

Использование искусственного интеллекта и компьютерного зрения в прикладных задачах - примеры и открытые вопросы

Заславский Марк Маркович, зам. зав. каф. МО ЭВМ, к.т.н.
Филатов Антон Юрьевич, доцент каф. МО ЭВМ, к.т.н.

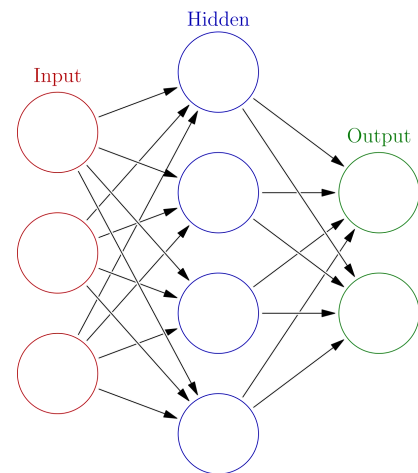
О чем будет доклад

- Есть многообещающие современные ИИ-решения ...
- Есть сложные проблемы из практики ...
- Почему они еще друг друга не нашли?
- Какие фундаментальные проблемы нужно решить в области ИИ?
- Как это проявляется в перспективных решениях и прототипах?



Оговорки

- Не весь ИИ - пространственные задачи и компьютерное зрение
- Компьютерное зрение - работа с данными специализированных сенсоров:
 - Радары
 - Лидары
 - RGB(D) камеры
 -
- Прикладные задачи -
 - связанные с повседневной деятельностью людей и организаций
 - имеют потенциал автоматизации
 - имеют ограничения и требования



Пространственный искусственный интеллект

- **Вопрос** - Готовы ли современные модели ИИ к решению комплексных проблем и какие есть перспективные направления?
- Пример - **пространственный ИИ**
- **Пространственный ИИ** - комплексное общее программно-аппаратное решение, подразумевающее
 - возможность **оперативной** работы в динамическом окружении,
 - одновременное решение **нескольких** пространственных задач,
 - **расширяемость** и возможность реализации новых сценариев,
 - только 3d.
- Пространственный ИИ про и для пользователей.

Пример приложений - роботы-уборщики

Роботы, в задачи которого входит “умная” уборка помещений в масштабе ЛЭТИ

- Убирает
 - полы и стены,
 - окна,
 - мебель.
- Двигает объекты (если нужно).
- Не мешает людям.
- Отслеживает заряд батареи.
- Ориентируется в географии ЛЭТИ.



Общие требования к пространственному ИИ

- Требования
 - Автономность,
 - Работа с минимальной задержкой,
 - Ориентация и движение в динамических условиях,
 - Способность решения нескольких высокоуровневых задач (с потенциальным расширением их числа)
- Подзадачи
 - Построение карты окружения (**Что вокруг меня?**)
 - Семантическая разметка (**Чем и кем я окружен?**),
 - Локализация (**Где я?**),
 - Навигация (**Как попасть в точку назначения?**),
 - Хранение и обработка данных (**Как согласовать мои наблюдения с другими агентами?**)

Построение карты окружения

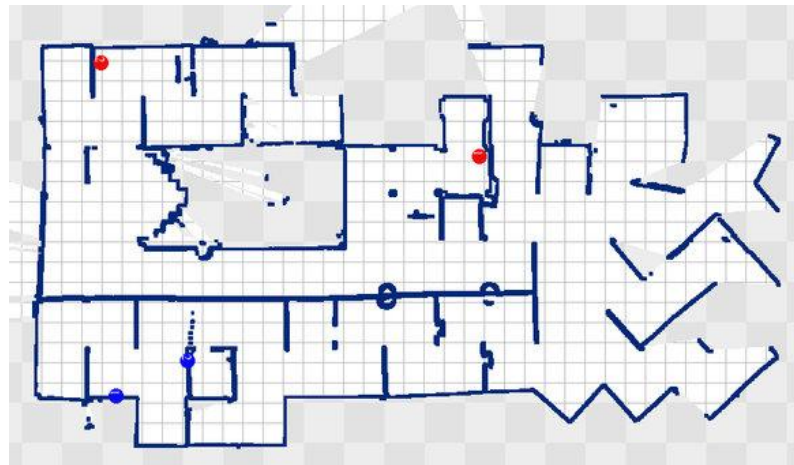
Подзадачи:

- формировать цифровое представление окружающей среды,
- прогнозирование карты.

