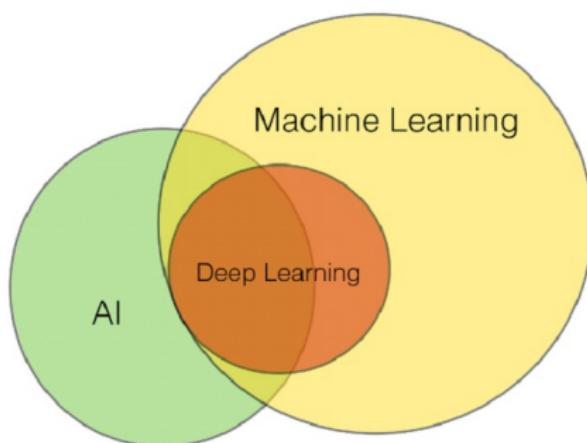


Введение в глубокое обучение

Виктор Китов
victorkitov.github.io



Глубокое обучение

- Традиционное машинное обучение полагается на человека при генерации признаков модели.

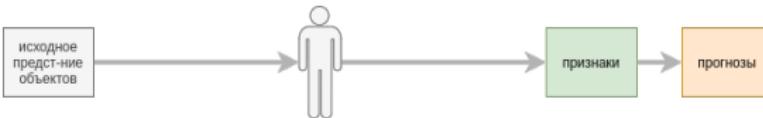
Традиционное машинное обучение (shallow learning):



Глубокое обучение

- Традиционное машинное обучение полагается на человека при генерации признаков модели.

Традиционное машинное обучение (shallow learning):



- Глубокое обучение настраивает не только модель, но и последовательность преобразований, генерирующих признаки.

Глубокое обучение (deep learning):



Глубокое обучение

- Преимущества:

- не нужно придумывать и вручную создавать признаки
- во многих задачах признаки слишком сложные
- признаки получаются лучше, итоговая модель точнее

- Недостатки:

- нужно гораздо больше обучающих примеров
- вычислительно сложнее

Содержание

1 Изображения

2 Текст

3 Звук

4 Графы

Исходное представление изображений

- Изображения - матрицы (grayscale) и тензоры (цветные) чисел.



150	90	80	150
110	75	85	110
120	40	35	150
150	100	130	150

150	90	80	150
110	75	85	110
120	40	35	150
150	100	130	150

- Сканы со спутников, медицинских устройств, сенсоров - тоже изображения.
- Видео - последовательность изображений (фреймов).

Классификация+локализация изображений



Classification

Cat

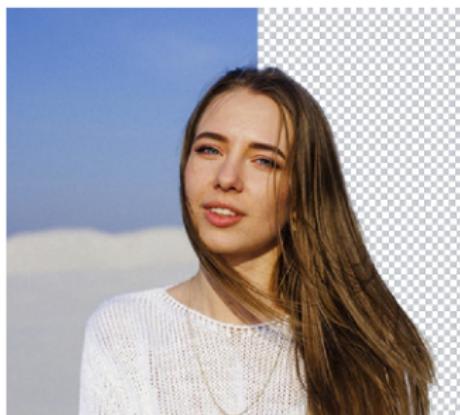


Classification, Localization

Cat

Выходы сети: C рейтингов классов + (x, y, h, w) .

Удаление+замена фона



Поиск похожих изображений (image retrieval)

Query	Top 10 Retrieved Images of DPFAH Method on Imagenet-100									
										
										
										
										
										
										
										
										

Семантическая сегментация

Детекция объектов

Сегментация объектов (instance segmentation)

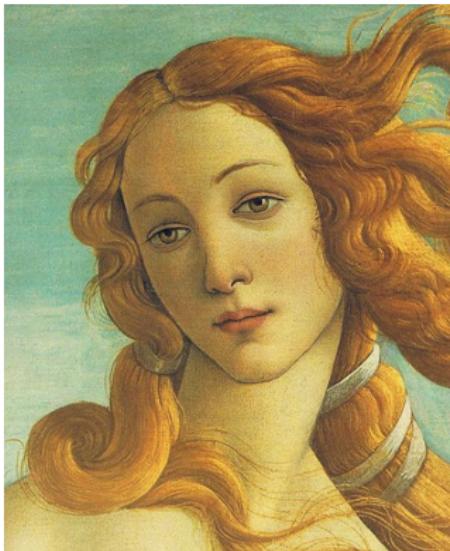
Распознавание жестов

Оценка глубины (depth estimation)

3D реконструкция по 2D снимкам

Стилизация изображений (style transfer)

Восстановление фото по картине¹



¹<https://mymodernmet.com/denis-shiryaev-neural-network-art/>

Оживление портретов (talking head generation)

Применение: виртуальные аватары, актёры, телеведущие.

