

Компьютерное Зрение
Лекция №0, осень 2021

Разработчик систем технического зрения –
профессия настоящего и будущего



Программа курса

1. Формирование изображений. Основные понятия
2. Введение в обработку сигналов и изображений
3. Глобальные и локальные признаки изображений
4. Параметрические модели
5. Задачи классификации, сегментации, обнаружения объектов
6. Обработка оптических потоков

Цели лекции

1. Приоткрыть область компьютерного зрения
2. Провести обзор основных задач и методов их решения

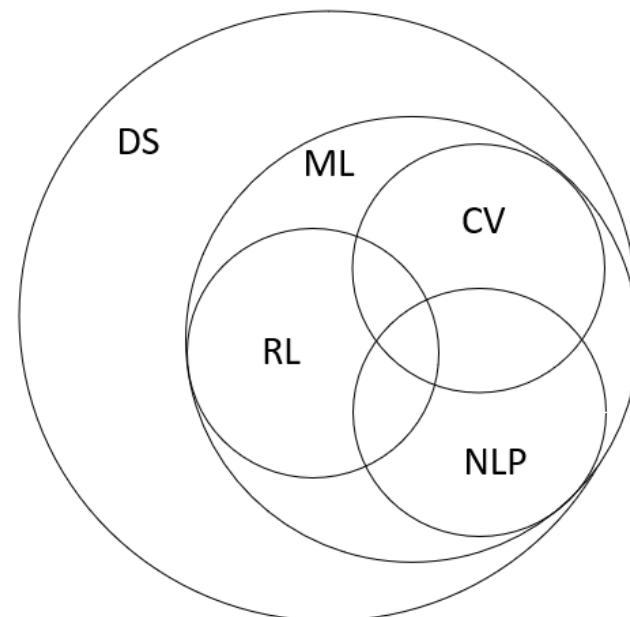
План лекции

1. Основные направления CV
2. Примеры решаемых задач
3. CV Pipeline
4. Заключение
5. Q&A

Введение

Три основных области исследований в ML (Machine learning)

1. CV (Computer Vision)
2. NLP (Natural Language Processing)
3. RL (Reinforcement Learning)



Основные направления в CV

- Сигнализация объектов
- Классификация
- Сегментация
- Детекция
- Multi-view geometry
- Трекинг и видеоаналитика
- Биометрия
- Оценка позы и взгляда человека
- Оценка качества изображений
- Улучшение и восстановление изображений
- Построение карты глубины

Основные направления в CV

Классификация

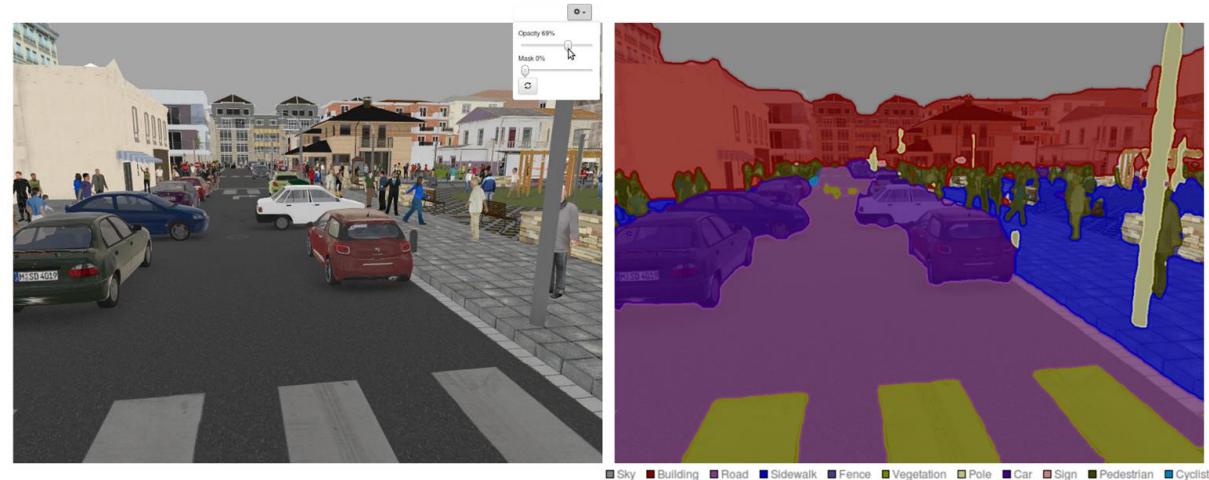
Решается задача
принадлежности
изображения одному из
классов



Основные направления в CV

Сегментация

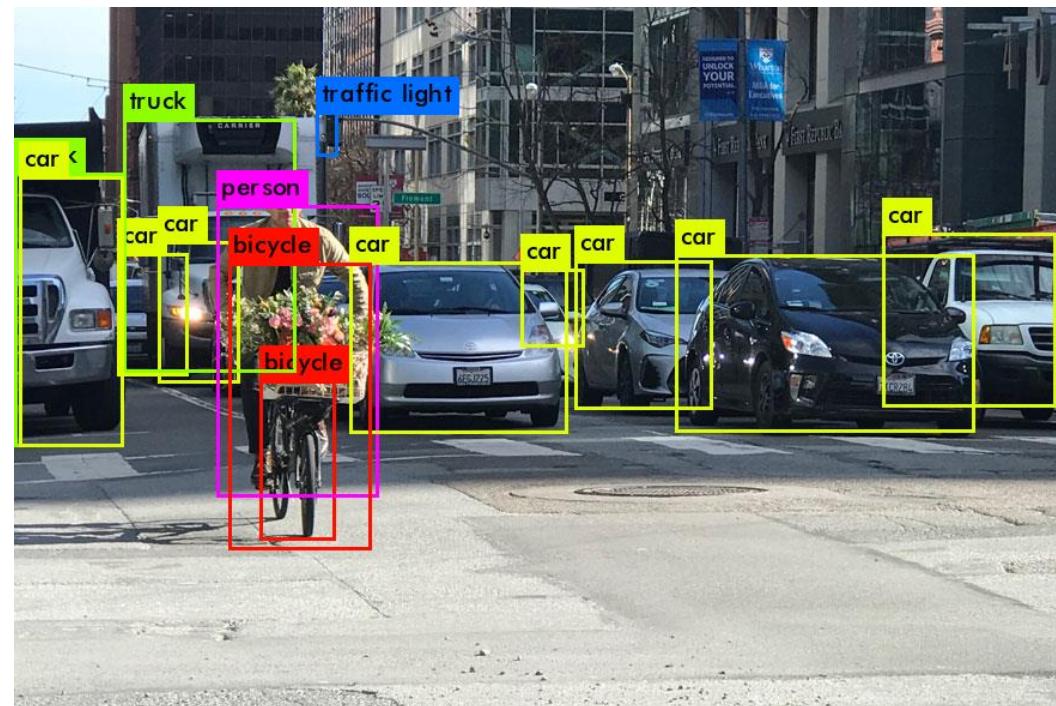
Решается задача классификации, но попиксельной



Основные направления в CV

Детекция

Решается задача
распознавания - найти и
классифицировать



Основные направления в CV

Сигнализация объектов

Есть ли кошечка на картинке?



Основные направления в CV

Трекинг и видео-аналитика

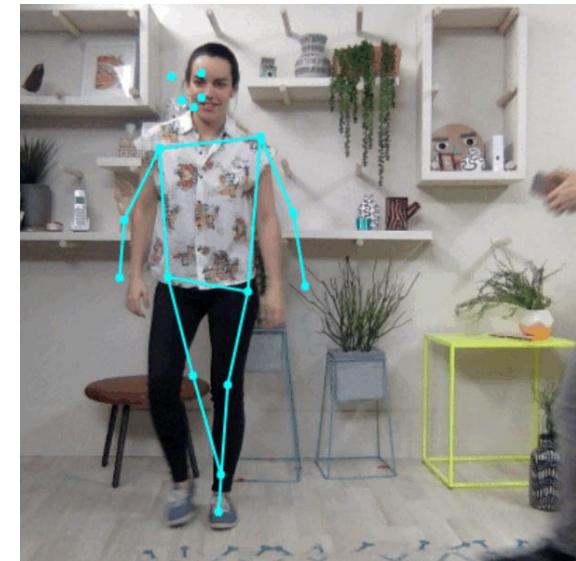
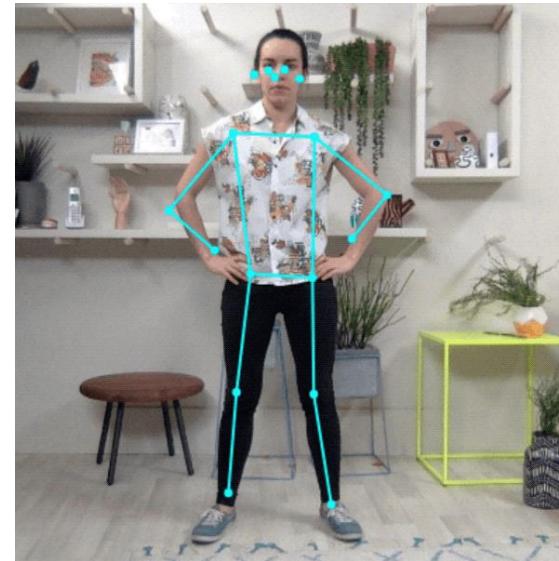
Распознавание объектов в видеопотоке