Проблемы:

- система не видит сразу все окружение,
- динамические условия,
- погрешности в работе сенсоров

Решения - визуальная одометрия + ML



https://www.researchgate.net/publication/326580040_RoboCup_Rescue_2018_Team_Description_Paper_AutonOHM

Семантическая разметка

Задачи:

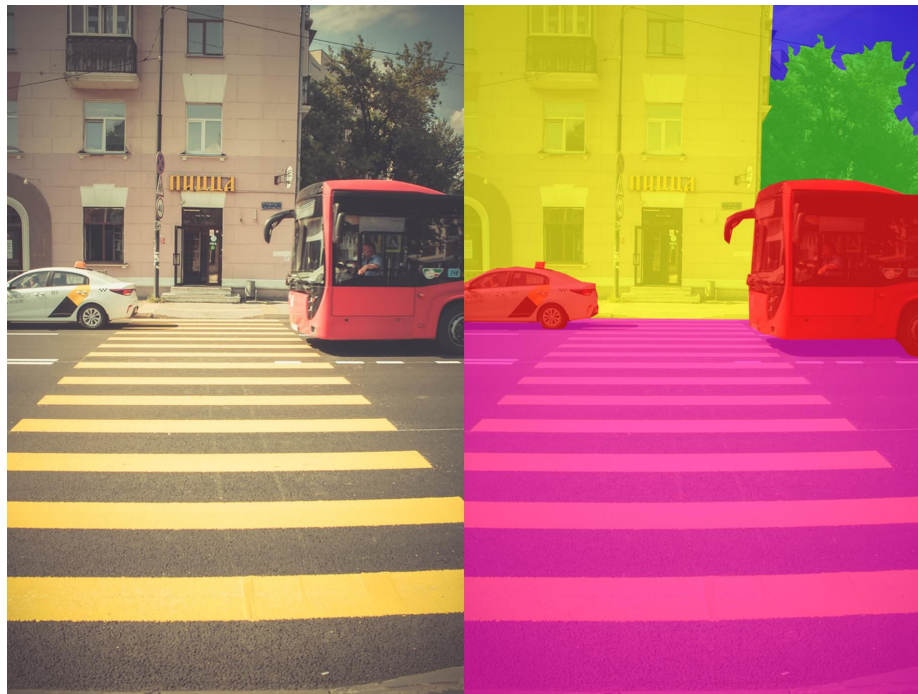
- семантическая сегментация карты,
- классификация объектов.

Проблемы:

- точность сегментации и классификации,
- повторяемость результатов.

Решения - DeepLearning

А какие классы нам нужны?



Локализация

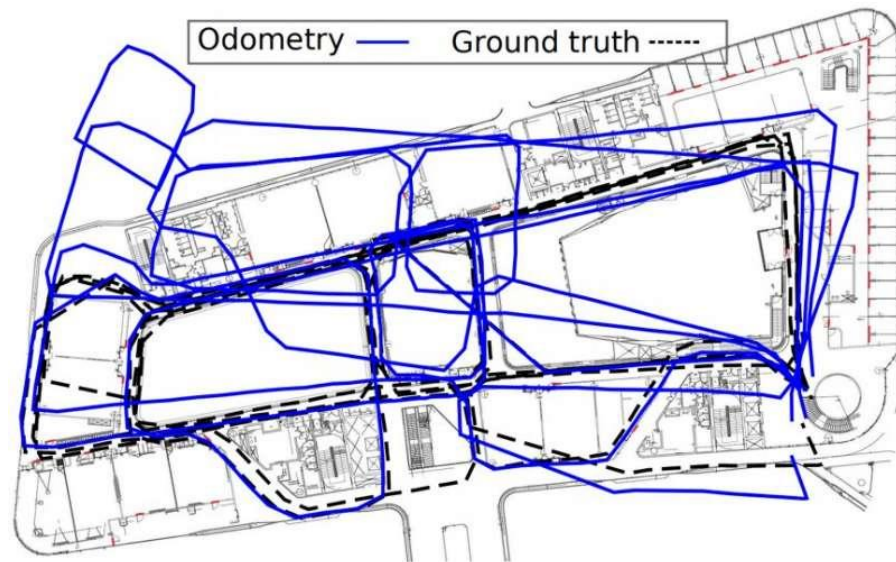
Максимальное точное позиционирование агентов системы

Проблемы:

- Разные сенсоры для разных контекстов (когда использовать какие датчики?),
- Точность и быстродействие,
- Как объединять оценки позиционирования из различных источников?