Описание изображений

input



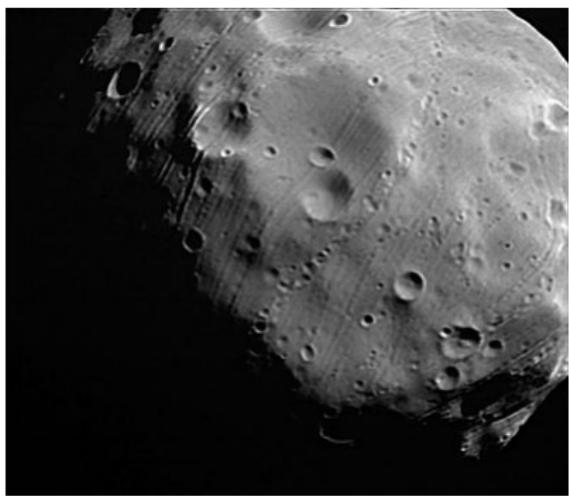
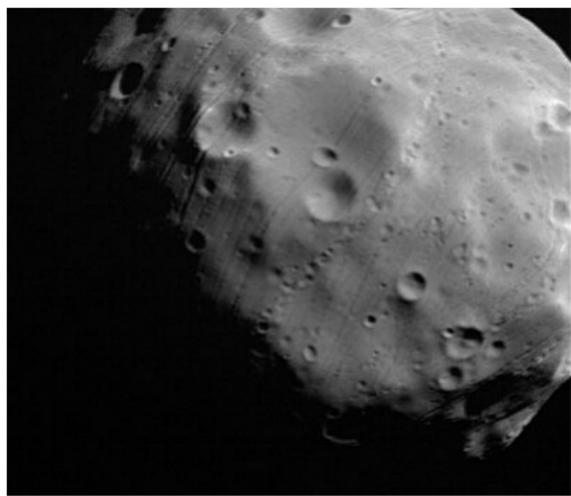
Report

Model
recognition



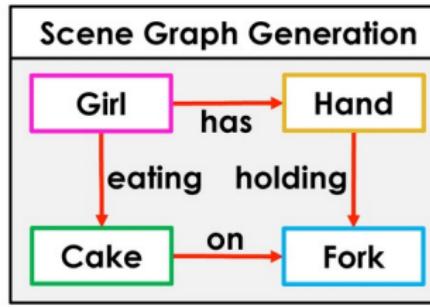
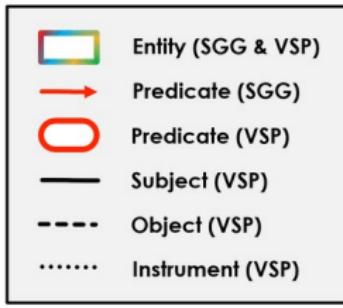
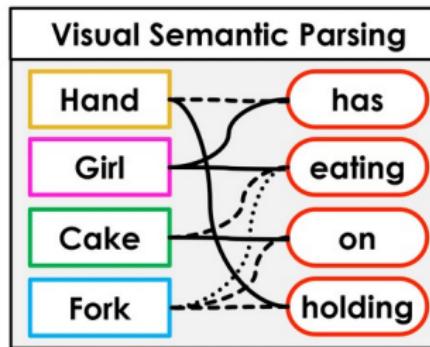
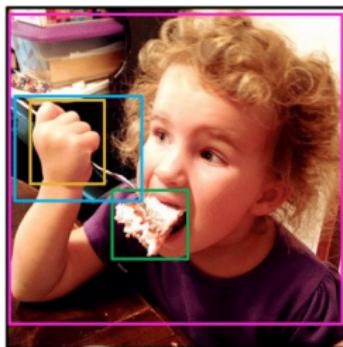
The heart size and pulmonary vascularity appear within normal limits. The lungs are free of focal airspace disease. No pleural effusion or pneumothorax. No acute bony abnormality.

Супер-разрешение (super-resolution)



Применения: телескопы, микроскопы, обычные фотоаппараты.

Анализ сцен (scene understanding)



Изображение -> граф объектов и их отношений друг к другу.