Основные направления в CV

Оценка позы и взгляда человека

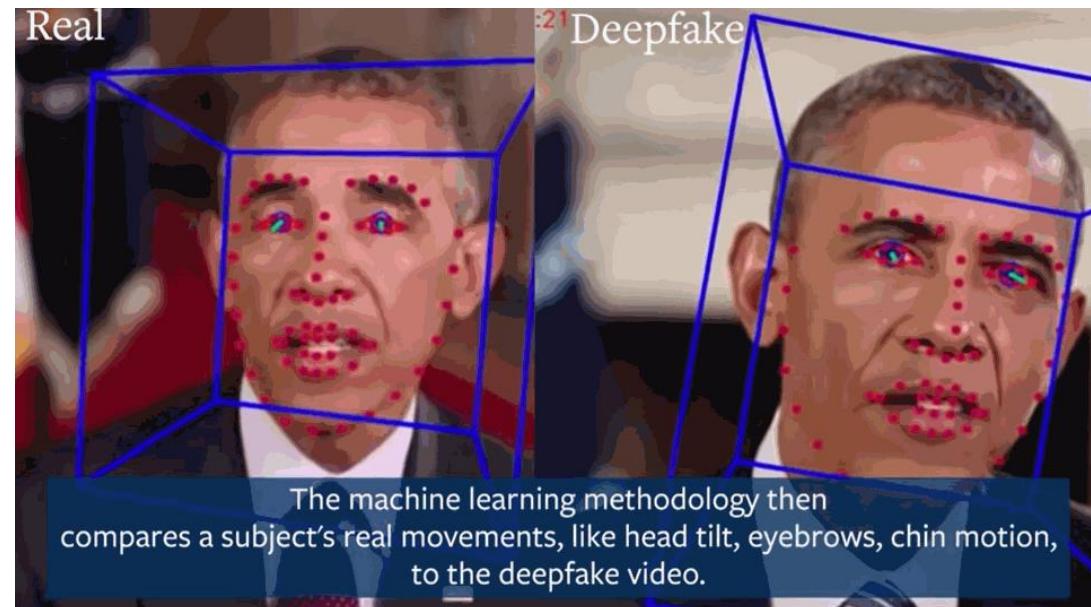
Решается задача
распознавания ключевых
точек, которые описывают
позы человека, его
положение, ориентацию в
пространстве



Основные направления в CV

Биометрия

Решается задача классификации на изображении или видеопотоке: фейк или нет



Основные направления в CV

Оценка качества изображений

Задача в том, чтобы
определить визуальное
качество изображения



Основные направления в CV

Улучшение и восстановление изображений

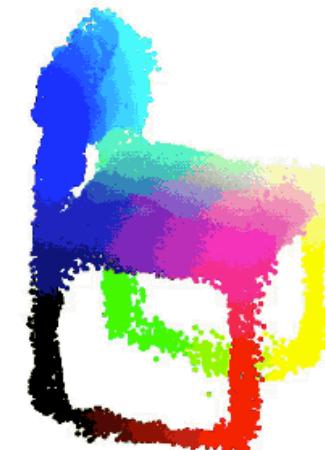
Поиск дефектов и
коррекция изображения



Основные направления в CV

Построение 3D Объектов

Решается задача
построения объемного
объекта (obj-файл) по
двумерному изображению



Основные направления в CV

Построение 3D Объектов

Решается задача
построения объемного
объекта (obj-файл) по ряду
двумерных изображений



Основные направления в CV

Построение карты глубины

Решается задача
построения карты
расстояний до
наблюдаемых объектов



Почему востребовано Кейс - Умный магазин

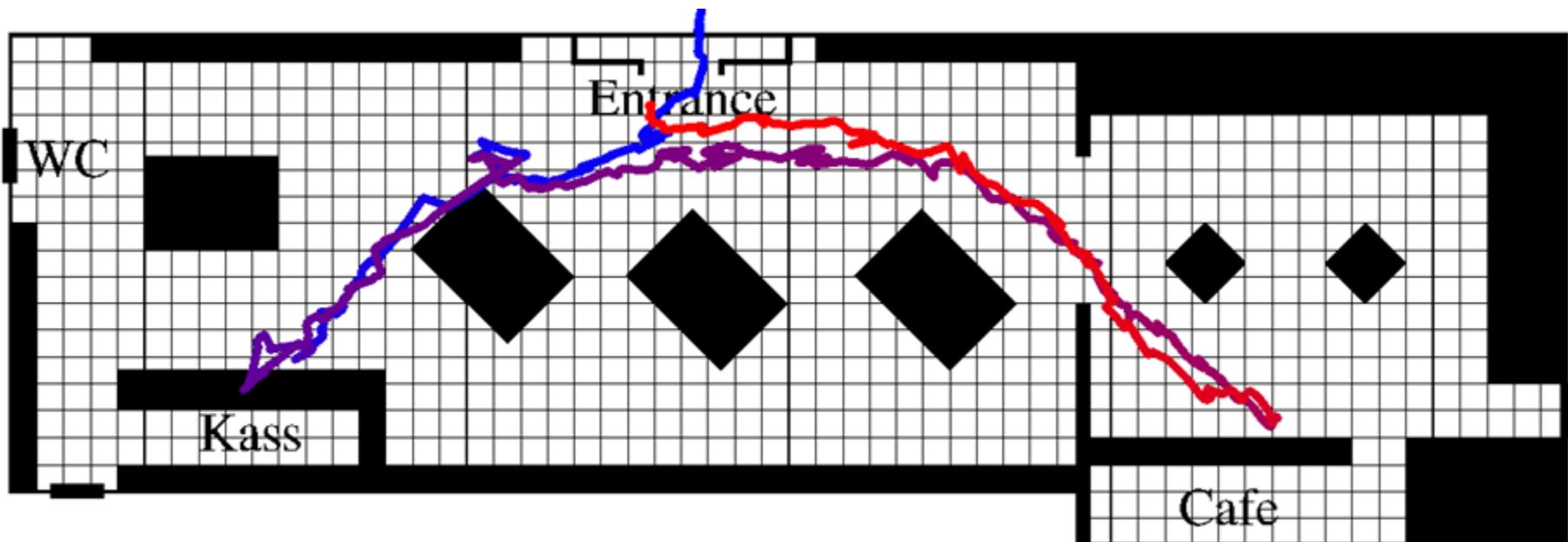


Области применения

- Трекинг по сети камер
- Подсчет количества уникальных людей
- Построение тепловых карт трафика

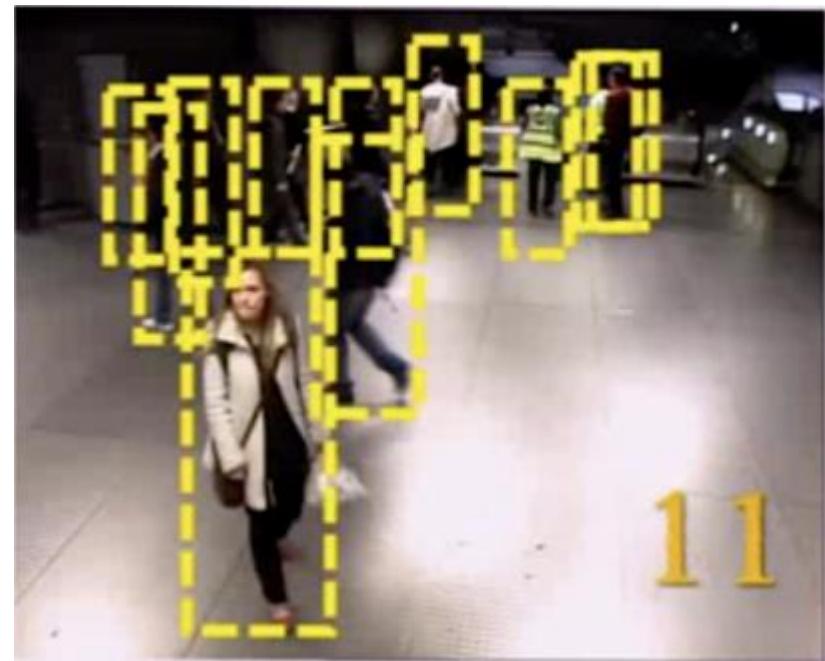


Траектория движения на карте



Области применения

- Трекинг по сети камер
- Подсчет количества уникальных людей
- Построение тепловых карт трафика



