

Введение в искусственный интеллект. Современное компьютерное зрение

Тема: Введение. Постановка основных задач компьютерного зрения

Бабин Д.Н., Иванов И.Е.

кафедра Математической Теории Интеллектуальных Систем

14 февраля 2023 г.



План лекции

- ① Организационные вопросы
- ② Историческая справка
- ③ Нейросетевой подход
- ④ Постановка основных задач компьютерного зрения
- ⑤ Примеры прикладных задач

Авторы курса



Руководитель курса: д.ф.-м.н. Бабин Дмитрий Николаевич



Лектор: к.ф.-м.н. Иванов Илья Евгеньевич



Лектор: к.ф.-м.н. Петюшко Александр Александрович

Сотрудничество

- Авторы имеют более 15 лет опыта участия в проектах, связанных с машинным обучением и компьютерным зрением
- Являются постоянными участниками группы распознавания образов кафедры МаТИС
- В качестве научных консультантов работают или работали с такими крупнейшими российскими и международными компаниями как Нейроком, LSI Research, Fotonation, Huawei и др.



- Данный курс является частью программы **SHARE**
 - **SHARE** = School of Huawei Advanced Research Education, или Школа опережающего научного образования Хуавэй
 - e-mail: share@intsys.msu.ru
 - Сайт SHARE: <http://sharemsu.ru>
 - Канал SHARE: <https://t.me/joinchat/9IzmCnQIyvs2NjUy>
 - Чат SHARE: <https://t.me/+aM2hI4zw78dkNDFi>



Зачем посещать этот курс

- ❶ Компьютерное зрение — одно из бурно развивающихся направлений в области искусственного интеллекта

Зачем посещать этот курс

- ➊ Компьютерное зрение — одно из бурно развивающихся направлений в области искусственного интеллекта
- ➋ Возможность получить state-of-the-art знания, которые пригодятся в работе

Зачем посещать этот курс

- ➊ Компьютерное зрение — одно из бурно развивающихся направлений в области искусственного интеллекта
- ➋ Возможность получить state-of-the-art знания, которые пригодятся в работе
- ➌ Шанс максимально использовать своё образование

Зачем посещать этот курс

- ❶ Компьютерное зрение — одно из бурно развивающихся направлений в области искусственного интеллекта
- ❷ Возможность получить state-of-the-art знания, которые пригодятся в работе
- ❸ Шанс максимально использовать своё образование
- ❹ Для лучших студентов возможны стажировки и бонусы от партнеров



Что же такое искусственный интеллект?

Естественный интеллект (человек)

- Может воспринимать информацию, ее анализировать, принимать решения на основе анализа



Что же такое искусственный интеллект?

Естественный интеллект (человек)

- Может воспринимать информацию, ее анализировать, принимать решения на основе анализа

Искусственный интеллект

- (Сильный) то же самое, что и естественный, только на месте человека — компьютер



Что же такое искусственный интеллект?

Естественный интеллект (человек)

- Может воспринимать информацию, ее анализировать, принимать решения на основе анализа

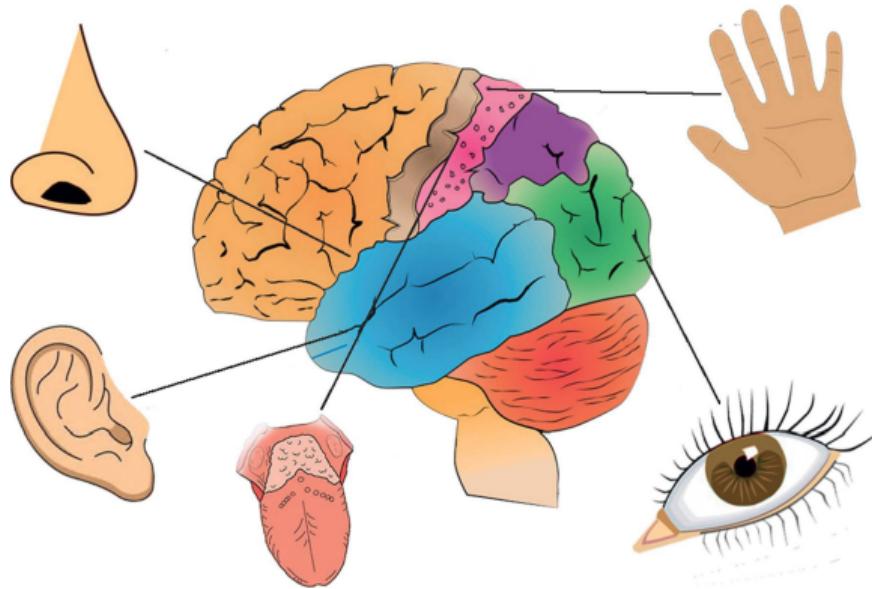
Искусственный интеллект

- (**Сильный**) то же самое, что и естественный, только на месте человека — компьютер
- (**Слабый**) алгоритм, способный обучиться на основе массива входных данных, чтобы затем выполнять задачу вместо человека