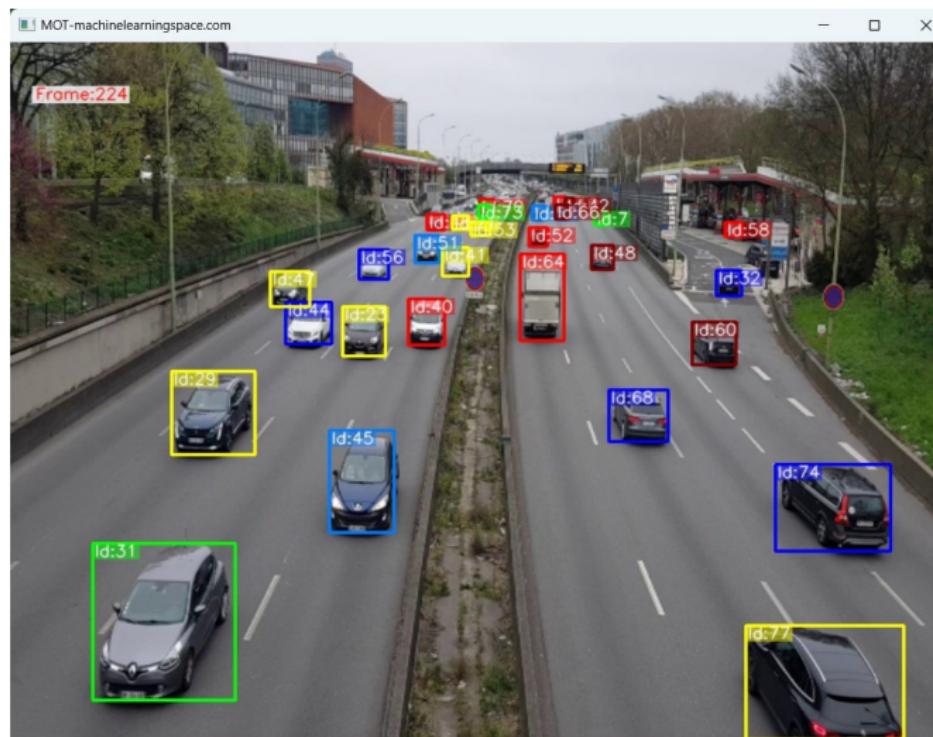
Применения полученного графа

- **поиск изображений:** *найди все картинки, где человек держит зонт.*
- **ответы на вопросы:** *сколько собак сидит на траве?*
- **мультимодальный вывод:** объединение текстовой и визуальной информации (например, анализ сайта).
- **перерисовка:** граф служит чертежом для генеративной модели
 - можно удалить объект/отношение, поменять местами, и модель перерисует картинку
- **задания для роботов:** *возьми чашку со стола и поставь в шкаф.*

Видео

- определения действия (activity recognition)
- генерация видео (video generation):
 - по текстовому описанию (text-to-video generation)
 - предсказание будущих кадров (video prediction)
 - по 1 изображению (image-to-video generation)
- повышение качества видео (video enhancement)
 - повышение разрешения (super-resolution)
 - повышение частоты кадров (frame interpolation)
- 3D-реконструкция по видеосъёмке с разных ракурсов

Трекинг объектов (object tracking)



Содержание

1 Изображения

2 Текст

3 Звук

4 Графы

Задачи на текстах²

- Классификация текста:

- спам/не спам
- положительный/отрицательный отзыв (sentiment analysis)



My experience
so far has been
fantastic!

POSITIVE



The product is
ok I guess

NEUTRAL



Your support team is
useless

NEGATIVE

²Источник

Задачи на текстах

- **Распознавание именованных сущностей (named entity recognition)**

Back in 2000 , People Magazine PUBLISHER highlighted Prince Williams' PERSON style who at the time was a little more fashion-conscious , even making fashion statements at times .

Now-a-days the prince mainly wears navy COLOR suits ITEM (sometimes double-breasted DESIGN) , light blue COLOR button-ups ITEM with classic LOOK pointed DESIGN collars PART , and burgundy COLOR ties ITEM .

But who knows what the future holds ...

Duchess Kate PERSON did wear an Alexander McQueen BRAND dress ITEM to the wedding OCCASION in the fall of 2017 SEASON .

- **Машинный перевод (machine translation):**
 - господа, я не ел шесть дней -> je ne mange pas six jours

Другие задачи

- определение автора по тексту
- суммаризация текста (text summarization)
- информационный поиск (information retrieval)
- ответы на вопросы (question answering)
 - генерация текста по теме (topic-guided text generation)
 - чат-боты (chat-bots)
- выявление тематики текста (topic modeling)
- извлечение связей между объектами: текст->граф знаний
- стилизация текста (text style transfer)
- генерация изображений по тексту (text-to-image)

