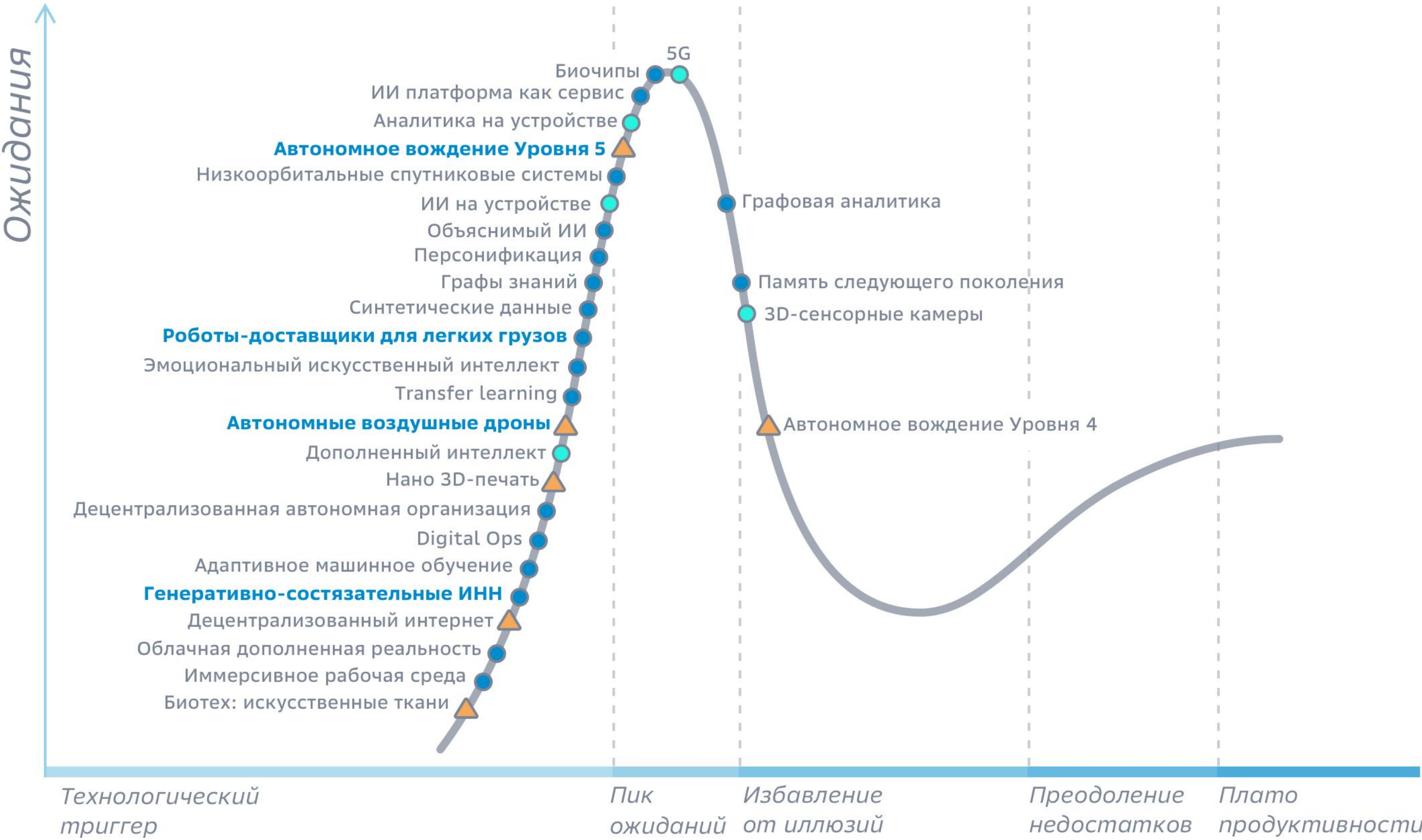


# Опыт применения продуктов Intel в робототехнике

*Алексей Гонноченко*  
*СТО*

**Sberbank**  
Robotics  
Laboratory

# Наш технологический форсайт



Выделяем основные технологические триггеры

на горизонте

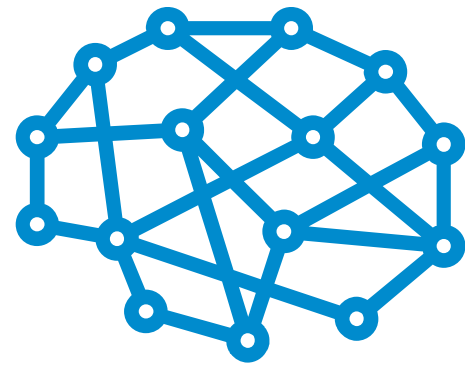
**5-10 лет**

Автономное вождение Уровня 5  
Роботы-доставщики для легких грузов  
Автономные воздушные дроны  
Генеративно-состязательные ИИ