Решения - SLAM-алгоритмы + аппаратные хитрости

Как быть в помещении?



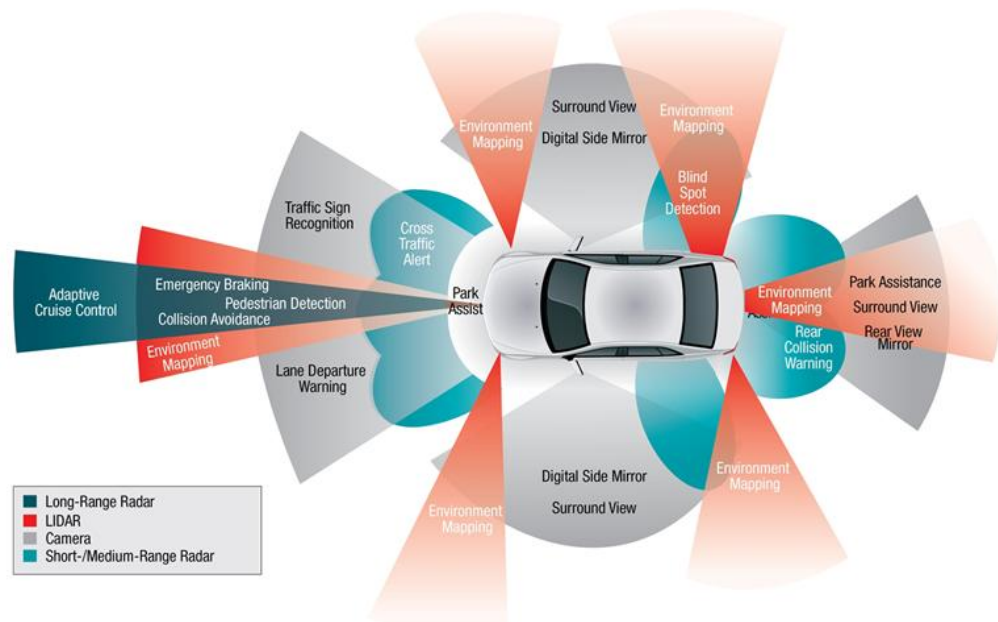
<https://techxplore.com/news/2020-01-collaborative-simultaneous-localization-technique-wi-fi.html>

Навигация

Перемещение к объекту интереса, учитывающее несколько конкурирующих метрик качества.

Проблемы:

- Поиск локальных решений и прогнозирование карты (== мы не знаем / не видим всю карту сразу),
- Учет динамических объектов (люди, птицы, автомобили) и избегание столкновений,
- Энергопотребление.



Как можно строить маршрут?

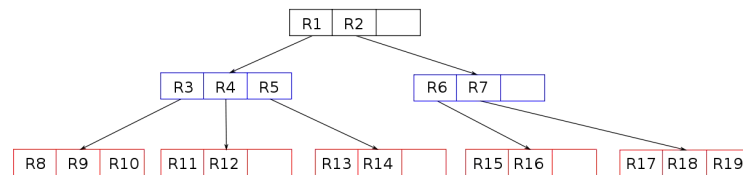
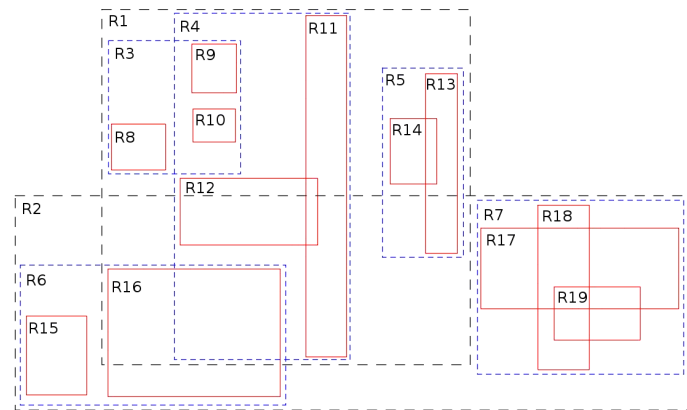
Хранение и обработка данных

Обеспечение бэкэнда для работы системы.