Содержание

1 Изображения

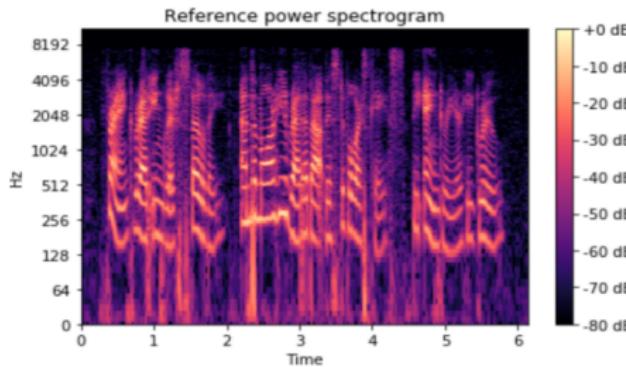
2 Текст

3 Звук

4 Графы

Исходное представление объектов

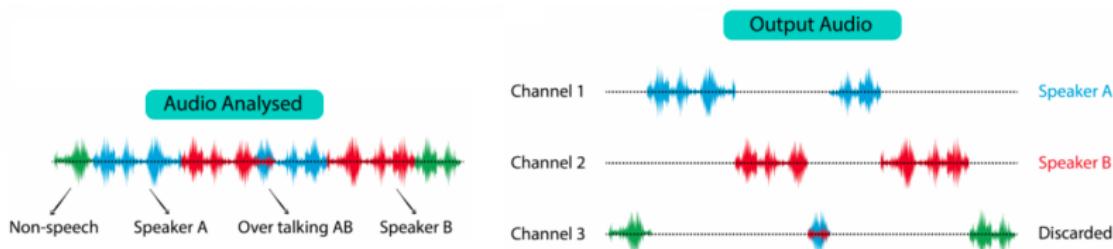
Исходное представление звука (речь, мелодия):



Не обязательно звук - может быть сигнал от любого сенсора.

Задачи

- Речь -> текст, текст -> речь
- Музыка -> распознать композицию (Shazam), стиль, жанр, исполнителя
- Стиль -> генерация музыки.
- Определение #спикеров, выделение временных фрагментов, когда кто говорит.



Содержание

1 Изображения

2 Текст

3 Звук

4 Графы

Источники графовых данных

- **Социальная сеть:** узлы - люди, связи - дружба, писали друг другу, лайкнули посты друг друга.
- **Интернет:** узлы - веб-страницы, связи - ссылки друг на друга.
- **Компьютерная сеть:** узлы - компьютеры, связи - маршруты пересылки сообщений.
- **Химические соединения:** атомы, связанные химической реакцией.
- **Машинное обучение:** узлы-объекты, связи между близкими объектами $\rho(x_i, x_j) \leq threshold$.
- **Транспорт:** узлы - дома, города, связи - транспортное сообщение между ними.

Источники графовых данных

- **Исследования:** узлы - статьи, связи - цитирования внутри статей.
- **Граф знаний:** узлы - сущности (люди, города, организации)+их связи друг с другом.
- **Экономический граф:** узлы - экономические агенты, связи - покупки, продажи.
- **Производственный граф:** узлы-компоненты, связи - что из чего составляется.
- **Логистический граф:**
узлы-производства/склады/магазины.
- **Энергетический граф**
генерации/распределения/потребления энергии.

Задачи на графах

- **Прогноз на узлах графа**
 - классификация вершин - откликнется ли каждый член сообщества соц. сети на предложение
- **Прогноз недостающих ребер**
 - рекомендация друзей соц. сети
- **Поиск подграфов с определёнными свойствами**
 - кластеризация узлов, поиск сообществ
- **Классификация графов**
 - молекула представлена графом, прогнозируем химические свойства
- **Генерация графов**
 - сгенерировать молекулу с заданными химическими свойствами

Подзадачи глубокого обучения

- **Многозадачное обучение** (multi-task learning)
- **Трансферное обучение** (transfer learning)
- **Последовательное обучение** (continuous learning)
- **Мета-обучение** (meta-learning)
 - Few-shot learning, one-shot learning, zero-shot learning.
- **Нейросетевой поиск архитектур** (neural architecture search)
- **Автоматическое машинное обучение** (AutoML)
- **Упрощение модели** (model simplification)
- **Распределённое обучение** (distributed learning)
- **Федеративное обучение** (federated learning):
 - обучение на сервере+пользовательских устройствах
 - на сервере - только агрегированная информация

Типы ИИ

- **Слабый ИИ (narrow AI)**: решение частных формализованных задач ML.
- **Общий ИИ (artificial general intelligence)**: способен решать широкий класс задачу путём самообучения и развития.
 - большие языковые модели (large language models)
- **Сильный ИИ (strong AI)**: способен мыслить и осознавать себя как отдельную личность.
 - пока существует как гипотеза

Литература

- Авторский учебник DeepMachineLearning.ru
- Учебник ШАД.
- Prince S. J. D. Understanding deep learning. – 2023.
- Bishop C. M., Bishop H. Deep learning: Foundations and concepts. – 2023.
- Zhang A. et al. Dive into deep learning. – 2023.
- paperswithcode.com, google scholar.