○ меньше 2 лет   ● 2-5 лет   ● 5-10 лет   ▲ более 10 лет   ● устаревает до плато

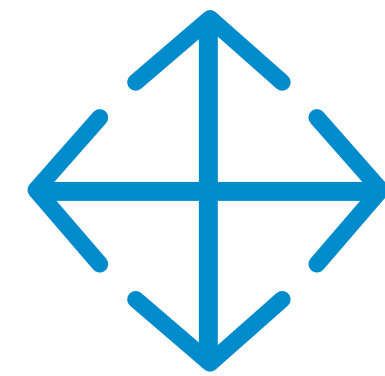
Источник: Gartner

# Стратегические направления



## Искусственный интеллект

Роботы способны самостоятельно обучаться и работать над ошибками



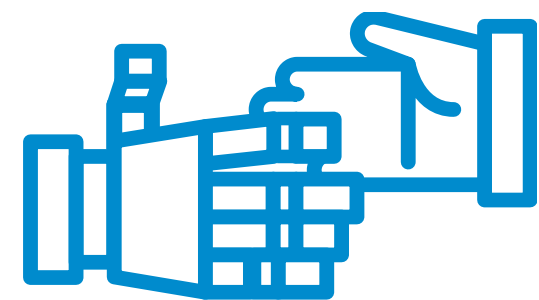
## Мобильность

Роботы умеют действовать без ограничений в мире людей



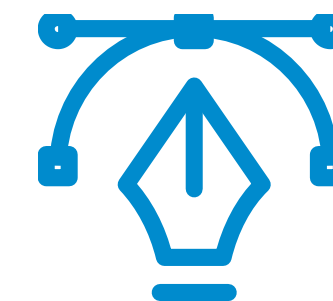
## Автономность

Роботы способны самостоятельно принимать решения в любых ситуациях



## Коллаборативность

Роботы полностью безопасны для людей



## Дизайн

Роботы удобны, понятны и предсказуемы

# Команда, которая делает роботов

**35**

инженеров  
и руководителей  
проектов

**25%**

имеют ученую степень или  
готовятся ее защитить

**70%**

пришли из  
стартапов

**99%**

работают по  
специальности



# Что умеют наши инженеры?

## Ключевые компетенции

### Электроника

Разработка и программирование модулей, исполнительных устройств, интеграция внешних модулей в роботов.

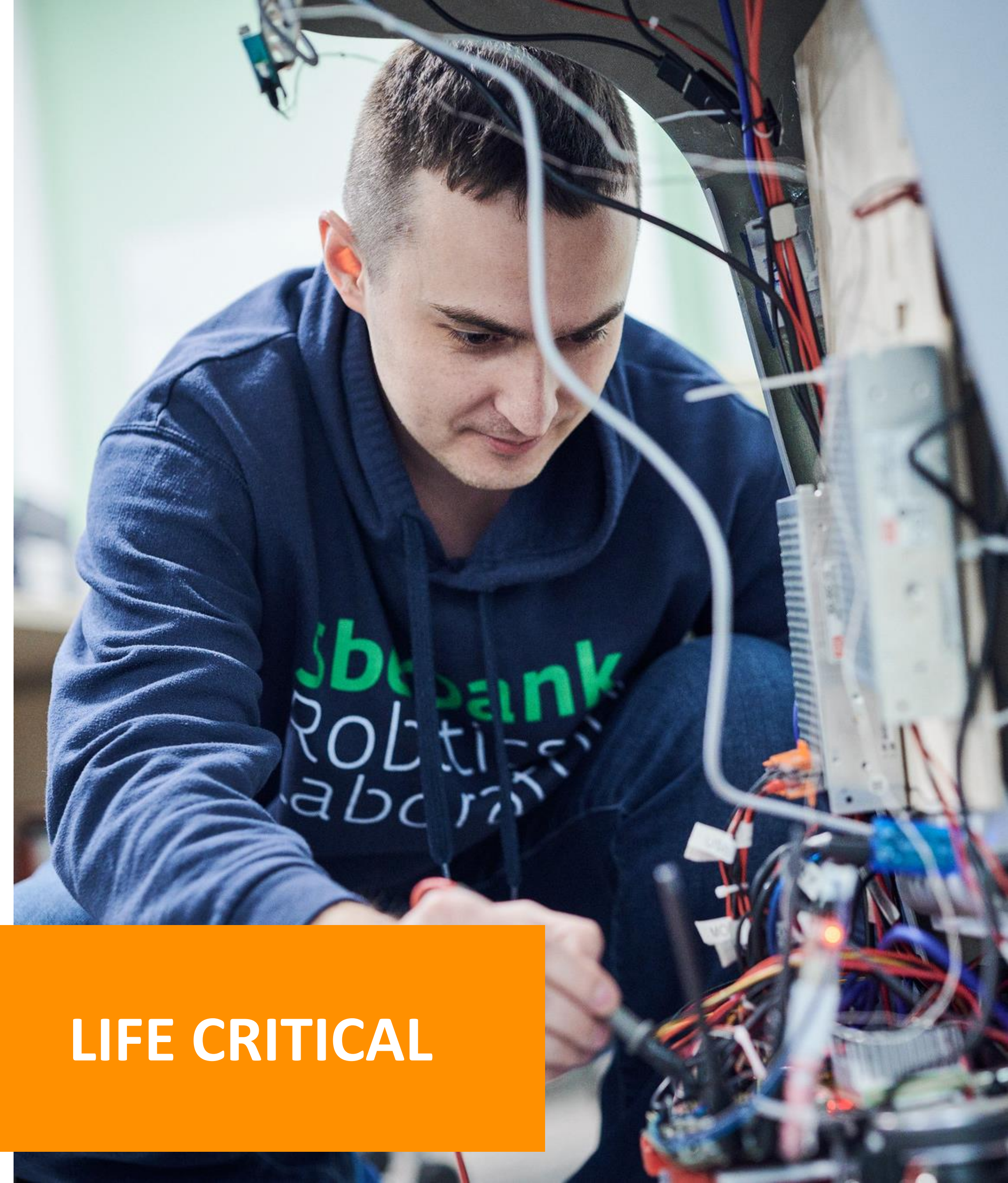
### Системы управления и восприятия

**Восприятие:** Глубокое обучение в обработке сенсорных данных, анализ и объединение разнородных данных. Работа с лидарами, камерами, ультразвуковыми, лазерными и радарными датчиками.

**Управление:** Классическое управление, машинное/глубокое обучение, Robotics learning, прогнозное управление.

### Конструирование

Прототипирование узлов и систем, разработка документации, производство прототипов. Исследования в области новых исполнительных механизмов и захватов.