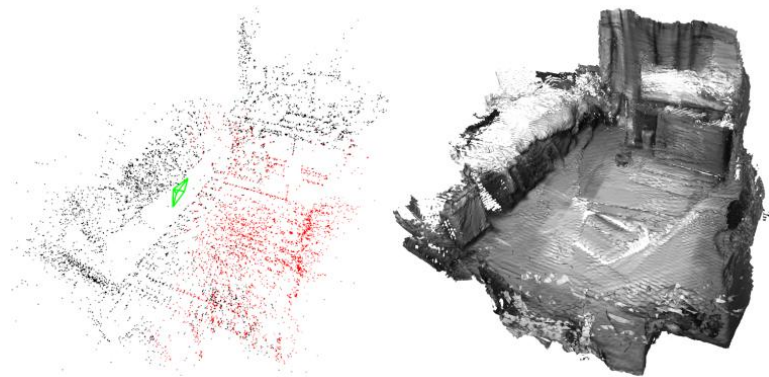
Проблемы:

- Объем карт и сложность их обработки / пересылки (что делать?).
- Слияние данных от агентов.
- Обеспечение агентов **достаточными** данными для автономной работы.

Структуры для работы пространственными данными?



Выводы



<https://arxiv.org/pdf/2107.08994.pdf>



(a) point cloud with semantic label

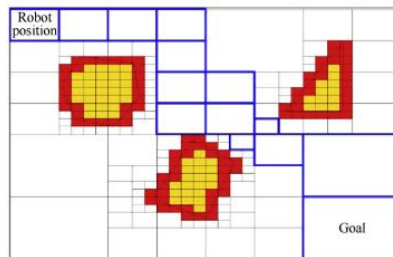


(b) segment result



cup
keyboard
monitor
mouse
teddy bear

(c) semantic map
https://github.com/qixuxiang/orb-slam2_with_semantic_label

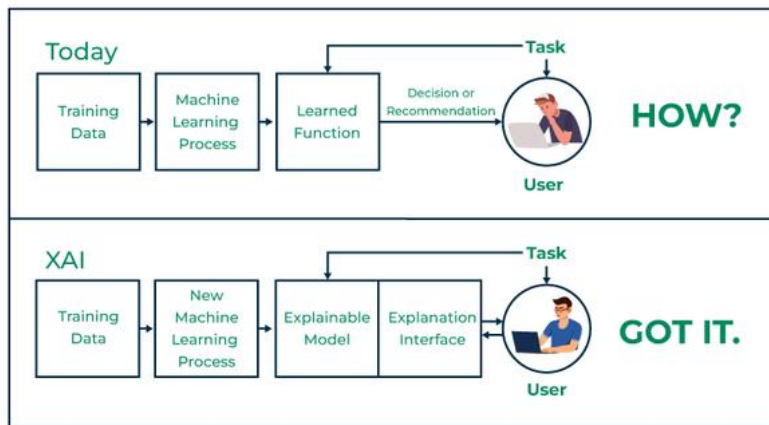
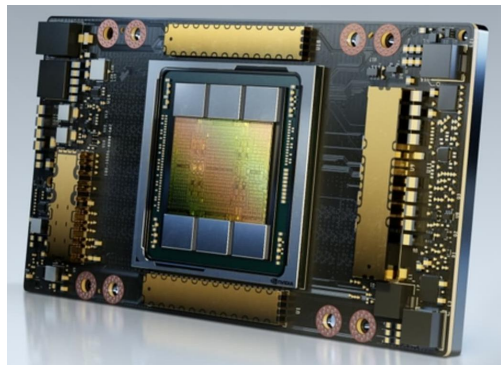