Области применения

- Трекинг по сети камер
- Подсчет количества уникальных людей
- Построение тепловых карт трафика



Построение тепловых карт трафика



Доступность товара на полках



Будущее уже близко



Почему востребовано Кейс - Автономные системы

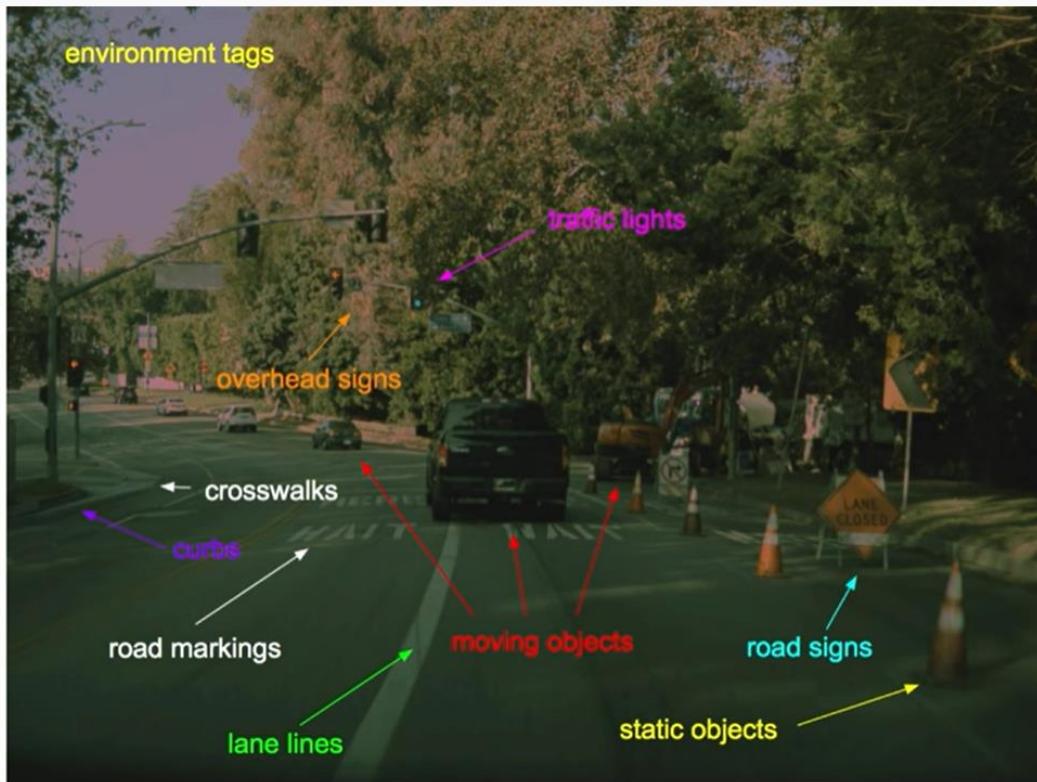
Автономные автомобили и роботы

Беспилотные летательные аппараты
(БПЛА)

Автономные подводные необитаемые
аппараты (АНПА)



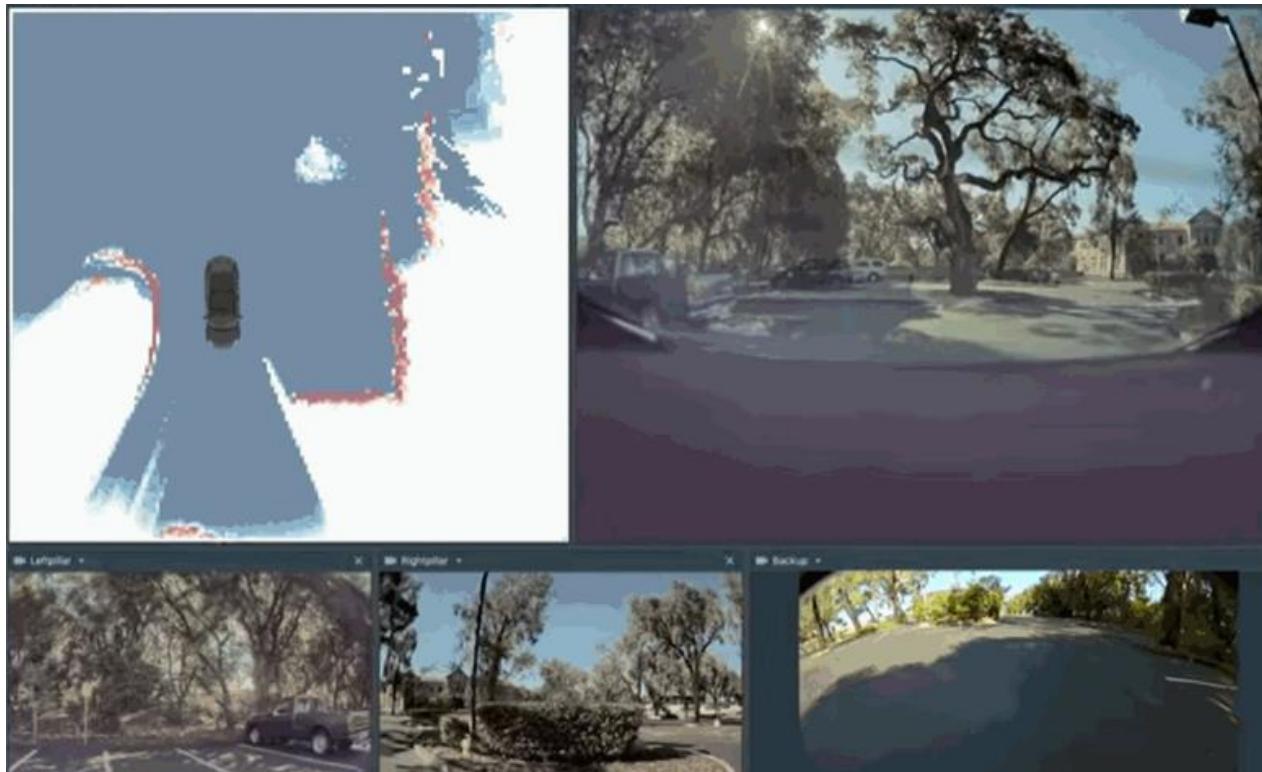
Беспилотный автомобиль



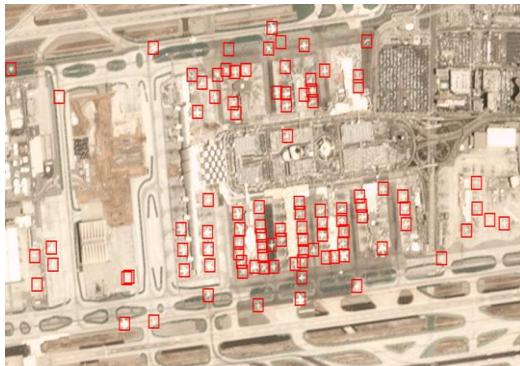
Задачи СВ для автономного управления автомобилем:

- распознавание ключевых объектов в пространстве
- построение карты местности
- ориентация в пространстве