Взаимодействие со средой

- Около 90 % информации поступает через **зрение**¹
- Около 9 % информации поступает через слух



¹https://www.rlsnet.ru/books_book_id_2_page_40.htm

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей
- Улучшение качества фотографий

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей
- Улучшение качества фотографий
- Зрение роботов

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей
- Улучшение качества фотографий
- Зрение роботов
- Поиск по фотографиям

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей
- Улучшение качества фотографий
- Зрение роботов
- Поиск по фотографиям
- Дополненная реальность

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей
- Улучшение качества фотографий
- Зрение роботов
- Поиск по фотографиям
- Дополненная реальность
- Автомобили без водителей

Компьютерное зрение в жизни

- Видеонаблюдение: нахождение и слежка за объектами
- Медицина: обнаружение злокачественных опухолей
- Улучшение качества фотографий
- Зрение роботов
- Поиск по фотографиям
- Дополненная реальность
- Автомобили без водителей
- Магазины без продавцов



Что будет в этом курсе

Теоретическая часть

- Основы построения нейронных сетей
 - Свёрточные слои, различные нормализации, механизм внимания
- Обучение нейронных сетей
 - Градиентный спуск, обратное распространение ошибки, инициализация
- Задачи компьютерного зрения
 - Классификация, детекция, сегментация
- Известные сверточные нейросети
 - AlexNet, VGG, ResNet, R-CNN, Mask R-CNN и более современные архитектуры
- Генеративные модели
 - Генеративные состязательные сети, (вариационный) автоэнкодер
- Ускорение нейронных сетей
 - Квантование, прунинг, дистилляция и др.

Что будет в этом курсе

Практическая часть

- Обработка изображений и нейросетевые фреймворки
 - Scikit-Image, opencv
 - PyTorch
- Реализация своего фреймворка обучения нейронных сетей
- Работа с данными: аугментации
- Соревнования по компьютерному зрению
- Исследовательские задания



Чего не будет в этом курсе

- Теория обработки изображений²
 - Преобразование Фурье, свертки

²См. “Математические основы обработки изображений”, Мазуренко И. Л.



Чего не будет в этом курсе

- Теория обработки изображений²
 - Преобразование Фурье, свертки
- Классическое компьютерное зрение
 - Теория фильтрации

²См. “Математические основы обработки изображений”, Мазуренко И. Л.



Оценки за курс

- Оценки за курс будут выставляться в соответствии с данными о посещении и набранными баллами за выполнение домашних заданий.



Оценки за курс

- Оценки за курс будут выставляться в соответствии с данными о посещении и набранными баллами за выполнение домашних заданий.
- В ходе курса будут предложены домашние задания трёх типов:
 - теоретические
 - практические
 - соревнования

Оценки за курс

- Оценки за курс будут выставляться в соответствии с данными о посещении и набранными баллами за выполнение домашних заданий.
- В ходе курса будут предложены домашние задания трёх типов:
 - теоретические
 - практические
 - соревнования
- В конце семестра состоится экзамен, на котором при желании можно будет повысить свою оценку

Оценки за курс

- Оценки за курс будут выставляться в соответствии с данными о посещении и набранными баллами за выполнение домашних заданий.
- В ходе курса будут предложены домашние задания трёх типов:
 - теоретические
 - практические
 - соревнования
- В конце семестра состоится экзамен, на котором при желании можно будет повысить свою оценку
- Предварительная шкала оценок:

Оценка	Процент выполненных заданий
Отлично	80 %
Хорошо	60 %
Зачет	40 %

Кодекс чести SHARE

- Спisyвать (у других студентов) категорически запрещается!

Кодекс чести SHARE

- Спisyвать (у других студентов) категорически запрещается!
- При подозрении на списанную работу ставится 0 баллов:
 - Списавшему
 - Давшему списать



Кодекс чести SHARE

- Спisyвать (у других студентов) категорически запрещается!
- При подозрении на списанную работу ставится 0 баллов:
 - Списавшему
 - Давшему списать
- При использовании дополнительных источников (ресурсы в Интернете, учебники) обязательно ссылаться на них

Полезные ресурсы

- Страница курса: <https://github.com/mlcoursemm/cv2023spring>
- Главный ресурс по курсам “Введение в компьютерный интеллект”:
<https://github.com/mlcoursemm>
- Почта курса: mlcoursemm@gmail.com

Нейросетевой подход

- На данный момент подавляющее большинство алгоритмов при работе с компьютерным зрением использует нейронные сети