Выводы

- **ИИ спотыкается о масштаб задач**
- **Технические моменты ограничивают применение**
- **Проблемы**
 1. **Технические:**
 - a. Формат и объем карт окружения
 - b. Накопление ошибки при инкрементальном режиме работы
 - c. Общая вычислительная сложность и скорость работы системы
 2. **Методические:**
 - a. Данные для обучения
 - b. Надежность моделей
 - c. Валидация моделей
 3. **Эксплуатационные:**
 - a. Настройка и адаптация (переносимость в другие условия) модели
 - b. Отладка
 - c. Ответственность за действия моделей
 - d. Дополнительные данные для работы

Перспективные направления

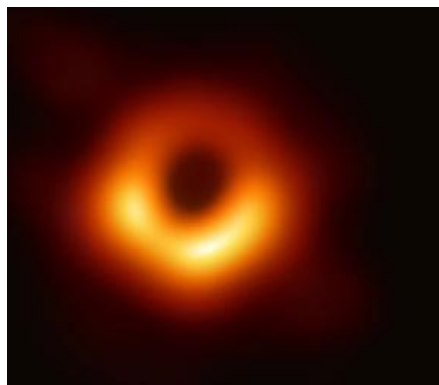
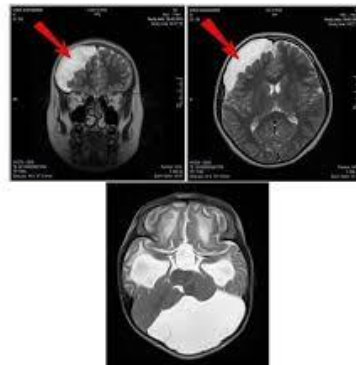
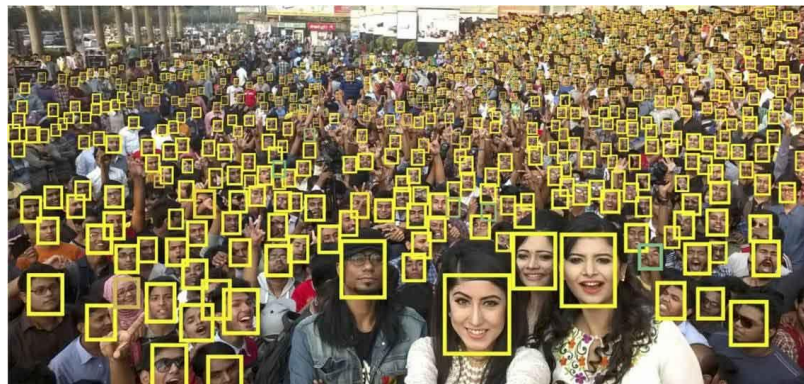
- Объяснимый и интерпретируемый ИИ
- Общие модели -> Частные модели
- Синтетические данные и аугментация (возможно, силами других ИИ-моделей)
- Новые подходы к аппаратному обеспечению, поддержка нейровычислений



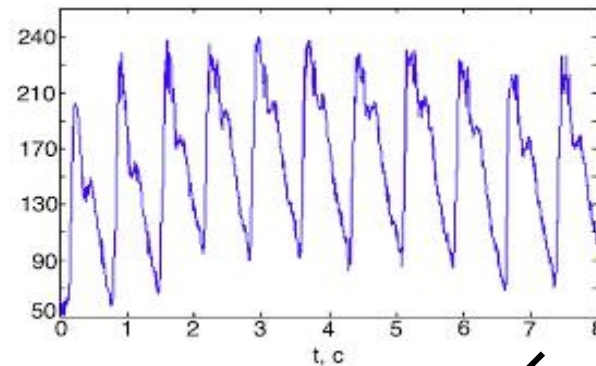
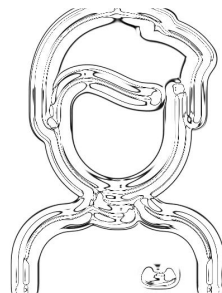
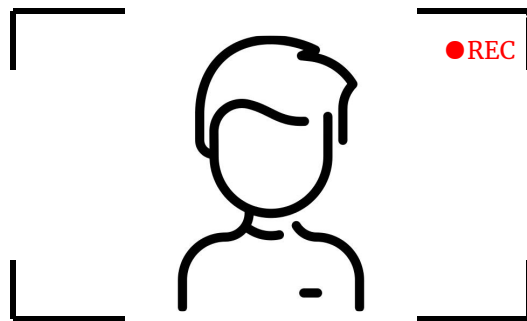
Часть 2 о других областях применимости

Филатов Антон Юрьевич, к.т.н.