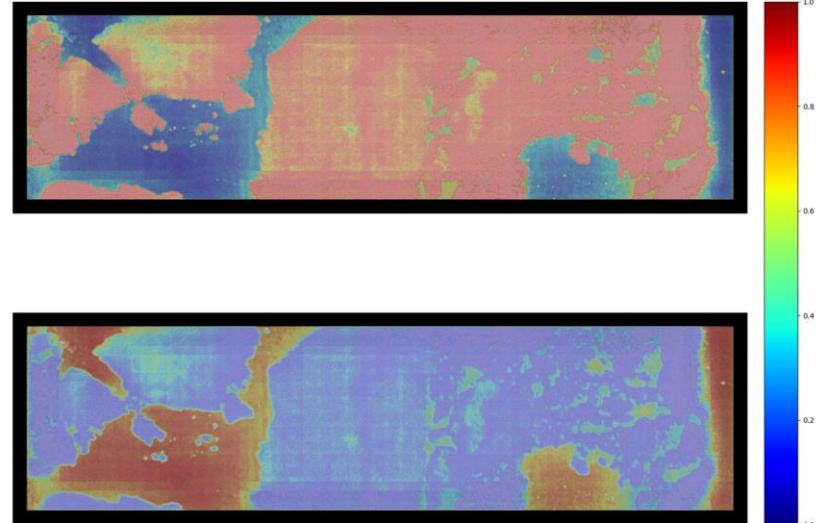
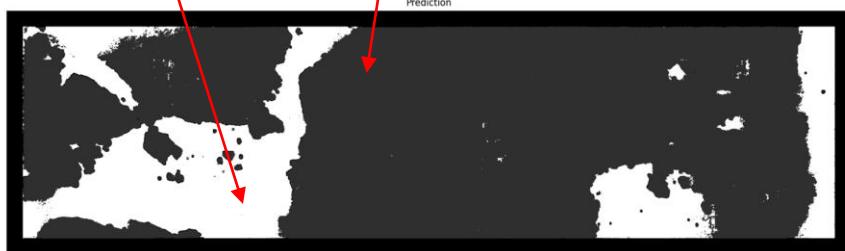
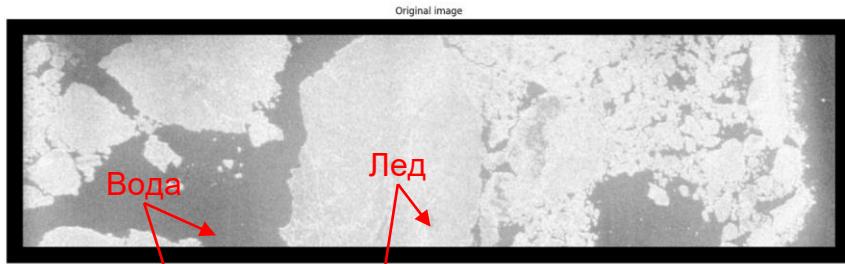
Задача анализа окружающей обстановки



Распознавание в оптическом диапазоне



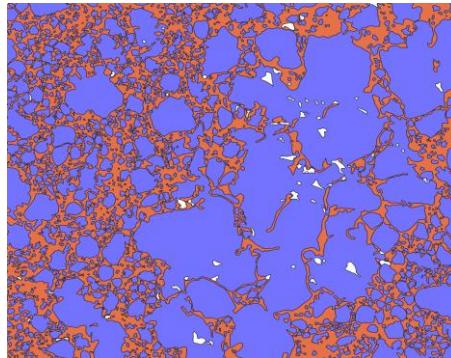
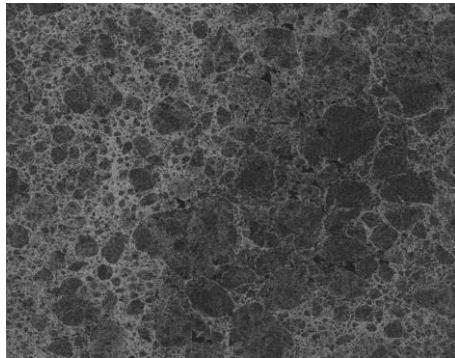
Распознавание типа подстилающей поверхности «Лед-вода»



Примеры распознавания подстилающей поверхности лед – вода на РЛИ С-диапазона

Тепловая карта распределения “уверенности” предсказания

Распознавание типа подстилающей поверхности «Мультикласс»



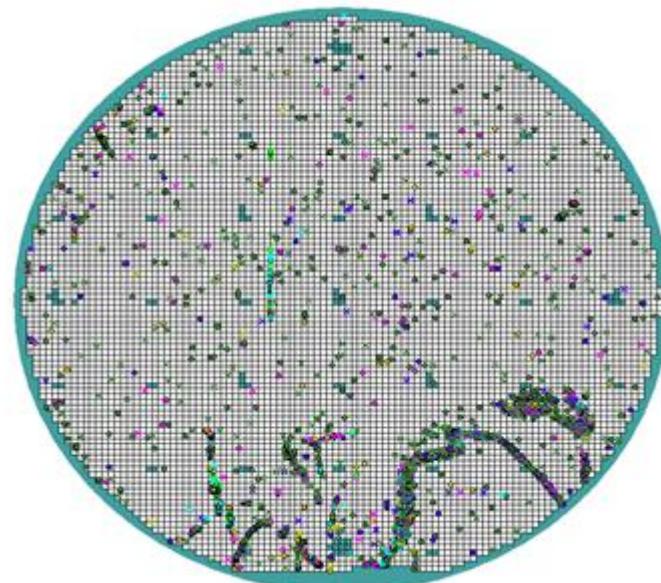
- █ – серый лед
- █ – тонкий вода
- █ – смешанный тип «лед-вода»
- █ – молодой (переходный) лед
- █ – детальная вода
- █ – вода

Почему востребовано

Кейс - Контроль качества продукции

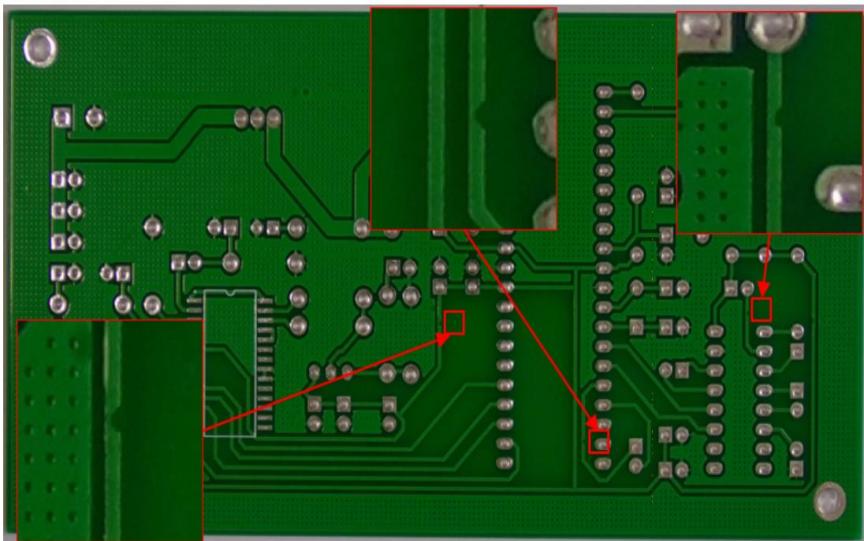


Электронный дефектоскоп DragonFly G2
фирмы «Rudolph Technologies»

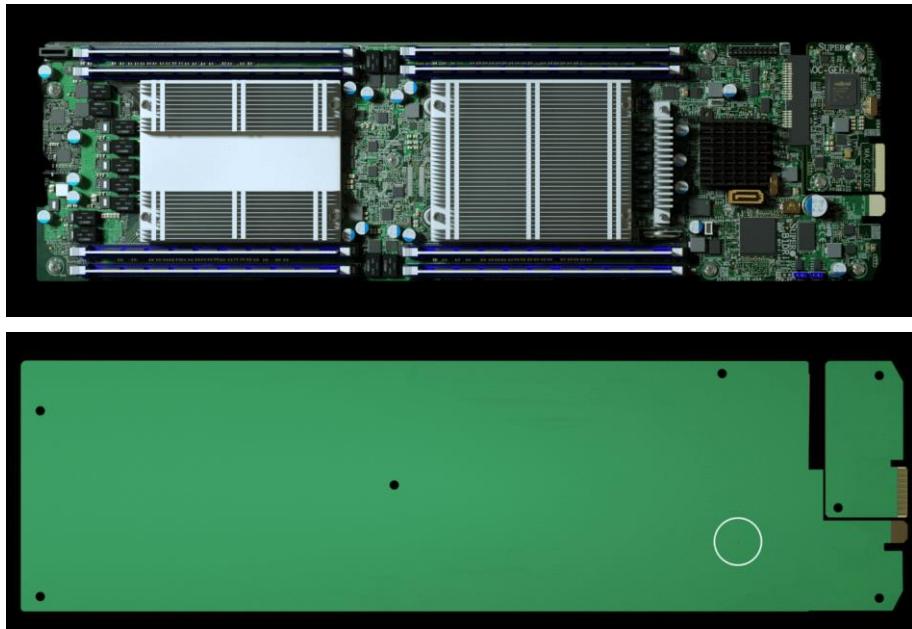


Результаты дефектоскопии микросхем
серии 3800AC

Обнаружение сигнатур дефектов по данным оптического микроскопа



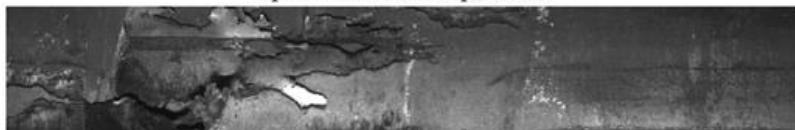
Выявление дефектов на печатных платах



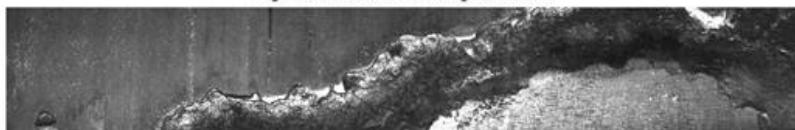
Выявление устройств перехвата данных
на печатных платах