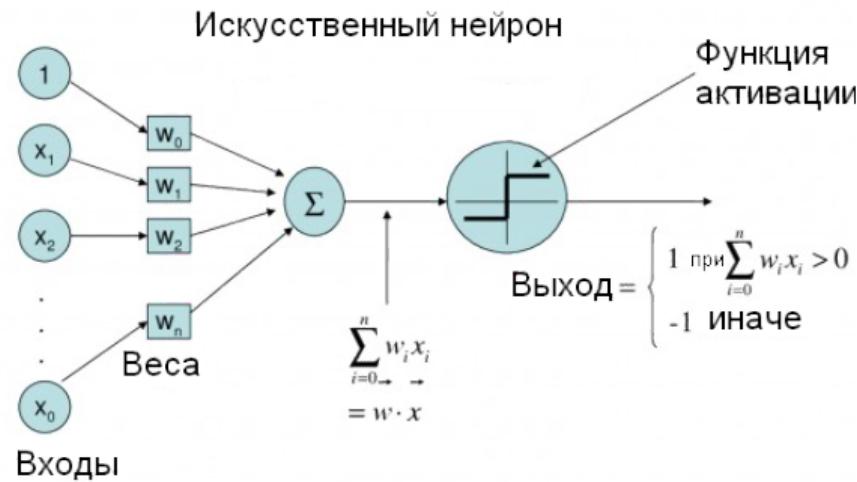
Нейросетевой подход

- На данный момент подавляющее большинство алгоритмов при работе с компьютерным зрением использует нейронные сети
- Простейшая нейронная сеть — однослойный перцептрон — состоит из линейного слоя и следующей за ним нелинейной функции



Нейросетевой подход

- На данный момент подавляющее большинство алгоритмов при работе с компьютерным зрением использует нейронные сети
- Простейшая нейронная сеть — однослойный перцептрон — состоит из линейного слоя и следующей за ним нелинейной функции
- В 1943 г. Маккалок и Питтс формализуют понятие нейрона



- 1949 — Хебб предлагает первый алгоритм обучения перцептрана

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)
- 1975–1980 — Фукусима предлагает мощную модель когнитрана и неокогнитрана (прототип сверточных сетей)

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)
- 1975–1980 — Фукусима предлагает мощную модель когнитрона и неокогнитрона (прототип сверточных сетей)
- 1989 — Лекун представляет первую работу по объединению свёрточных нейросетей и метода обратного распространения ошибки

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)
- 1975–1980 — Фукусима предлагает мощную модель когнитрана и неокогнитрана (прототип сверточных сетей)
- 1989 — Лекун представляет первую работу по объединению свёрточных нейросетей и метода обратного распространения ошибки
- 2012 — Крижевский и Хинтон публикуют первую успешную глубокую классификационную нейросеть AlexNet

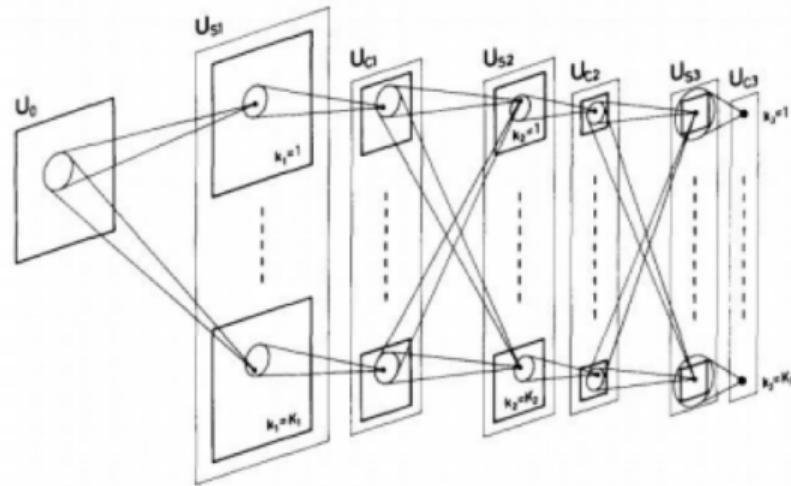
Неокогнитрон

- Фукусима предложил практически современный метод построения архитектуры нейросетей, заимствованный им из модели первичной зрительной коры

- Фукусима предложил практически современный метод построения архитектуры нейросетей, заимствованный им из модели первичной зрительной коры
- Два вида нейронов:
 - Простые (Simple), отвечающие за локальные признаки
 - Сложные (Complex), отвечающие за компенсацию искажения
 - Организованы в каскадную структуру S-C-S-C-S-C...
 - В сверточной сети S=свертка, C=субдискретизация

Неокогнитрон

- Фукусима предложил практически современный метод построения архитектуры нейросетей, заимствованный им из модели первичной зрительной коры
- Два вида нейронов:
 - Простые (Simple), отвечающие за локальные признаки
 - Сложные (Complex), отвечающие за компенсацию искажения
 - Организованы в каскадную структуру SCSCSC...
 - В сверточной сети S=свертка, C=субдискретизация



- Главный минус: не было предложено метода обратного распространения ошибки для обучения

Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году

³<https://arxiv.org>

⁴<https://github.com>

Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:

³<https://arxiv.org>

⁴<https://github.com>



Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:
 - Появились большие наборы данных для обучения ($\sim 10^6$ изображений)

³<https://arxiv.org>

⁴<https://github.com>

Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:
 - Появились большие наборы данных для обучения ($\sim 10^6$ изображений)
 - Появились мощные видеокарты для обучения нейронных сетей

³<https://arxiv.org>

⁴<https://github.com>



Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:
 - Появились большие наборы данных для обучения ($\sim 10^6$ изображений)
 - Появились мощные видеокарты для обучения нейронных сетей
 - Открытость исследований (arxiv³, github⁴)

³<https://arxiv.org>

⁴<https://github.com>

Ресурсы для работы с нейросетями

- На данный момент обучение нейросетей немыслимо без использования **GPU** (Graphical Processing Unit) от NVIDIA, **TPU** (Tensor Processing Unit) от Google, AI Chip (Da Vinci) от Huawei и пр.

⁵<https://colab.research.google.com>

⁶<https://aws.amazon.com/ru/ec2/>

Ресурсы для работы с нейросетями

- На данный момент обучение нейросетей немыслимо без использования **GPU** (Graphical Processing Unit) от NVIDIA, **TPU** (Tensor Processing Unit) от Google, AI Chip (Da Vinci) от Huawei и пр.
- Для обучения промышленных нейросетей (например, Mask R-CNN) необходим объем памяти GPU от 12-16 ГБ

⁵<https://colab.research.google.com>

⁶<https://aws.amazon.com/ru/ec2/>

Ресурсы для работы с нейросетями

- На данный момент обучение нейросетей немыслимо без использования **GPU** (Graphical Processing Unit) от NVIDIA, **TPU** (Tensor Processing Unit) от Google, AI Chip (Da Vinci) от Huawei и пр.
- Для обучения промышленных нейросетей (например, Mask R-CNN) необходим объем памяти GPU от 12-16 ГБ
 - В силу этого подходящие карты на данный момент: Titan X / V / RTX, GTX 1080 Ti, RTX 2080 Ti / 3080 / 3090, P100, V100, A100 (и более мощные)
 - Для любого проекта необходимо обеспечить доступ к данному оборудованию (например, удаленно через ssh)
 - Либо воспользоваться облачными сервисами (например, через Google Colaboratory⁵ либо Amazon Web Services⁶)

⁵<https://colab.research.google.com>

⁶<https://aws.amazon.com/ru/ec2/>

Задача классификации

- **Задача:** отнести входное изображение к одному из классов
 - По фотографии человека определить пол, возраст и т. д.
 - По фотографии автомобиля определить марку и модель

Задача классификации

- **Задача:** отнести входное изображение к одному из классов
 - По фотографии человека определить пол, возраст и т. д.
 - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
 - Решается с помощью сверточных нейросетей (СНС)

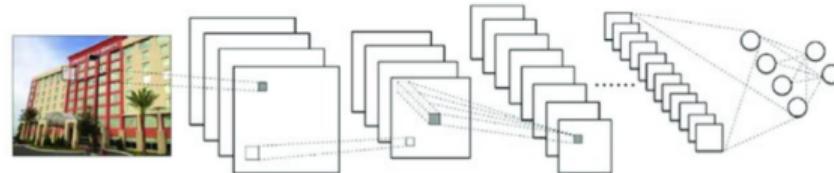
Задача классификации

- **Задача:** отнести входное изображение к одному из классов
 - По фотографии человека определить пол, возраст и т. д.
 - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
 - Решается с помощью сверточных нейросетей (СНС)
- **Проблема:** двусмысленность изображений
 - Не любое изображение можно описать одним классом

Задача классификации

- **Задача:** отнести входное изображение к одному из классов
 - По фотографии человека определить пол, возраст и т. д.
 - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
 - Решается с помощью сверточных нейросетей (СНС)
- **Проблема:** двусмысленность изображений
 - Не любое изображение можно описать одним классом

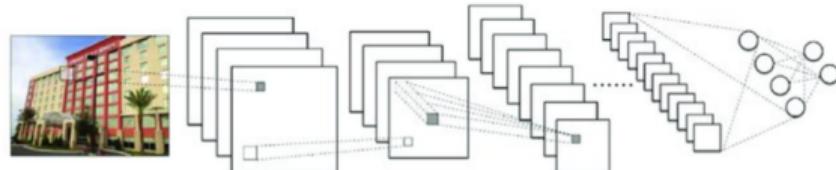
Пример функционирования СНС



Задача классификации

- **Задача:** отнести входное изображение к одному из классов
 - По фотографии человека определить пол, возраст и т. д.
 - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
 - Решается с помощью сверточных нейросетей (СНС)
- **Проблема:** двусмысленность изображений
 - Не любое изображение можно описать одним классом

Пример функционирования СНС



Основные представители:

- 2012 — AlexNet
- 2014 — VGGNet
- 2014 — Inception
- 2015 — ResNet
- 2016 — ResNeXt
- 2017 — MobileNet
- 2017 — (P)NASNet
- 2019 — EfficientNet
- 2020 — ViT
- 2021 — Swin Transformer
- 2022 — ConvNeXts
- 2023 — ???



