

Что дальше?

План занятия

Сегодня мы обсудим:

- Какие еще задачи могут решать нейросети:
 - CV
 - NLP/звук
 - Другие задачи
 - Мультимодальность
- Как/где можно легко использовать предобученные нейросети для разных задач;
- Что делать дальше после курса, куда идти?
- Ответы на ваши вопросы

Какие еще задачи могут решать нейросети

Computer Vision

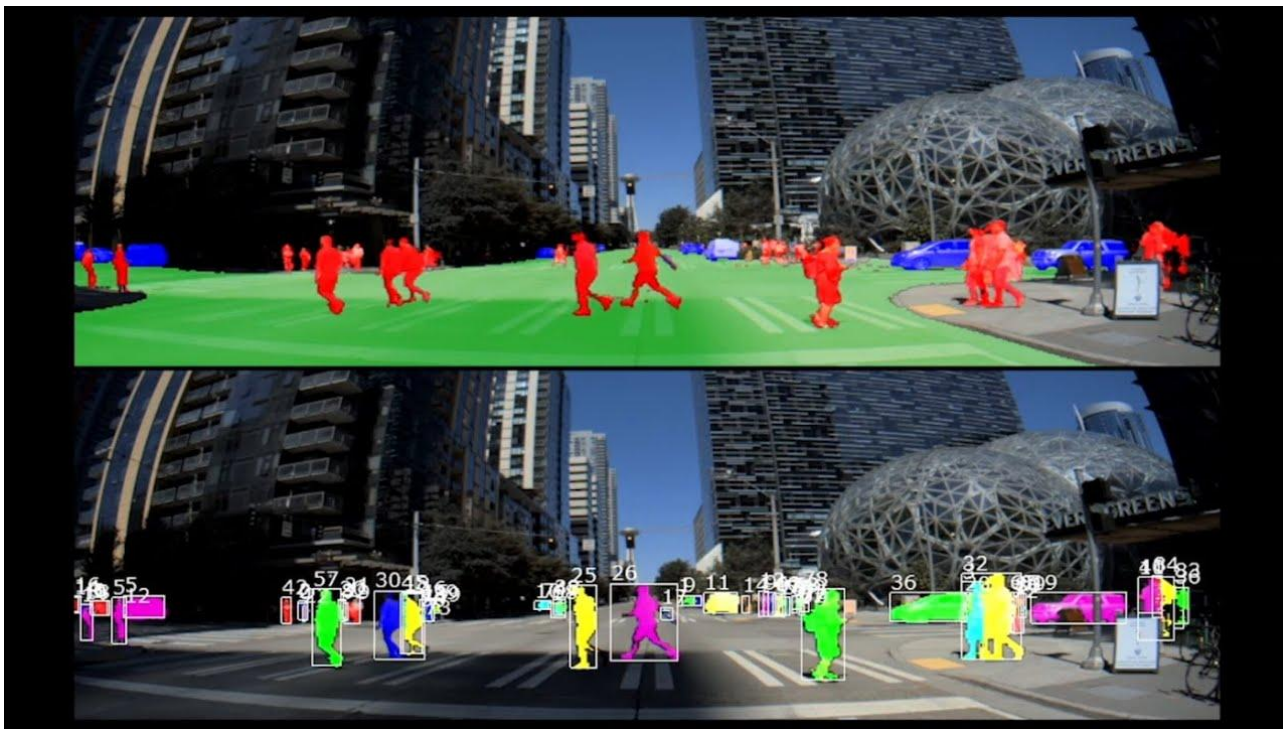
Беспилотные автомобили

В беспилотниках решается много различных задач CV:

- Классификация;
- Детекция;
- Сегментация;
- Поиск по изображениям;
- Оценка положения
- ...