Дефекты стального проката

Оригинальное изображение



Оригинальное изображение



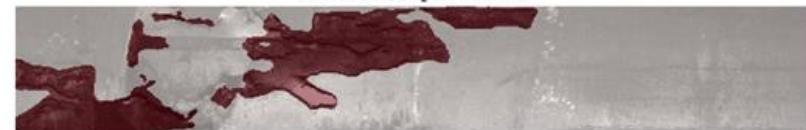
Оригинальное изображение



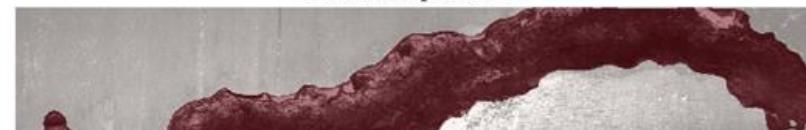
Оригинальное изображение



Маски с дефектом #4



Маски с дефектом #4



Маски с дефектом #3



Маски с дефектом #3

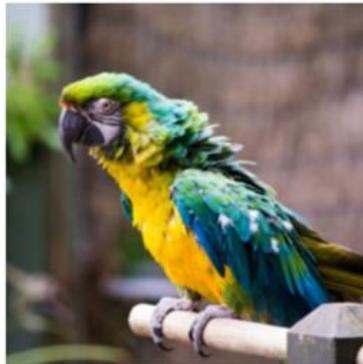


Почему востребовано

Кейс - Смартфоны

Примеры применения:

- высокое разрешение
- стабилизация съемки
- детализированные изображения



Задача улучшения качества изображения



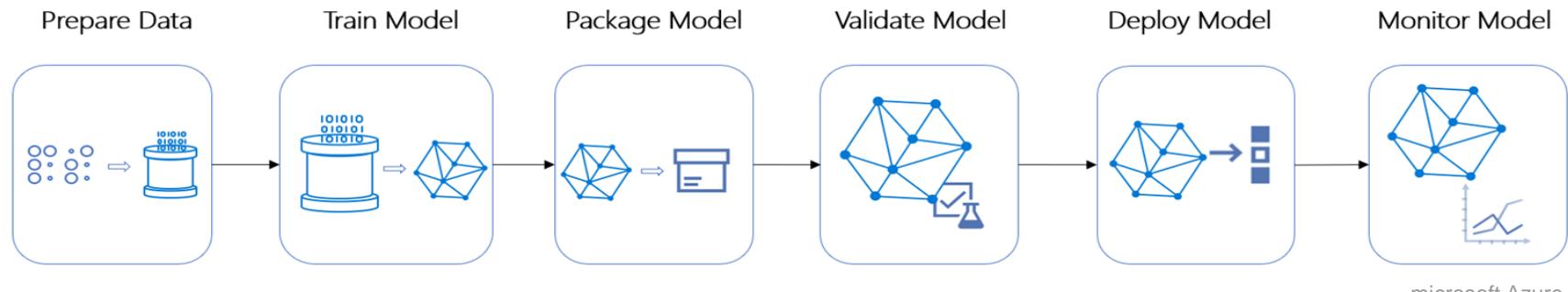
Снято с помощью Honor View 30 Pro (2019)



Снято с помощью Samsung A5 (2017)

CV Pipeline

Пример решения Computer Vision задачи

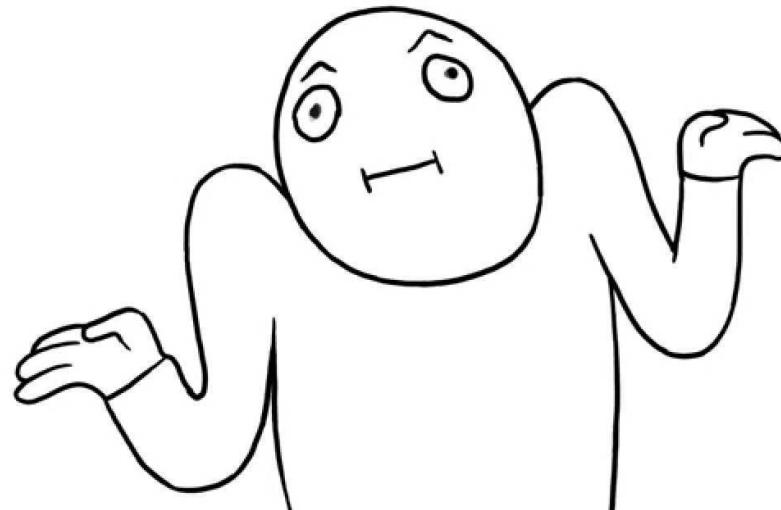


1. Общение с заказчиком.
2. Предварительная аналитика
3. Подготовка данных.
4. Построение моделей (fit-predict).
5. Интеграция моделей в железо или облако.
6. Поддержка и доработка.

CV Pipeline

Общение с заказчиком

Уже на этом месте вся разработка может закончиться.

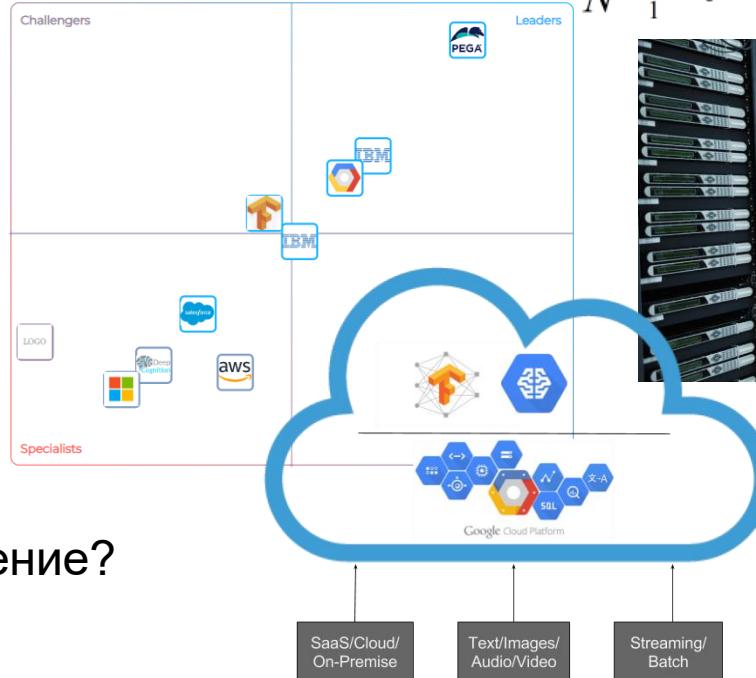


CV Pipeline

Общение с заказчиком

Ну а вообще:

1. Какая метрика?
2. Какие объекты?
3. Какие ограничения?
4. Платформа?
5. Сервисы?
6. Производительность?
7. Как производить дообучение?