**LIFE CRITICAL**

# Лабораторию определяют три вещи



Люди, которые способны  
создавать уникальные  
вещи



Правильные инструменты  
для работы этих  
людей



Умение создавать новые  
точки генерации новых  
идей

# Откуда приходят хорошие идеи?...

Конференции



Стажировки в лаборатории



Стажировки в компаниях



Правление Сбербанка



**30+**  
стажеров пришли к нам из  
лучших учебных заведений  
страны

МГУ  
МГУ 1830  
МИРЭА  
Российский технологический университет  
МФТИ  
МИСИС  
Университет науки и технологий  
Станкин  
Московский Государственный Технологический Университет  
Сколтех  
Сколковский институт науки и технологий  
21  
МОСКОВСКИЙ ПОЛИТЕХ





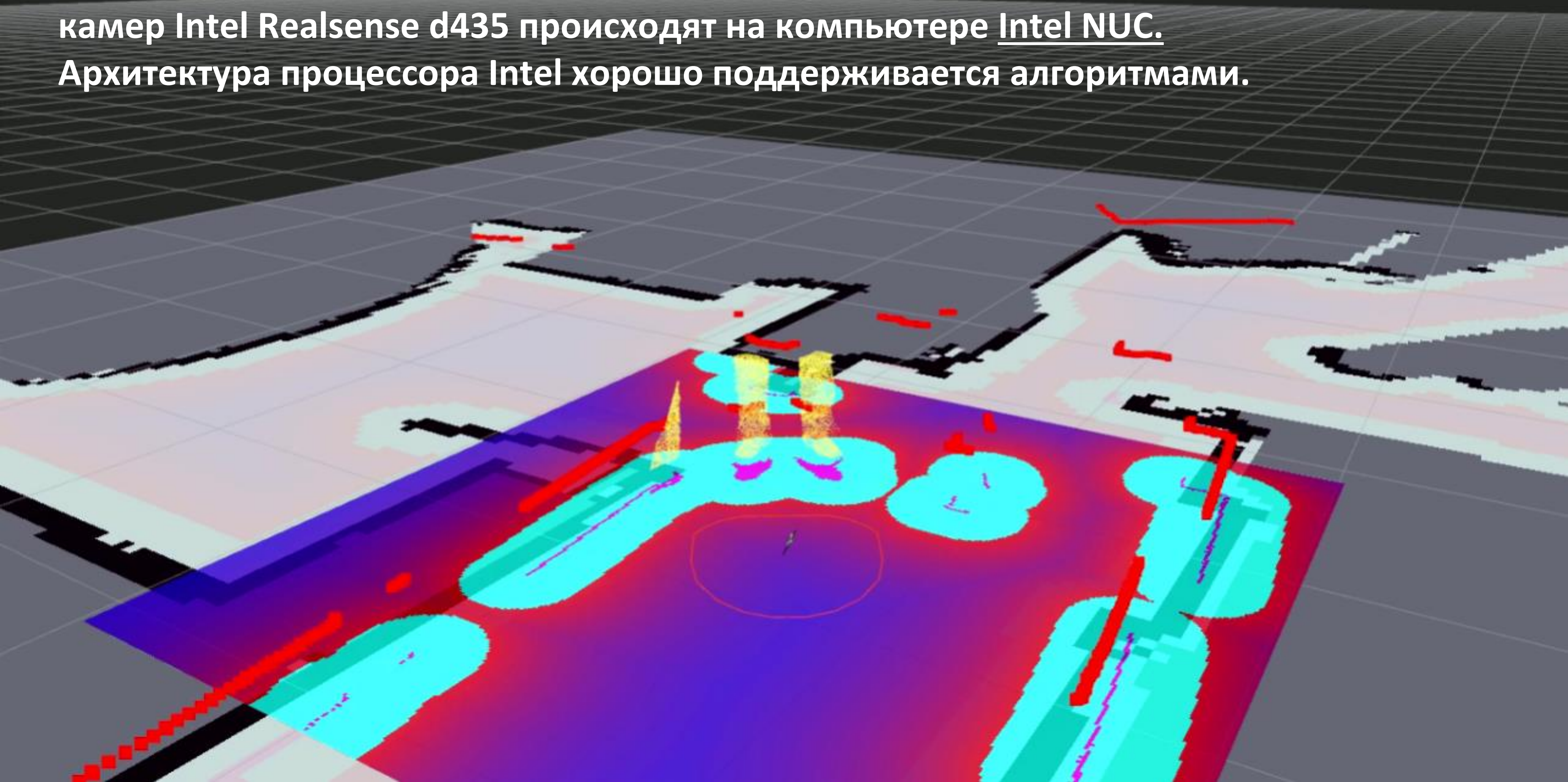
# Проект: логистический робот «ГЛЕБ»