Проблемная классификация

Что изображено: лошадь или женщина?



Проблемная классификация

Что изображено: лошадь или женщина?



Правильный ответ: сено

Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником



Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект

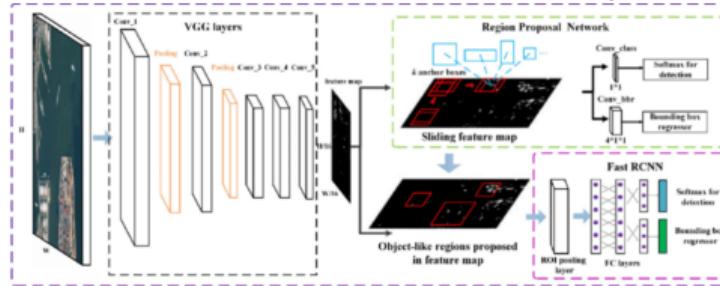
Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект
- Существуют алгоритмы, которые решают эту задачу в т. ч. в реальном времени
 - Faster R-CNN, YOLO, SSD

Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект
- Существуют алгоритмы, которые решают эту задачу в т. ч. в реальном времени
 - Faster R-CNN, YOLO, SSD

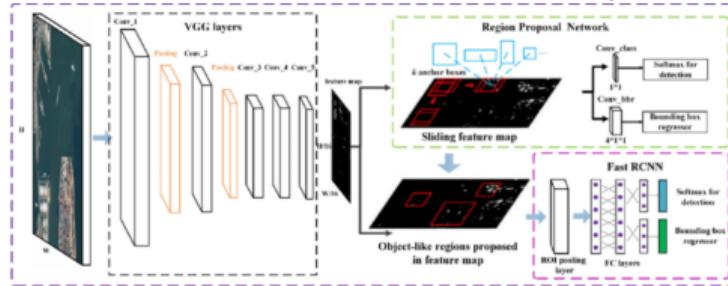
Пример функционирования детектора (Faster R-CNN)



Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект
- Существуют алгоритмы, которые решают эту задачу в т. ч. в реальном времени
 - Faster R-CNN, YOLO, SSD

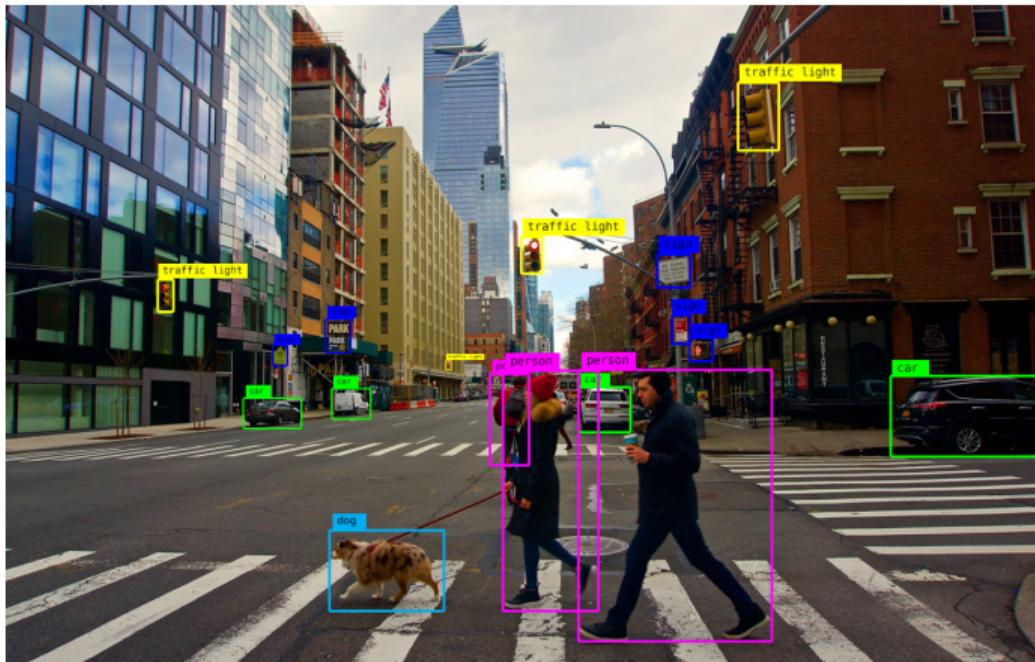
Пример функционирования детектора (Faster R-CNN)



Основные представители:

- 2013 — Overfeat
- 2014 — R-CNN
- 2015 — Faster R-CNN
- 2015 — YOLO
- 2015 — SSD
- 2016 — R-FCN
- 2017 — FPN
- 2017 — RetinaNet
- 2019 — CenterNet
- 2019 — EfficientDet
- 2020 — DETR