Беспилотные автомобили



То, как “видит” окружающий мир беспилотный автомобиль

Перенос стиля (Style Transfer)

Исходное
изображение



+

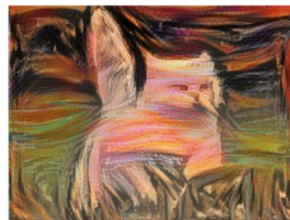
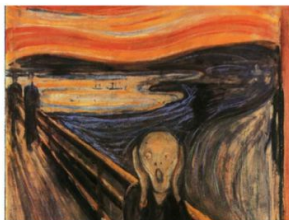
Картинка стиля



Итоговая
картинка



+



+



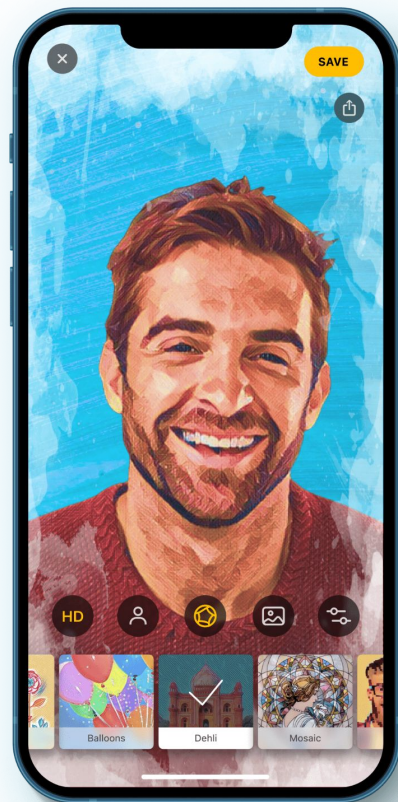
Перенос стиля (Style Transfer)

Prisma: приложение для переноса
стиля изображений

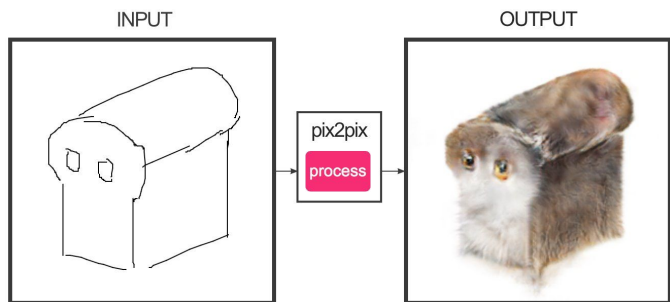


[Приложение для iOS](#)

[Приложение для Android](#)



- Генерация новых изображений;
- Дополнение частей изображений;
- Удаление частей изображений и заполнение пропусков



Распознавание лиц

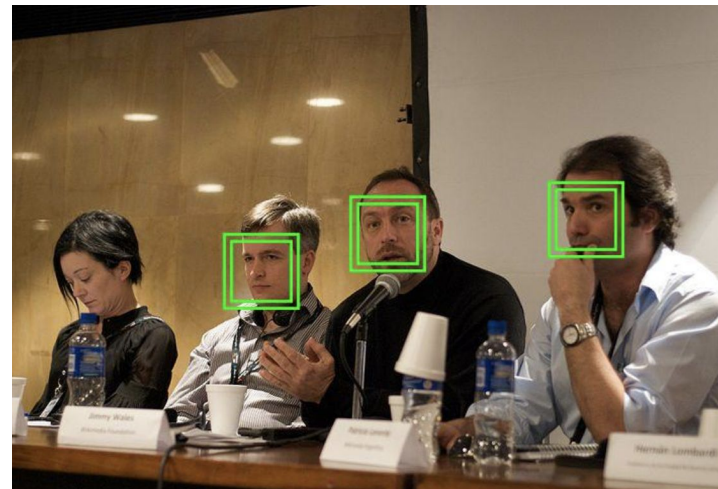
Система распознавания лиц состоит из трех этапов:

- Детекция;
- Alignment;
- Распознавание (классификация).

Распознавание лиц

Система распознавания лиц состоит из трех этапов:

- **Детекция;**
- Alignment;
- Распознавание (классификация).

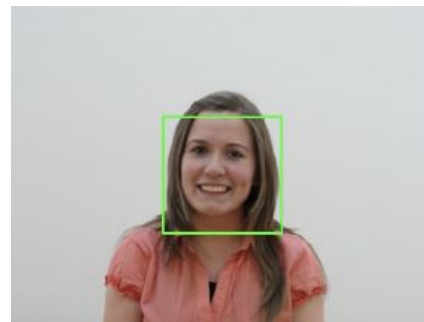


Нейросеть ищет на изображении все лица

Распознавание лиц

Система распознавания лиц состоит из трех этапов:

- Детекция;
- **Alignment**;
- Распознавание (классификация).