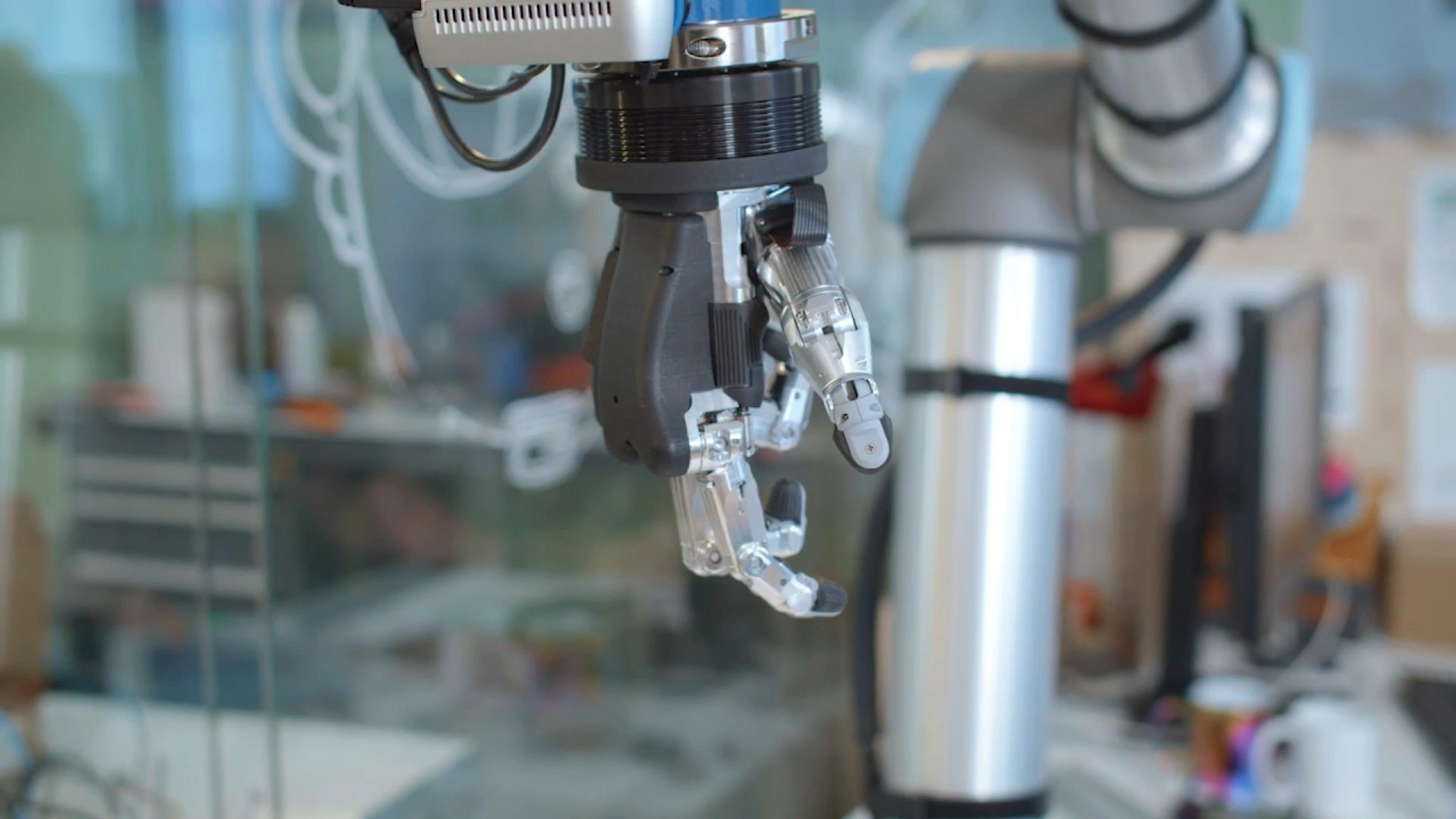
«ГЛЕБ» – робот для автономной доставки предметов малого и среднего размера внутри помещений

- Применены две камеры
- **Intel Realsense D435** для избегания столкновений с препятствиями. Карта глубины с камеры имеет хорошую точность и стойкость к помехам вроде солнечного света, одновременная работа нескольких камер. Intel предоставляют хороший драйвер для ROS и обновления прошивки.
- Возможны промышленные варианты исполнения камеры.



Вычисления SLAM алгоритмов и обработка изображений глубины с камер Intel Realsense d435 происходят на компьютере Intel NUC.  
Архитектура процессора Intel хорошо поддерживается алгоритмами.





# Проект: kNOwTeacher



Исследовательский проект по глубокому обучению искусственных нейросетей для манипулирования объектами разнообразной формы

- Робот учится захватывать и перемещать объекты любой формы, включая мягкие и деформируемые
- Применяются искусственные нейросети и компьютерное зрение
- Процесс обучения автоматический

# Проект: kNOwTeacher

## Object Grasping and Manipulating According to User-defined Method Using Key-points

A. N. Semochkin  
Sberbank Robotics Laboratory  
PJSC Sberbank  
Moscow, Russian Federation  
Semochkin.A.N@sberbank.ru

S. Zabihifar  
Sberbank Robotics Laboratory  
PJSC Sberbank  
Moscow, Russian Federation  
Zabikhifar.S@sberbank.ru

A. R. Efimov  
Sberbank Robotics Laboratory  
PJSC Sberbank  
Moscow, Russian Federation  
AREfimov@sberbank.ru

**Abstract**— This paper presents an approach for grasping objects by the robotic manipulator at predefined points by user-defined grasping methods for cases of arbitrary position and orientation of objects. The key idea is to implement a detection algorithm for user-defined key-points on the reference object image. As these points can be detected and localized on the object surface, it is possible to generalize the initial sequence of robotic manipulations, pre-recorded for the single scene, to the set of situations with arbitrary object orientation and location within the workspace. The robotic arm had a depth camera, mounted on the end effector for 3D key-points detection and anthropomorphic robotic hand.

**Keywords**— object grasp, neural network, robotic manipulator

### I. INTRODUCTION

A task of object grasping and subsequent manipulation in robotics requires prior knowledge of object location, orientation, and topology. Moreover, the solution itself can be generalized to different shapes, sizes, locations and orientations combinations. Optic sensors, capable to generate color images and depth maps, seem to be the best option for data acquisition and evaluation. A large body of work on objects grasping addresses only computer vision and optic sensors, either stationary or mounted on the robotic arm. At a category level, the task of detecting and manipulating certain objects can be solved by instance segmentation, as it was demonstrated in the Amazon Robotics Challenge (ARC) [1,2]. Object-level segmentation, however, does not provide any information about object structure, and hence may not be suitable for solving more complex problems.

In this article, we use predefined key points on the object image to determine the location and orientation of the object in the real world. To detect certain points and grab an object in its vicinity, we use a robotic arm with end-effector consisting of a depth camera and a gripper system. That way, we can always determine the camera orientation for each object image and localize the detected key points on the real object using the depth camera. This data is enough to calculate object location and orientation. While not previously applied to the anthropomorphic robotic manipulation domain, recent studies have demonstrated advances in learning dense visual object descriptors [3,4], including self-supervision from raw RGBD data [4], using 3D key-point determination and specific grasping method which inspired our present work.