Пример работы детектора



Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта

Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку

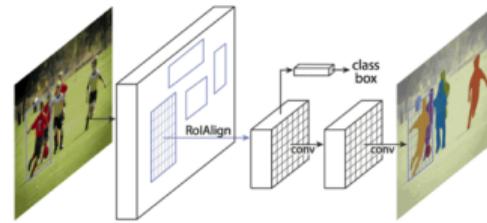
Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку
- На данный момент существуют алгоритмы-комбайны (решающие одновременно задачи классификации, детекции и сегментации)

Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку
- На данный момент существуют алгоритмы-комбайны (решающие одновременно задачи классификации, детекции и сегментации)

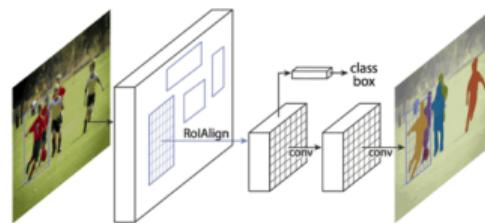
Пример функционирования решения-комбайна (Mask R-CNN)



Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку
- На данный момент существуют алгоритмы-комбайны (решающие одновременно задачи классификации, детекции и сегментации)

Пример функционирования решения-комбайна (Mask R-CNN)



Основные представители:

- 2014 — FCN
- 2015 — U-Net
- 2015 — SegNet
- 2016 — RefineNet
- 2016 — PSPNet
- 2017 — Mask R-CNN
- 2018 — DeepLabv3+
- 2020 — ResNeSt
- 2020 — DetectoRS
- 2021 — SETR

Пример работы решения-комбайна



Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра



Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео



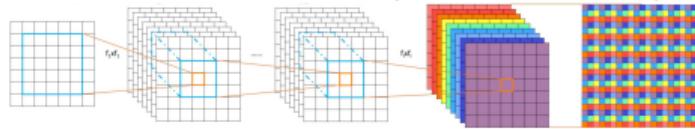
Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео
- **Цель:** улучшить классические алгоритмы интерполяции

Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео
- **Цель:** улучшить классические алгоритмы интерполяции

Пример функционирования алгоритма сверхразрешения (ESPCN, 2016)

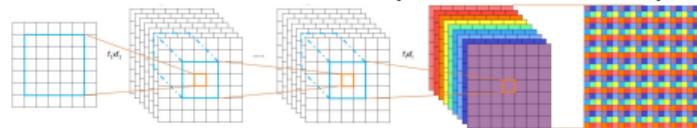


Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео
- **Цель:** улучшить классические алгоритмы интерполяции

Метод сверхразрешения	PSNR
Бикубическая интерполяция	24.32 dB
Бикубическая интерполяция + GIMP	24.69 dB
Бикубическая интерполяция + Photoshop	24.72 dB
<i>Нейро сверхразрешение</i>	<i>27.20 dB</i>

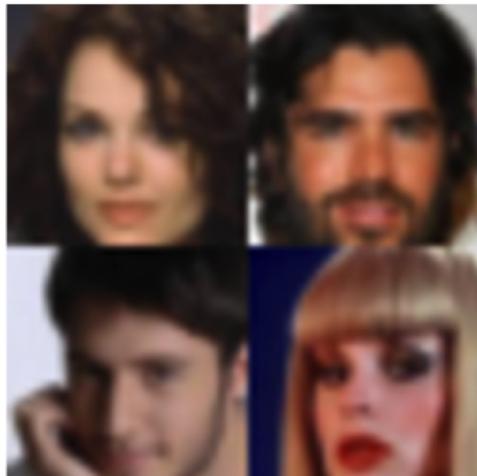
Пример функционирования алгоритма сверхразрешения (ESPCN, 2016)