$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

$$mAP = \frac{1}{N} \sum_1^N \int_0^1 precision(recall) dr$$



CV Pipeline

Предварительная аналитика

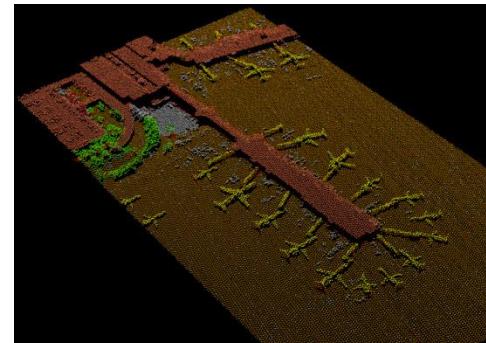
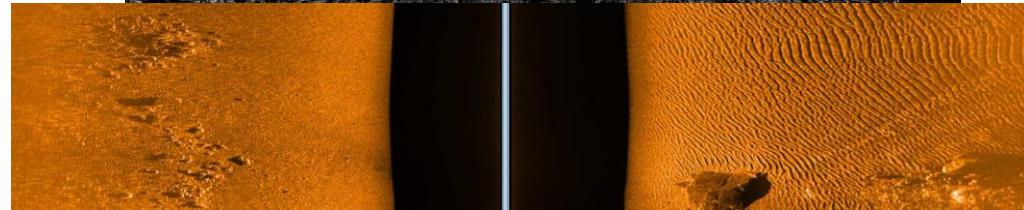
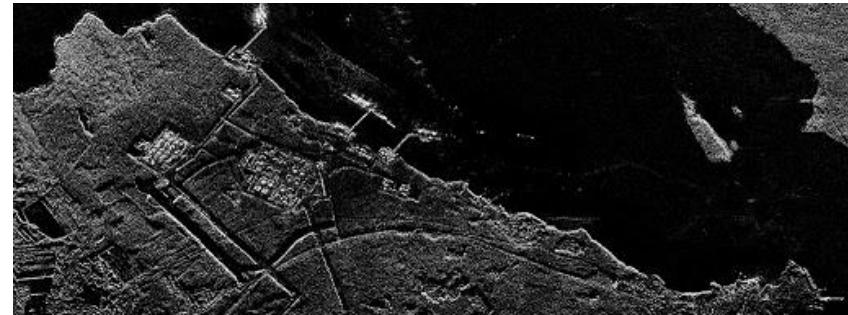
1. Оценка вычислительной сложности.
2. Оценка нагрузок со стороны пользователей.
3. Оценка требуемого и доступного объемов
данных для обучения моделей.
4. Оценка существующих решений.
5. Выявление структуры данных.

CV Pipeline

Подготовка данных

Какие данные могут на вас свалиться?

- Фото
 - One-photon камеры
 - Стереокамеры
- Видео
 - Высокоскоростные камеры
- 3D объекты
- Карты
- Гологramмы (Радиолокаторы и Гидролокаторы)
 - Фазовые данные
 - Амплитудные данные
 - Доплеровские данные
- Данные с лидара
- Данные с дальномера
- Мультиспектральные данные

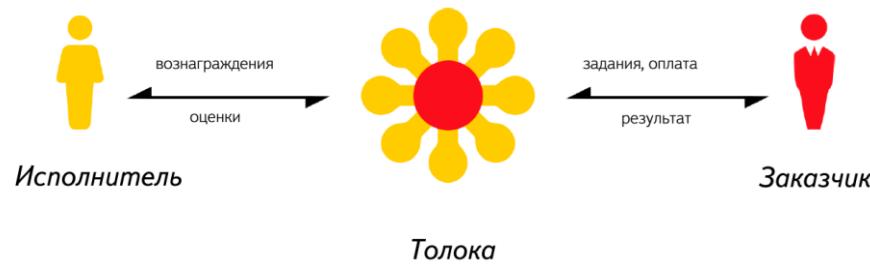


CV Pipeline

Подготовка данных

Как теперь все это разметить?

scale



LIONBRIDGE

clickworker

TranscribeMe!

Handl

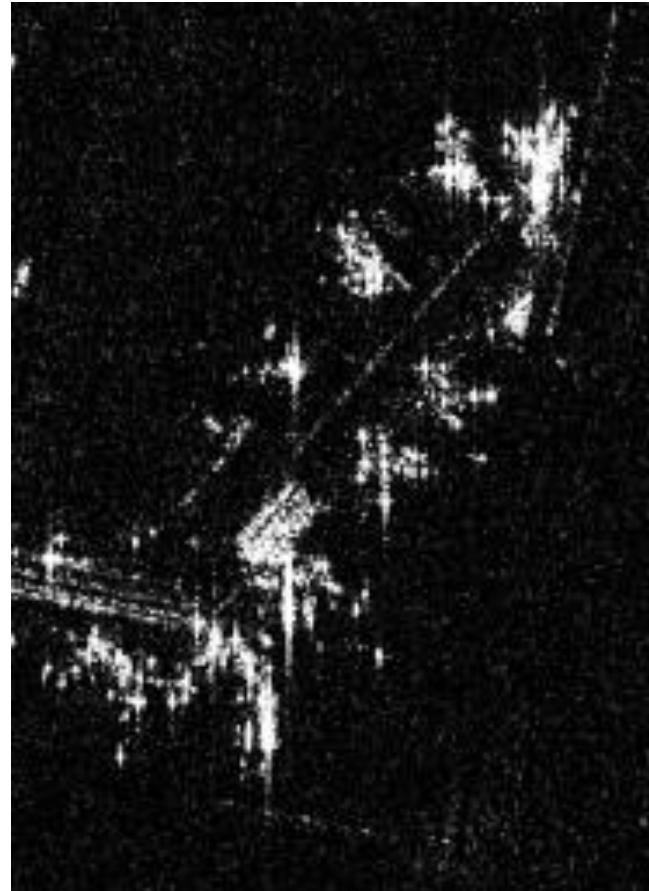
Dbrain

Amazon Mechanical Turk

CV Pipeline

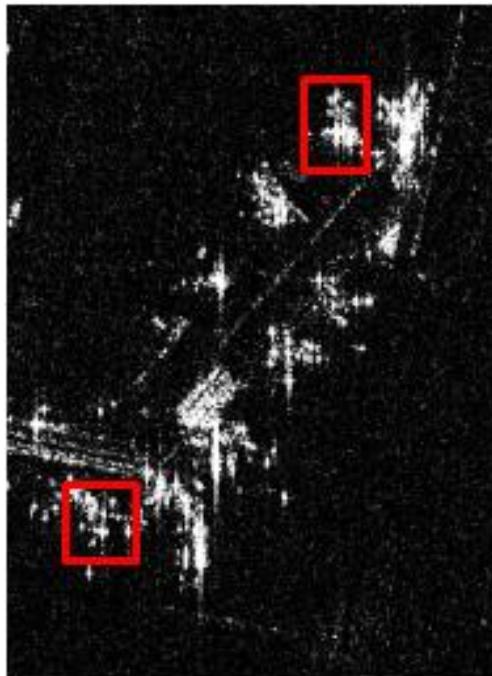
Подготовка данных

А если в данных ничего не понятно?