Convolutional neural networks proved to be a promising tool for the object key-point detection task. Using a neural network [5], we can map color image space to the descriptors space to minimize the distance between two different pixel descriptors from two images, corresponding to a single point on the object surface. As a result, the pre-trained network generates shape representations based on object images. A

physical robotic manipulator equipped with a camera can create the training datasets for such neural network automatically.

For grasping the object after determination its pose and location, it is necessary to keep in mind its pose and use a gripping method that suits best for subsequent manipulations. For example, the task of putting shoes on a shoe rack and aligning it to one side is considered in [6]. Exact key-point data enable the use of copying approach when human operator “teach” the robot to grasp the object in “free drive mode” by recording movements' sequence. That way, it is sufficient to reorient the object relative to discovered key-points to perform the grasping task for different object orientations.

In this work, we present the prototype of the system, capable of training robotic manipulator to object grasping without coding. The entire pipeline including neural network learning, key-points detection, and object grasping learning procedure is available through the graphical user interface.

### II. RELATED WORK

Previous works address four main domains: grasping algorithms, descriptors learning, manipulation methods and pose estimation. The combination of depth camera data and neural nets is a standard approach for the object grasping task. In grasping algorithms domain in [7], the RGBD dataset has been used to train two neural networks to detect the grasping areas for different objects. In [6], a grasp planner has been developed, which uses the detected key-points in combination with local dense geometric information from a point cloud. The grasp planner utilizes the detected key-points to reduce the search space of grasps. Key-points determined from a trained neural network of 2D images input data and depth value using the integral network [8]. After discovering the local geometry part to focus on, a grasp can be found by a variety of geometric or learning-based and object-agnostic grasping algorithms [9,10]. A novel formulation of robotic pick and place task has been proposed in [11] as a deep reinforcement learning (RL). It solves the problem of unknown exact manipulation objects geometry when prior knowledge of the general class of objects for which the system was trained is needed. Dexterous in-hand manipulation policies have been introduced in [12] that are capable of performing vision-based object reorientation on a physical Shadow Dexterous Hand using reinforcement learning.

In our work, grasping procedure using anthropomorphic robotic arm and hand has been learned step by step for one intra-category object relative to its key-points and generalized to other objects. A humanoid 5-finger hand gripper allows learning object grasping in a human-like manner.