Задачи улучшения изображений

Сверхразрешение для лиц — пример

Бикубическая интерполяция: PSNR = 24.32



Нейросверх-разрешение: PSNR = 27.20



Малоразмерный вход

Задачи улучшения изображений

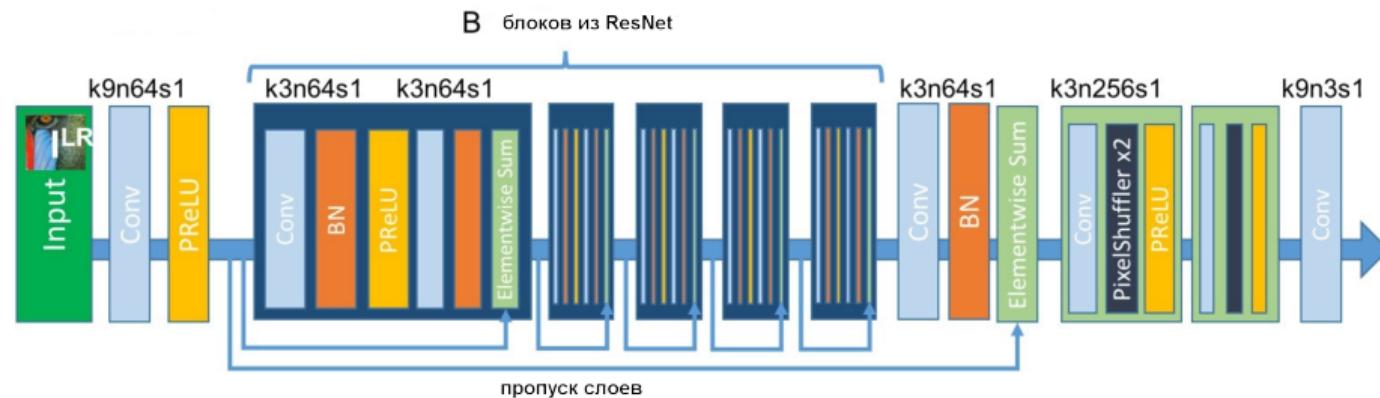
- **Задача:** улучшить качество дефектных изображений
 - Устранить смаз (motion blur)
 - Устранить размытие (defocus blur)
 - Удалить артефакты, являющиеся результатом попытки устраниить дефекты зашитым в устройстве алгоритмом (image reconstruction)

Задачи улучшения изображений

- **Задача:** улучшить качество дефектных изображений

- Устранить смаз (motion blur)
- Устранить размытие (defocus blur)
- Удалить артефакты, являющиеся результатом попытки устраниить дефекты зашитым в устройстве алгоритмом (image reconstruction)

Пример функционирования алгоритма деблюра (SRGAN, 2016)



Задачи улучшения изображений

Исходное
размытие



Деблор для лиц — пример
Классический
Wiener



Коммерческое
решение



Нейросетевое
решение



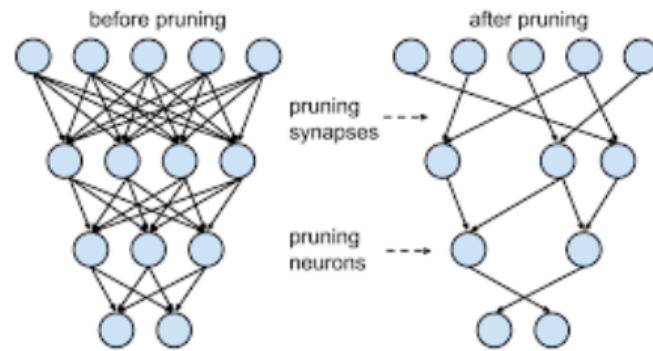
Задачи ускорения нейронных сетей и оптимизации "под железо"

Проблема

Большинство SoTA моделей практически нельзя использовать в реальных проектах

Решение

На практике нейронные сети можно существенно уменьшить, если увеличить допустимый уровень ошибки.



Распознавание автомобильных номеров

- **Задача:** по видеопотоку детектировать транспортное средство и распознать госномер

Распознавание автомобильных номеров

- **Задача:** по видеопотоку детектировать транспортное средство и распознать госномер
- Зачастую итоговое решение состоит из следующих модулей:
 - Детектирование автомобиля
 - Распознавание марки и модели (опционально)
 - Трекинг автомобиля
 - Детектирование номера
 - Распознавание номера
 - Объединение результатов с разных кадров / поиск ключевых кадров
 - Оптимизация моделей и портирование на вычислитель

Распознавание автомобильных номеров

- **Задача:** по видеопотоку детектировать транспортное средство и распознать госномер
- Зачастую итоговое решение состоит из следующих модулей:
 - Детектирование автомобиля
 - Распознавание марки и модели (опционально)
 - Трекинг автомобиля
 - Детектирование номера
 - Распознавание номера
 - Объединение результатов с разных кадров / поиск ключевых кадров
 - Оптимизация моделей и портирование на вычислитель



Примеры прикладных задач

Ре-идентификация пешеходов

- **Задача:** идентифицировать пешеходов на камерах наблюдения



Ре-идентификация пешеходов

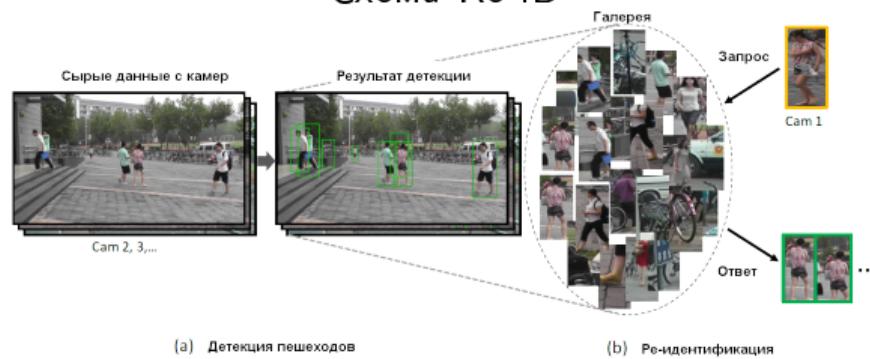
- **Задача:** идентифицировать пешеходов на камерах наблюдения
- **Метод решения:**
 - Детектирование пешеходов
 - Трекинг пешеходов
 - Извлечение признаков
 - Поиск в пространстве признаков