На изображении лица ищутся ключевые точки: точки глаз, носа, рта

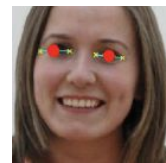
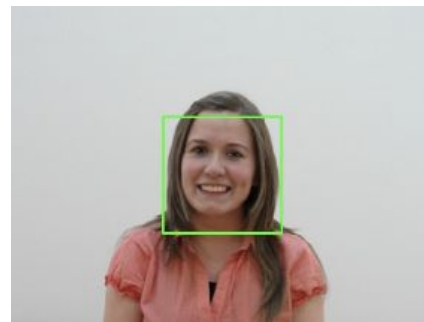
Распознавание лиц

Система распознавания лиц состоит из трех этапов:

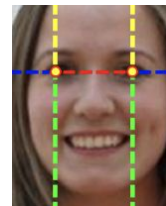
- Детекция;
- **Alignment**;
- Распознавание (классификация).



Изображение поворачивается и растягивается так, чтобы глаза занимали определенное положение на картинке. А также чтобы между глазами было определенное расстояние



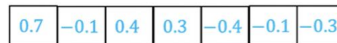
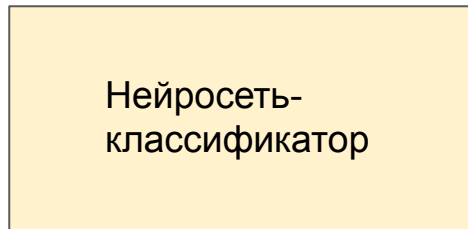
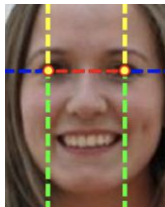
На изображении лица ищутся ключевые точки: точки глаз, носа, рта



Распознавание лиц

Система распознавания лиц состоит из трех этапов:

- Детекция;
- Alignment;
- **Распознавание (классификация).**



эмбеддинг лица

Этот эмбеддинг затем можно сравнивать по метрике L2 с эмбеддингами других лиц из базы данных



Изображение подается в нейросеть для классификации

Все рассмотренные задачи применимы и к **видео**



Дипфейки (deepfake)

Дипфейк — видео, созданное с помощью нейросетей, которое похоже на настоящее. Обычно дипфейки призваны обмануть людей: убедить их, что видео настоящее.

Дипфейки часто используются в мошеннических целях. Вот [история](#) из 2020 года о том, как мошенники с помощью подобного видео разводили людей на деньги



Natural Language Processing

Задачи NLP

Некоторые задачи сферы NLP:

- Классификация текстов, анализ тональности;
- Языковое моделирование;
- Машинный перевод;
- Named Entity Recognition;
- Question-answering, диалоговые системы;
- Суммаризация текста

Albert Einstein **PER** Albert Einstein was born in **Ulm LOC** in **Germany LOC** on March 14, 1879. Six weeks later the family moved to **Munich LOC**, where he later on began his schooling at the **Luitpold Gymnasium ORG**. In 1896 he entered the **Swiss Federal Polytechnic School ORG** in **Zurich LOC** to be trained as a teacher in physics and mathematics.

Пример NER

GPT-3

GPT-3 — огромная языковая модель от OpenAI. Ее обучали на очень большом количестве данных.

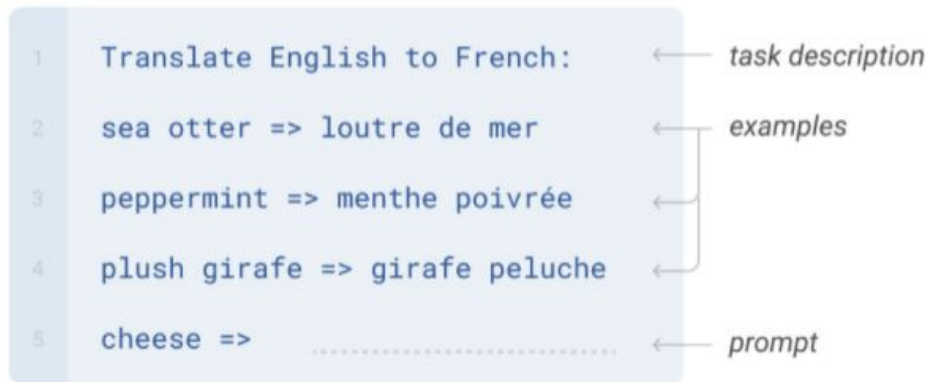
Она вышла в 2019 году, и про нее тогда очень много писали.