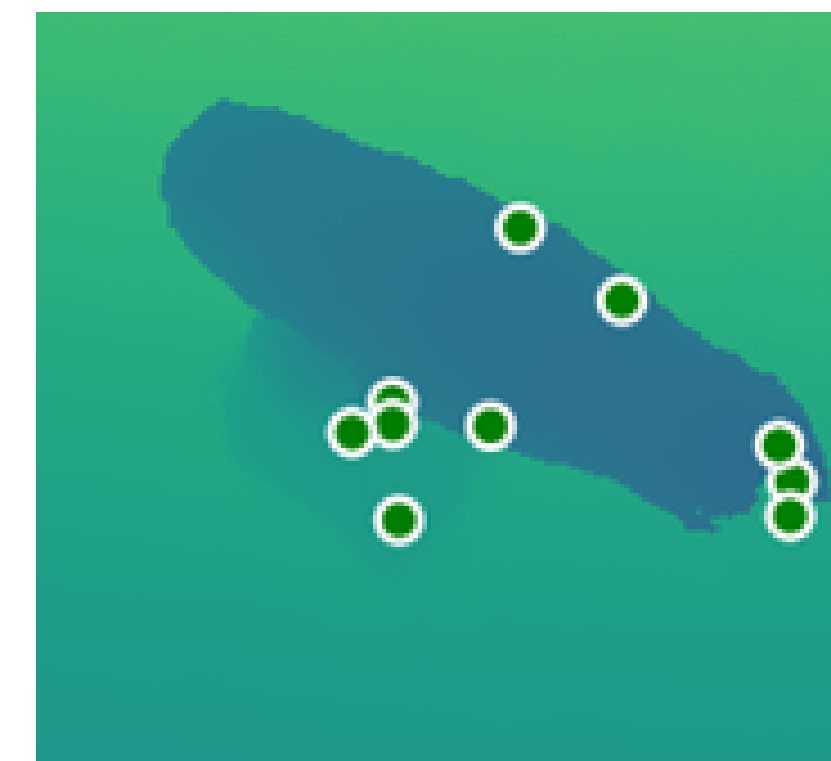
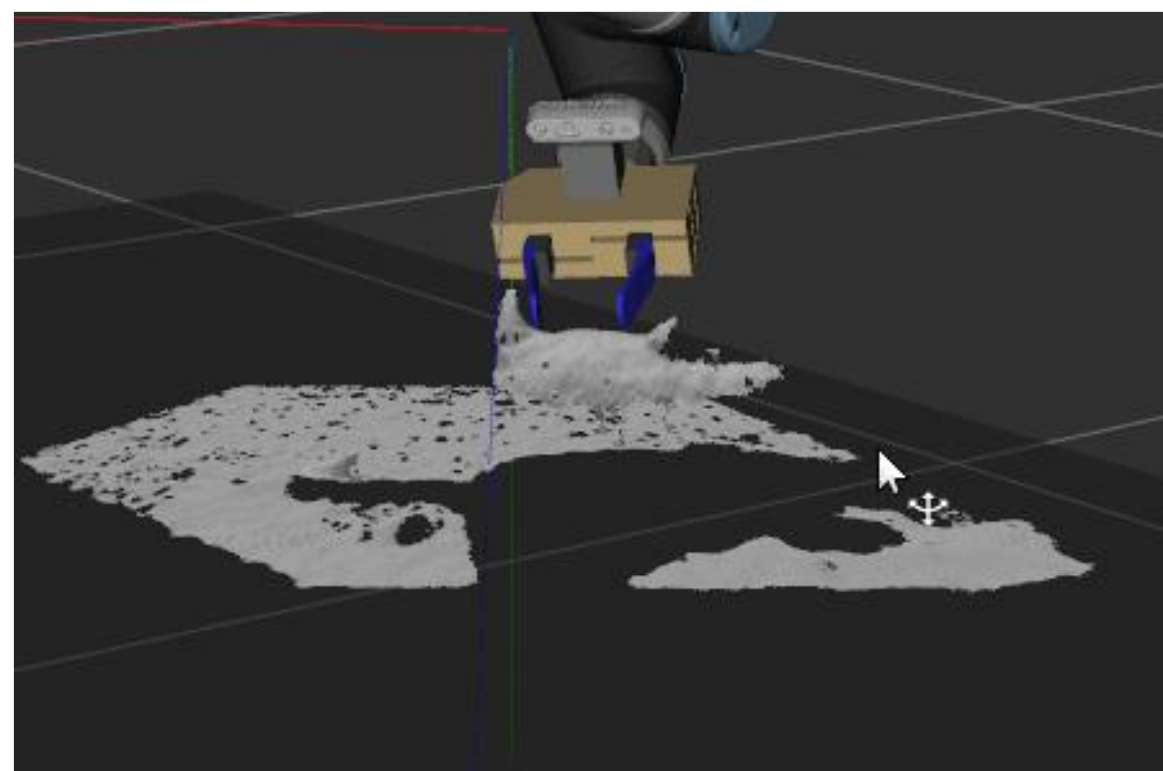