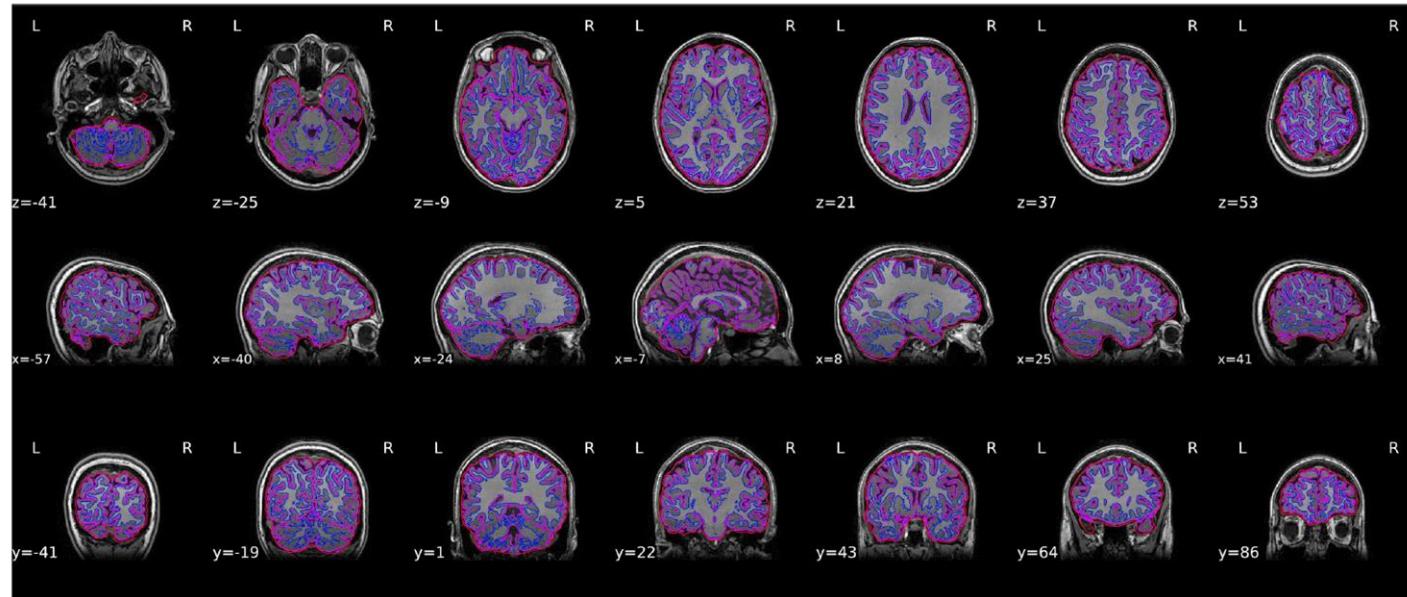
CV Pipeline

Подготовка данных



CV Pipeline

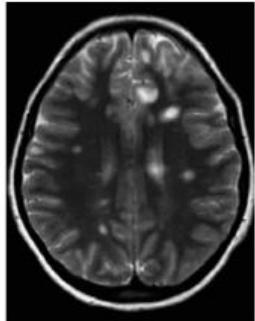
Подготовка данных



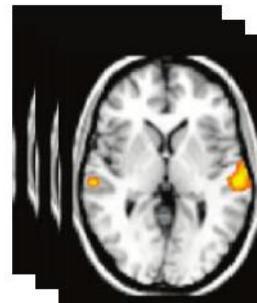
Pominova, Marina, et al. "Voxelwise 3D Convolutional and Recurrent Neural Networks for Epilepsy and Depression Diagnostics from Structural and Functional MRI Data." 2018 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 2018.

CV Pipeline

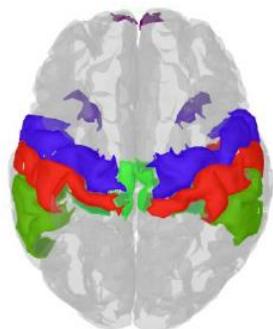
Подготовка данных



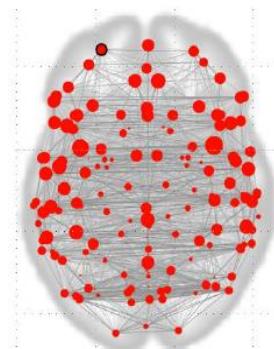
109 регионов
7-9 характерных зон



117 регионов
1D временная серия для
одной зоны



Морфологические
особенности:
Толщина, изгибы и
площади поверхности

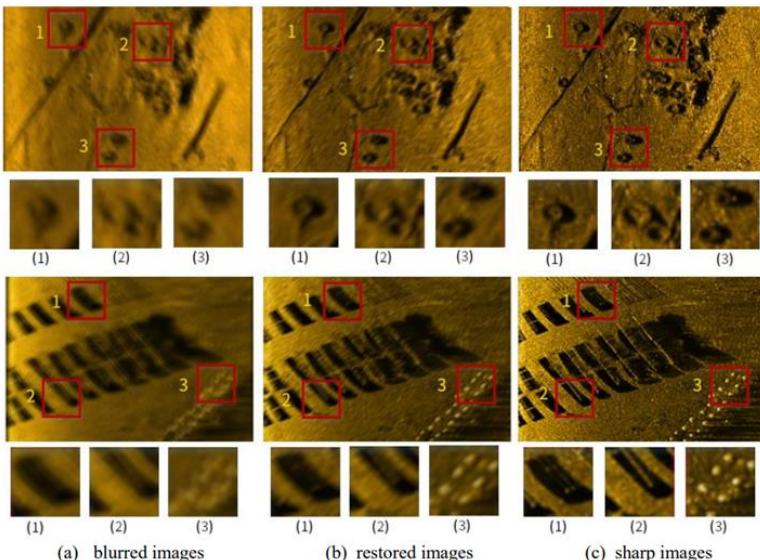


Функционально зависимая
матрица 117x177:
Корреляция между двумя
зонами во времени

CV Pipeline

Подготовка данных

А если данных нет и их нельзя снять или найти?



Algorithm 1 Create blur image

Parameters:

Velocity - Side scan sonar relative speed

Angle - Side scan sonar relative motion angle

EPS - Very small interference

holder - Correction threshold

```
1: procedure Create.blur(Image, Velocity, Angle,holder)
2:   sx =  $\lfloor \lfloor Velocity * \cos(Angle) + a * holder - Velocity * EPS \rfloor \rfloor$ 
3:   sy =  $\lfloor \lfloor Velocity * \sin(Angle) + a * holder - Velocity * EPS \rfloor \rfloor$ 
4:   for i = 0 to sy do
5:     for j = 0 to sx do
6:       psf[i][j] = i * |cos(Angle)| - j * sin(Angle)
7:       if  $\sqrt{i^2 + j^2} > Velocity/2$  and |psf[i][j]|  $\leq$  holder:
8:         temp = Velocity/2 - |(j + psf[i][j]) * sin(Angle)| / cos(Angle)
9:         psf[i][j] =  $\sqrt{(psf[i][j])^2 + temp^2}$ 
11:        psf[i][j] = holder + EPS - |psf[i][j]|
12:        if psf[i][j] < 0
13:          psf[i][j] = 0
14:        if Angle < 90 and Angle > 0
15:          psf  $\leftarrow$  flip left and right(psf)
16:          Anchor = (sx - 1, 0)
17:        elif angle > -90 and angle < 0
18:          psf  $\leftarrow$  flip up and down(flip left and right(psf))
19:          Anchor = (sx - 1, sy - 1)
20:        elif angle < -90
21:          psf  $\leftarrow$  flip up and down(psf)
22:          Anchor = (0, sy - 1)
23:        Kernel  $\leftarrow$  sub pixel interpolation(psf)
24:        Blurred image  $\leftarrow$  conv(Kernel, [Image.x+Kernel.x-Anchor.x, image.y+Kernel.y-Anchor.y])
25:      return Blurred image
```

Using Conditional Adversarial Networks to Deblur the Sonar Image of the Unknown Motion Blur Kernels

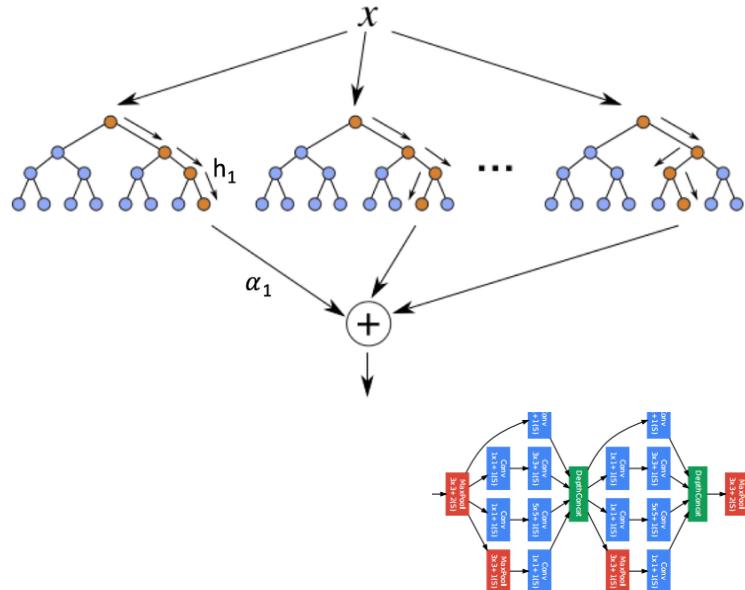
CV Pipeline

Построение моделей

- Выбор метода
- Выбор архитектуры
- Оптимизация гиперпараметров

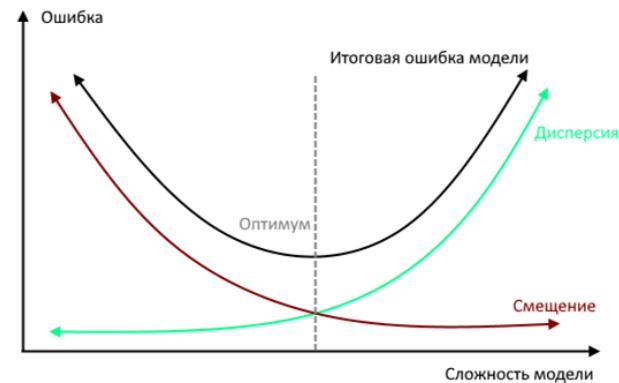
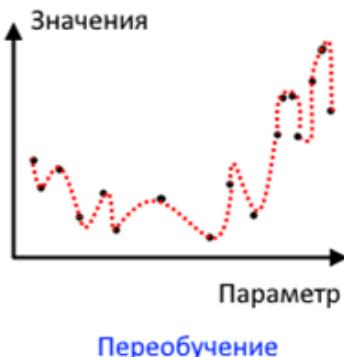
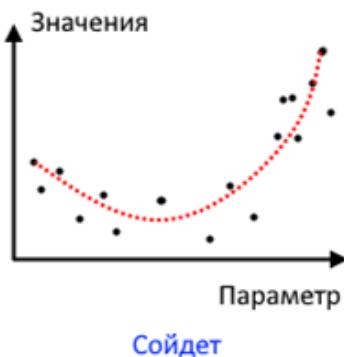
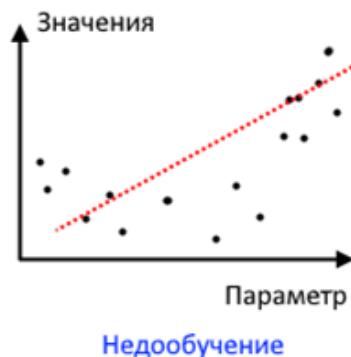
CV Pipeline