GPT-3

У GPT-3 были пара интересных особенностей:

- Ее можно было заставить генерировать не только текст. Например, сайты (html-страницы), код и т.д.;
- Ее можно заставить решать задачи в zero-shot режиме



ChatGPT

ChatGPT — модель на основе GPT-3, которую дообучили лучше взаимодействовать с пользователем. Она вышла в конце 2022 года. И она понимает русский!



Если бы Шекспир опоздал на работу, как бы он объяснил свое отсутствие?



Возможно, Шекспир мог бы объяснить свое отсутствие так:

"Мое невезение, как невезение траура, занимает меня с утра. Моя шляпа, принадлежащая мне, покрыта пеплом от запаха дыма, который выходит из горелки, которую я невольно поджег вчера вечером. Мое облако настроения напоминает мне о вечном темном небе. Но самое главное - величайшая неудача нашего времени обрушилась на меня. Моя лошадь, которую я купил недавно, погибла на дороге. Я был вынужден идти пешком, и так как я не знаю этого города так хорошо, что бы найти быстро путь, опоздание стало неизбежным. Прошу прощения за мои неудачи и опоздание."



Назови пять причин надеть клоунский нос на корпоратив



1. Для повышения настроения и развлечения. Клоунский нос может принести радость и смех на корпоратив, что может способствовать улучшению настроения участников и созданию более дружелюбной атмосферы.
2. Для создания необычного оформления. Клоунский нос может стать уникальным аксессуаром, который поможет создать необычный и запоминающийся образ на корпоративе.
3. Для продвижения бренда. Если ваш бренд связан с юмором и развлечением, то нос клоуна может стать хорошим способом продвижения бренда и укрепления его имиджа.
4. Для выделения себя из толпы. Если вы хотите выделить себя на корпоративе и привлечь к себе внимание, то нос клоуна может стать хорошим способом сделать это.
5. Для участия в косплей-ко

ChatGPT

ChatGPT — модель на основе GPT-3, которую дообучили лучше взаимодействовать с пользователем

Ссылки:

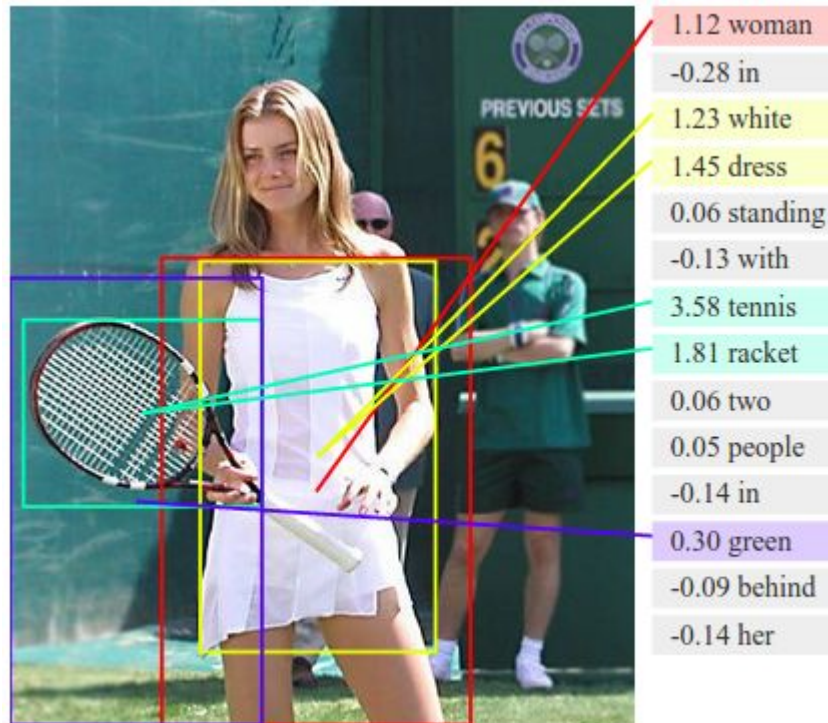
- [Статья медузы о том, как работает ChatGPT;](#)
- [Поболтать с ChatGPT;](#)
- [Статья](#) о том, что ChatGPT справляется с поиском лучше, чем Google Search;
- ChatGPT [умеет решать задачи](#)! Например, задачи с экзамена по математике или биологии.