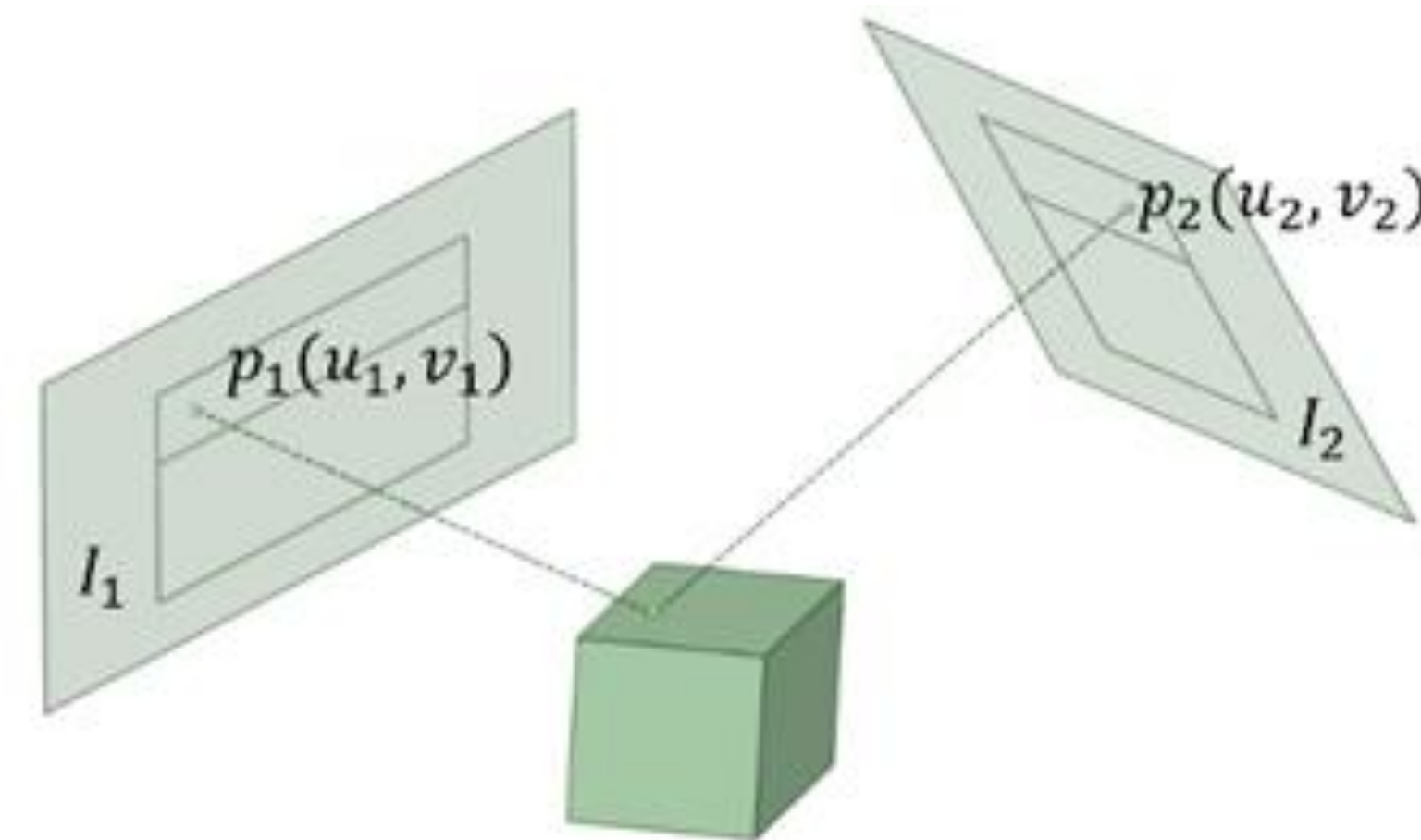
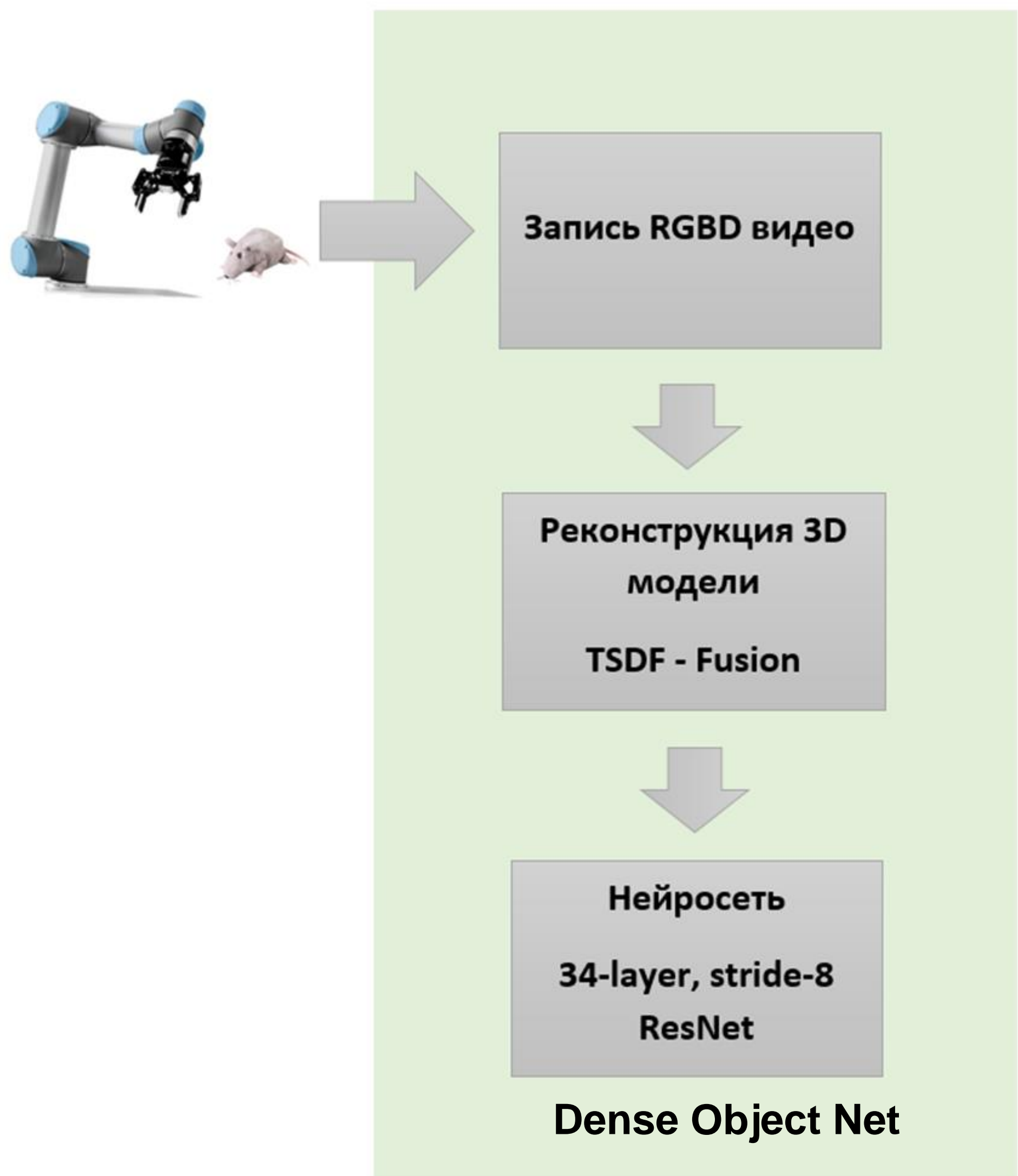


