu_i = \frac{(\sum_j h_{i,j})}{D} \quad (\text{скаляр})$$

$$\sigma_i = \left(\frac{(\sum_j h_{i,j} - \mu_i)^2}{D} \right)^{1/2} \quad (\text{скаляр})$$

$$z_i = (h_i - \mu_i) / \sigma_i \quad (\text{форма D})$$

$$y_i = \gamma \cdot z_i + \beta \quad (\text{форма D})$$

Va et al, 2016



Дорожная карта

Обязательно:

- Математический анализ
- Теория функций комплексного переменного
- Линейная алгебра и аналитическая геометрия
- Теория вероятностей и математическая статистика
- Дискретная математика
- Теория информации
- Программирование (Python), алгоритмы и структуры данных
- Теория оптимизации
- Машинное обучение
 - Обучение с учителем
 - Обучение без учителя
 - Обучение с подкреплением
 - Ансамбли
 - Глубокое обучение
 - Большие данные

Желательно:

- многомерный анализ
- функциональный анализ
- вариационное исчисление
- численные методы