Мультимодальность

Image captioning

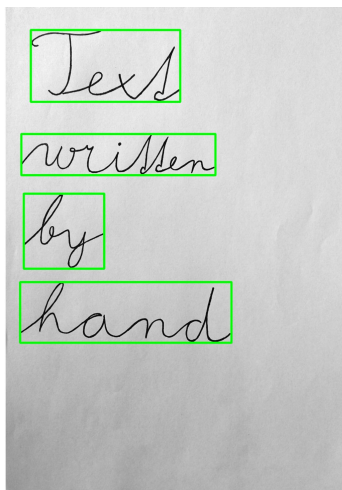
- Генерация описания к картинке/видео;
- Генерация субтитров к видео

В задаче image captioning начинают применять графовые нейросети. Граф отлично подходит для передачи структуры сцены и взаимодействий ее элементов



Optical Character Recognition (OCR)

Оптическое распознавание символов



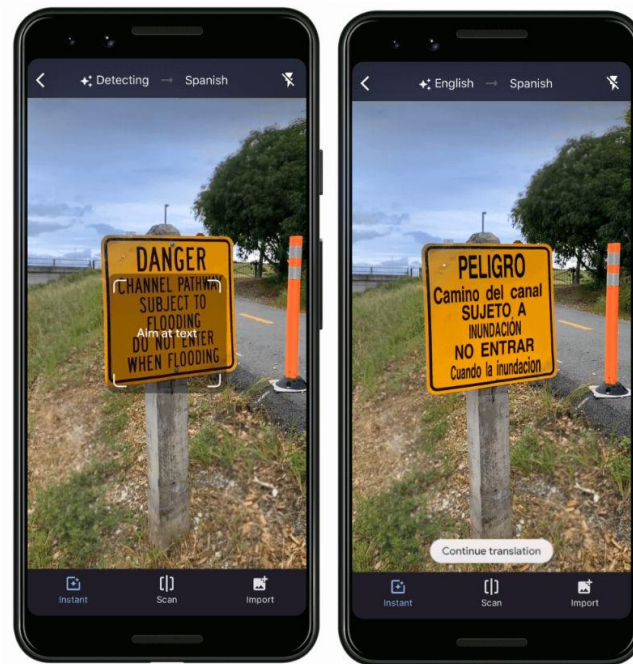
Text written by hand



Optical Character Recognition (OCR)

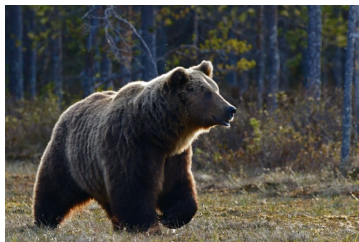
Применений OCR много:

- Автоматическое распознавание визиток, резюме;
- Автоматическая оцифровка текста, написанного от руки;
- Автоматическое распознавание текста и его перевод;
- Улучшение качества документов;
- ...

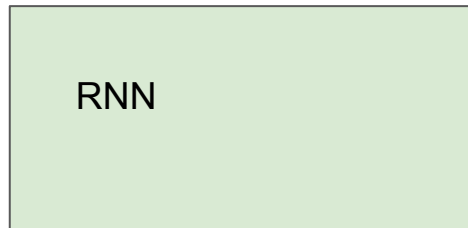
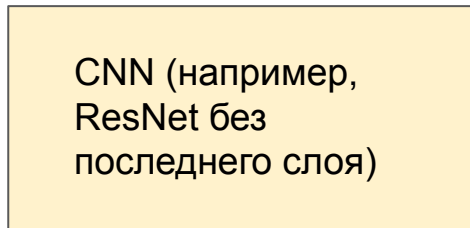


Автоматический перевод текста с английского на испанский с помощью камеры телефона