Построение моделей



CV Pipeline

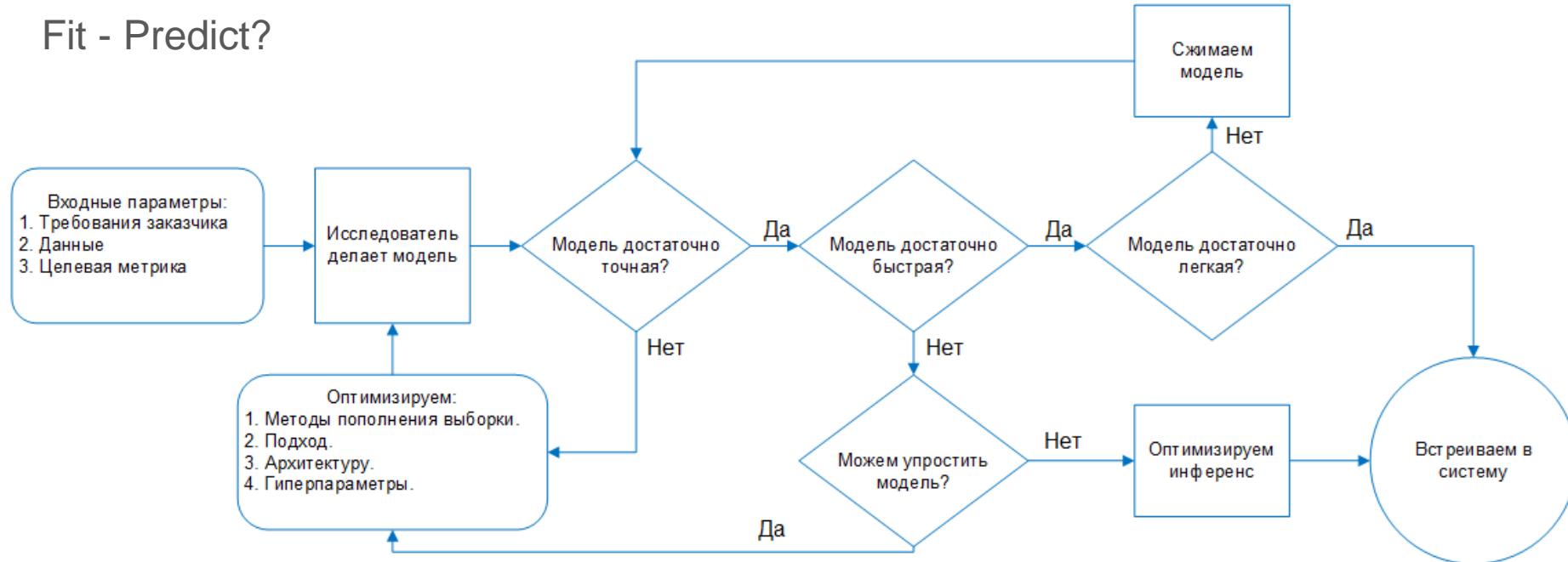
Построение моделей



CV Pipeline

Построение моделей и интеграция в железо

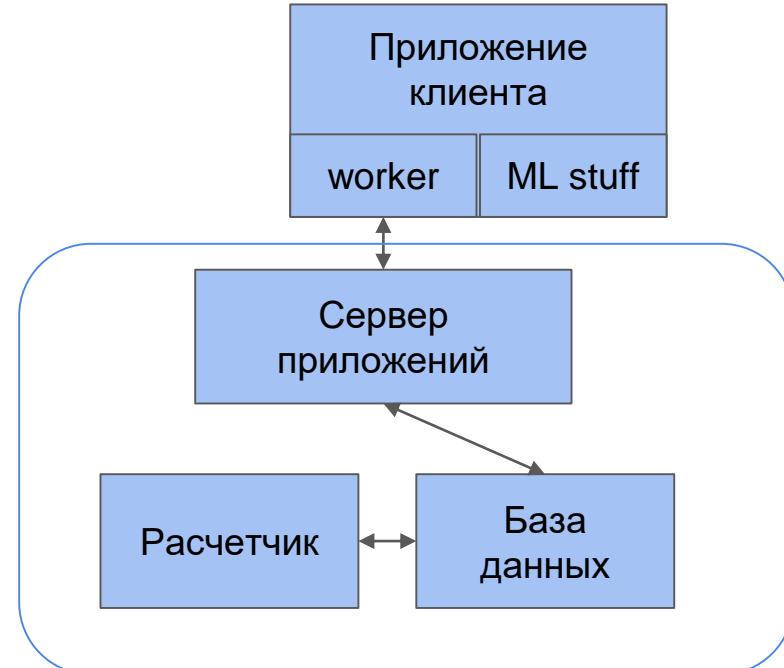
Fit - Predict?



CV Pipeline

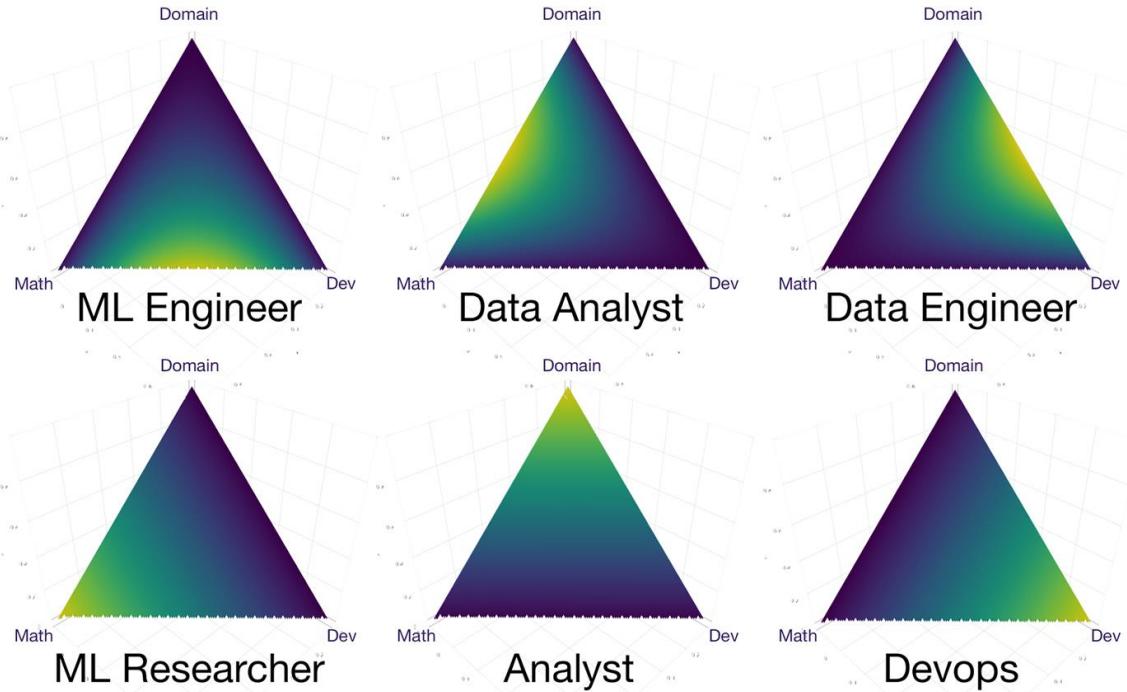
Поддержка и доработка

1. Сбор и хранение логов.
2. Дообучение моделей.
3. Обеспечение стабильности.
4. Масштабирование.
5. Управление ресурсами.



CV Pipeline

Рабочий процесс



Заключение

Выводы и напутствие

1. Познакомились с компьютерным зрением и его место в современном мире
2. Рассмотрели основных задач и кейсы из реальной жизни
3. Выявили первые шаги обучения компьютерному зрению