Примеры прикладных задач

Ре-идентификация пешеходов

- **Задача:** идентифицировать пешеходов на камерах наблюдения
- **Метод решения:**
 - Детектирование пешеходов
 - Трекинг пешеходов
 - Извлечение признаков
 - Поиск в пространстве признаков

Схема Re-ID



Распознавание лиц

- **Задача:** по входному изображению лица распознать человека. Есть два подтипа задач:
 - **Верификация** — принадлежат ли две входные картинки одному человеку?
 - **Идентификация** — нахождение в базе данных человека, фотография которого подана на вход

Распознавание лиц

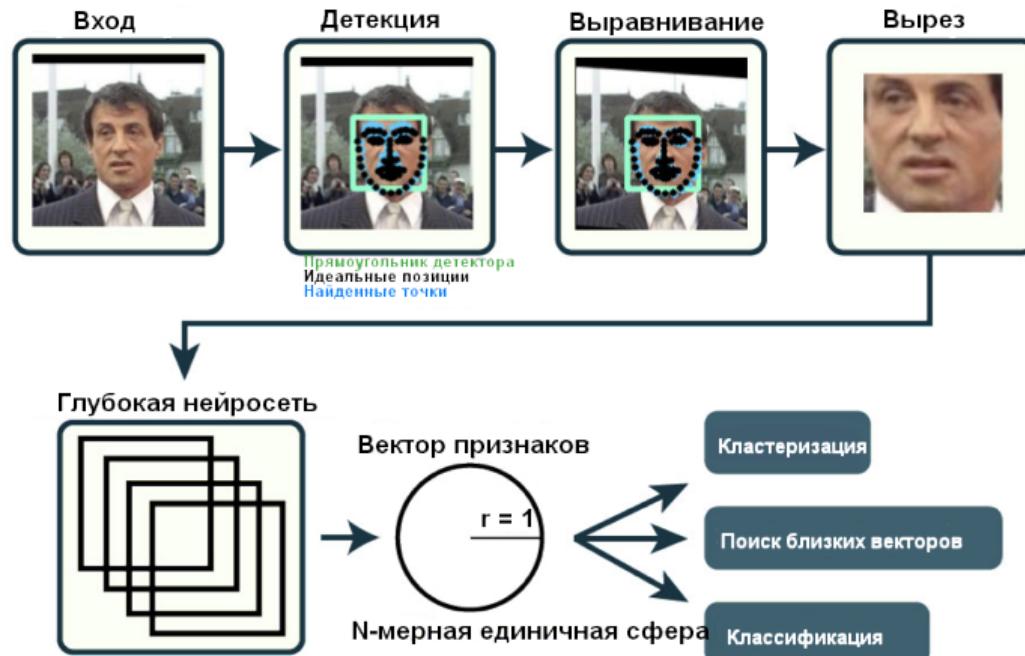
- **Задача:** по входному изображению лица распознать человека. Есть два подтипа задач:
 - **Верификация** — принадлежат ли две входные картинки одному человеку?
 - **Идентификация** — нахождение в базе данных человека, фотография которого подана на вход
- **Метод решения:** с помощью каскада СНС

Распознавание лиц

- **Задача:** по входному изображению лица распознать человека. Есть два подтипа задач:
 - **Верификация** — принадлежат ли две входные картинки одному человеку?
 - **Идентификация** — нахождение в базе данных человека, фотография которого подана на вход
- **Метод решения:** с помощью каскада СНС
- Главная особенность — нахождение N -мерного вектора признаков, соответствующего данному человеку

Примеры прикладных задач

Типичная система распознавания лиц:



Синтез (создание) лиц

- **Задача:** сгенерировать набор синтетических лиц, неотличимых от реальных фотографий
 - Зачастую нужно для расширения обучающей выборки для задачи распознавания лиц
 - Необходимо генерировать как можно больше разных личностей

Синтез (создание) лиц

- **Задача:** сгенерировать набор синтетических лиц, неотличимых от реальных фотографий
 - Зачастую нужно для расширения обучающей выборки для задачи распознавания лиц
 - Необходимо генерировать как можно больше разных личностей
- **Метод решения:** с помощью специальных генеративных состязательных нейросетей (GAN)

Примеры прикладных задач

Какое из этих лиц — синтетическое⁷?



⁷<https://thispersondoesnotexist.com>

Время для вопросов



Спасибо за внимание!