Идея устройства подобных нейросетей



Бурый медведь шагает по лесу

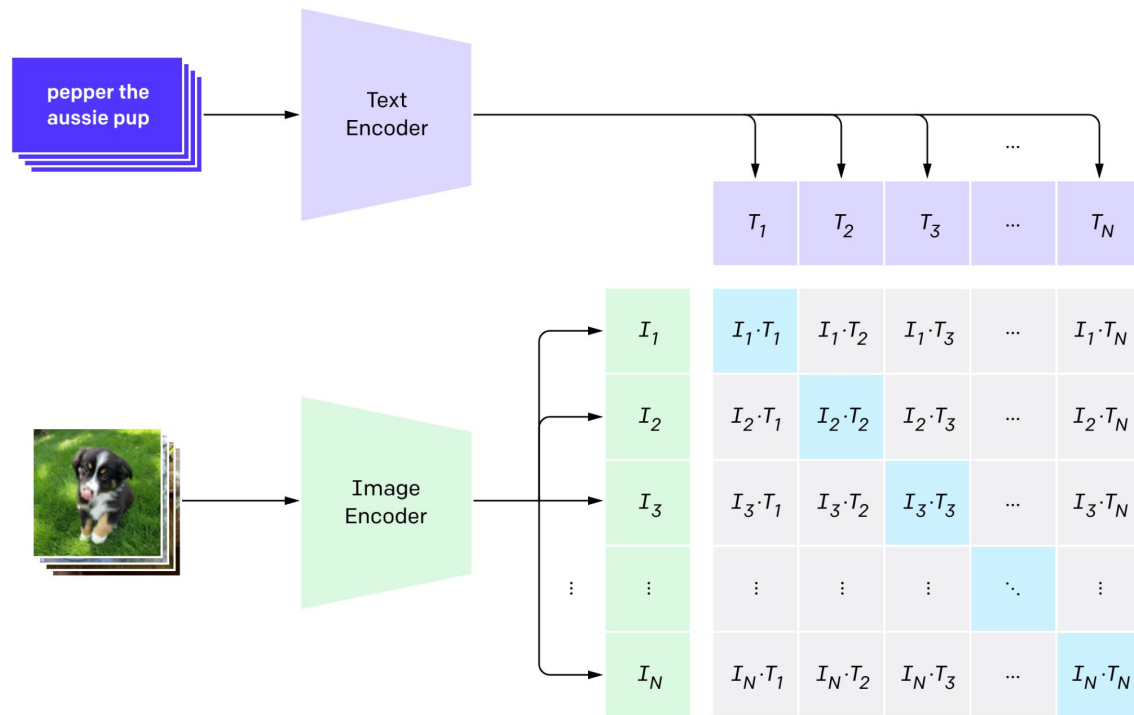


эмбединг
картинки

-0.8	-0.4	-0.5	0.1	-0.9	0.3	0.8
------	------	------	-----	------	-----	-----

CLIP

Идея: выучить
совместное
пространство
эмбеддингов для
картинок и текста



Генерация изображений по тексту



Фото ежика, сидящего в маленькой лодочке посреди озера. На нем гавайская рубашка и соломенная шляпа. Ежик читает книгу. На его фоне — листья.



Плюшевый мишка в белой футболке серфит в океане



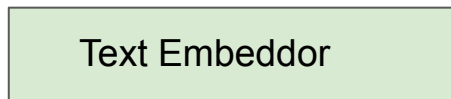
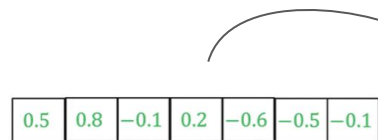
Красный Феррари, едущий по гравию в лесу. На фоне сияет радуга.

Вы можете сгенерировать картинки по своему текстовому запросу на сайте neural.love

Генерация изображений по тексту

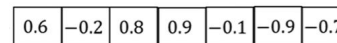
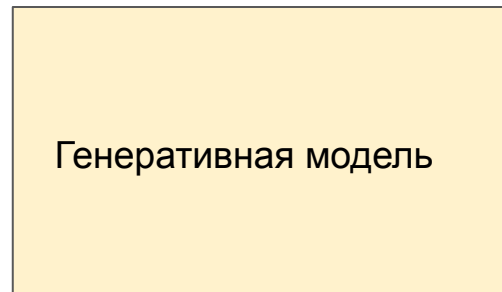
Идея, как может быть устроена нейросеть для генерации картинок

Например, BERT
или CLIP



Плюшевый мишка в белой футболке серфит в океане

condition



Случайный шум



Генерация изображений

Нейросети, с помощью которых можно делать генерацию изображений:

- Автоэнкодеры;
- GAN;
- Flow-based models;
- Диффузионные модели

В последний год большую популярность приобрели именно диффузионные модели.

В [этой статье](#) постарались очень просто и наглядно объяснить, на какой идее работает генерация картинок в диффузионных моделях

Поздравляем с завершением курса!