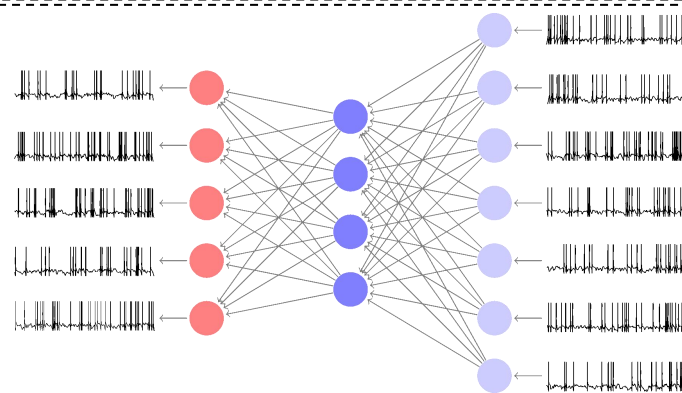
Где ещё применяется компьютерное зрение



Определение стресса по видео



0:00 - 0:25	Не стресс
0:25 - 0:54	Стресс
0:54 - 3:47	Не стресс



Определение пульса по видео 1

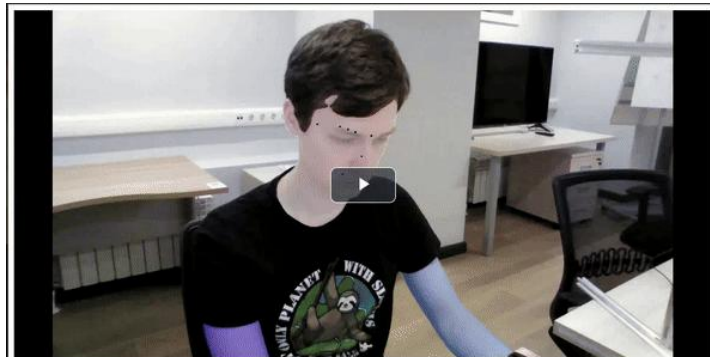
Сегментация изображения с помощью ИС (1 Hz)

Прогнозирование
сегментов в
промежутках
между
сегментациями
(30 Hz)

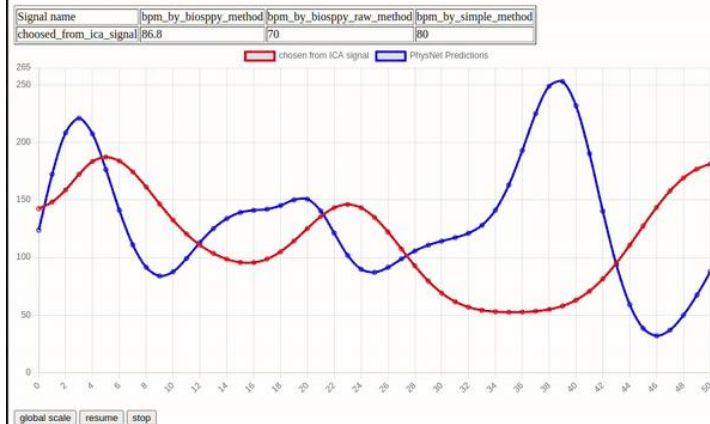
Поиск точек интереса (pose
detection) на теле человека с
помощью ИС

Сопоставление точек интереса
и сегментов

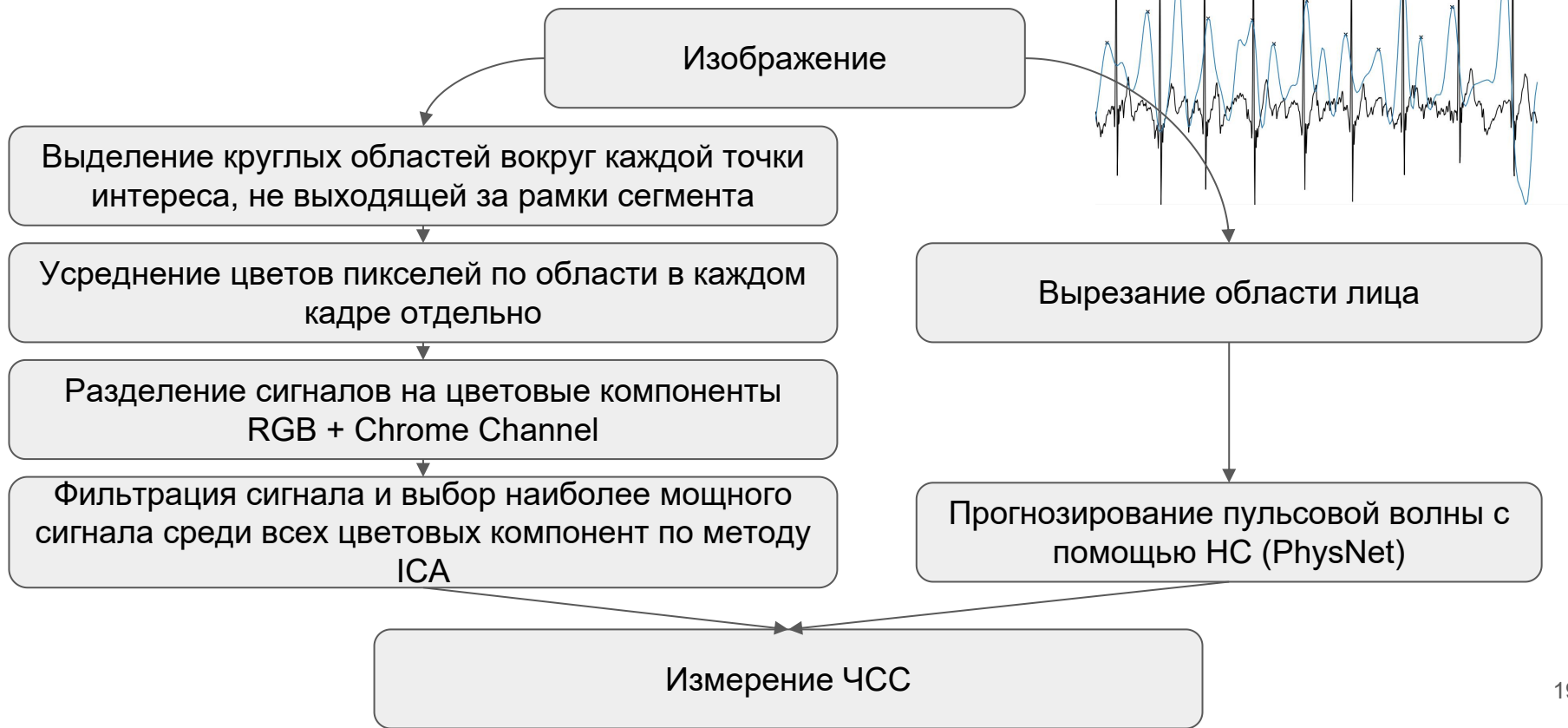
Смещение сегмента на новый
кадр с помощью оптического
потока



Here's your result:



Определение пульса по видео 2



Точность определения пульса

Точность зависит от среднего пульса человека.

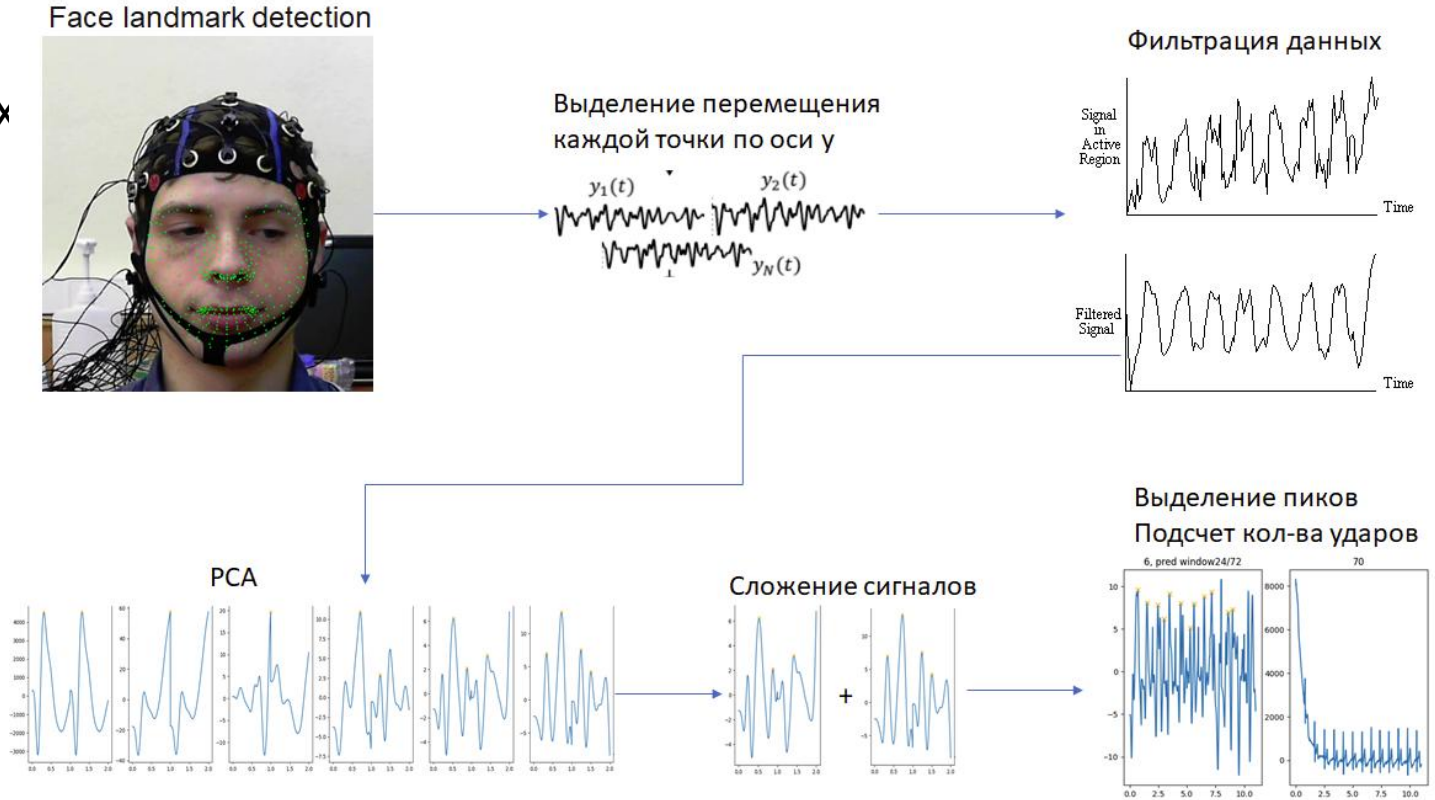
При пульсе порядка 70-80 ударов в минуту погрешность составляет 2-3 удара в минуту

При пульсе порядка 100 уд/мин - погрешность порядка 8-9 уд/мин.

Это связано с параметрами, которые подбираются исходя из знания о длине видео, качества картинки, интенсивности движений человека.

Другой возможный подход без машинного обучения

Детекция мельчайших движений



Определение стресса по пульсовой волне

- Стресс - это что вообще такое
- Существует множество статей, где стресс измеряют по КГР и ЭЭГ
- мы разработали метод, основанный на персонификации, который позволяет распознавать стресс по пульсовой волне с точностью до 98,5%

Работа с ИИ

1. Подготовка:

- a. Выбор формата данных для обучения
- b. Определение достаточного количества данных
- c. Выбор архитектуры модели ИИ для обучения

2. Обучение

- a. Определение количества требуемых вычислительных ресурсов

3. Использование

- a. Определение количества требуемых вычислительных ресурсов
- b. Определение граничных состояний входных данных, при которых гарантируется работа

Этапы 2 и 3 (Обучение и использование) - есть сложности

- Современный ИИ - недешёвое удовольствие! Мощностей персональных ПК уже не хватает
- Как понять, когда нейросеть ошибётся?
- Нельзя не валидировать вывод нейросети



Всё ли может Искусственный Интеллект?

- контекстозависимость
- ограниченность валидации
- узконаправленность
- зависимость от формата



Спасибо за внимание!