# Состав экспериментальной установки

- Коллаборативный робот манипулятор UR10
- Камера глубины **Intel RealSense 435**
- Антропоморфный захват Shunk SVH Hand
- Стол или рабочее пространство робота
  
- Компьютер **Intel Core i7**, 64Gb,  
Nvidia GeForce GTX 1080Ti
  
- Ubuntu 16.04, docker, ROS, python

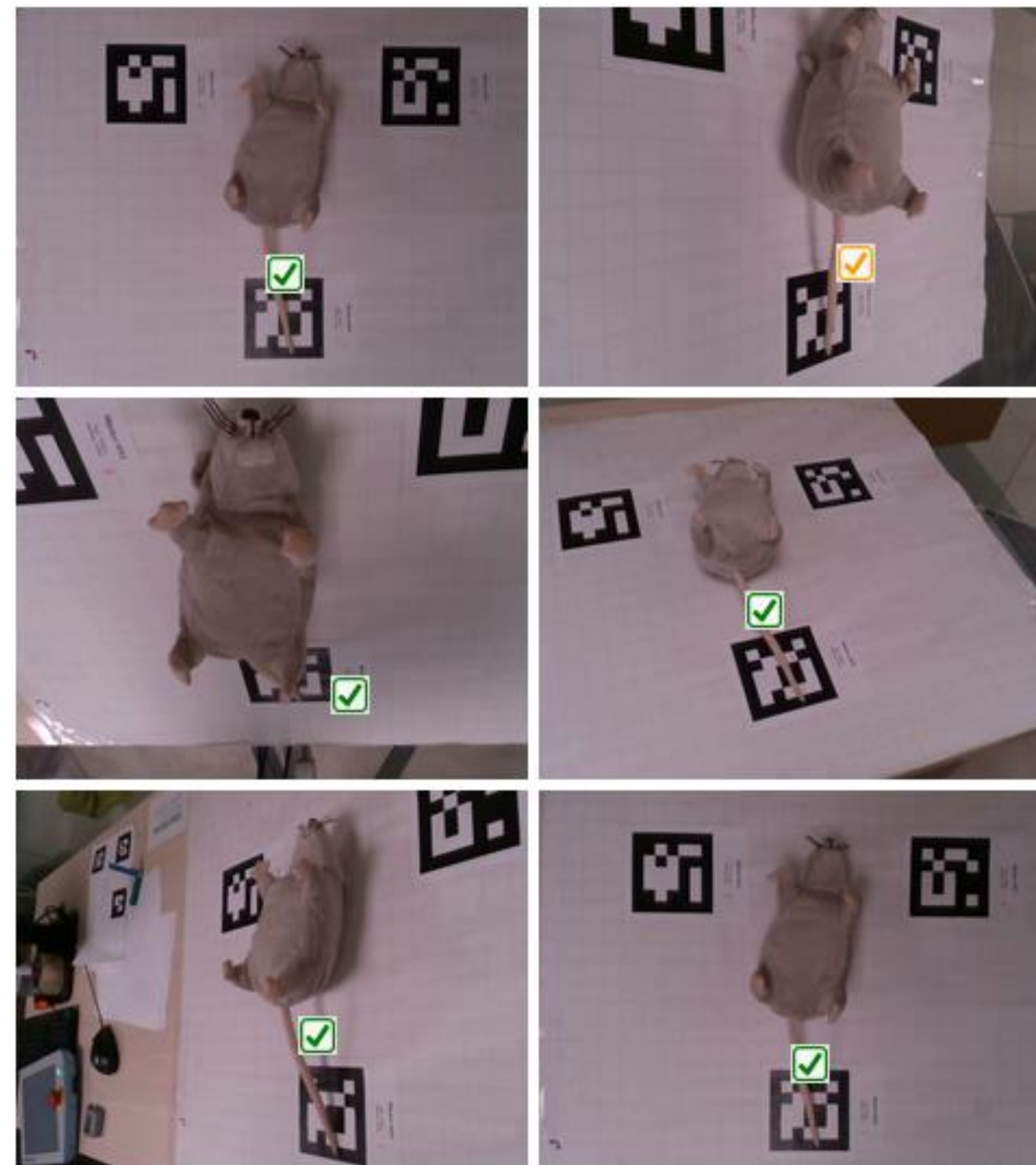
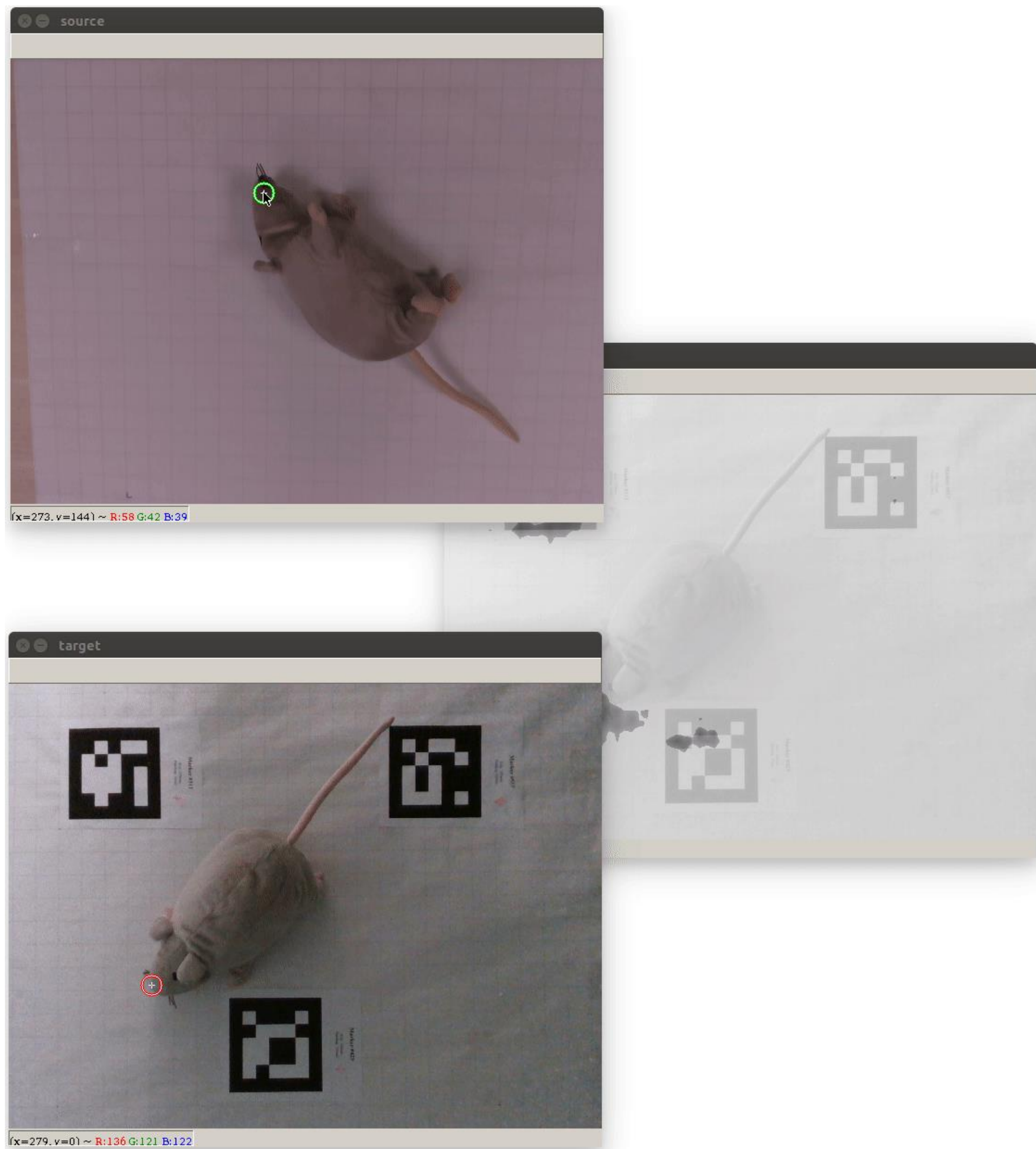


# Обучение нейронной сети



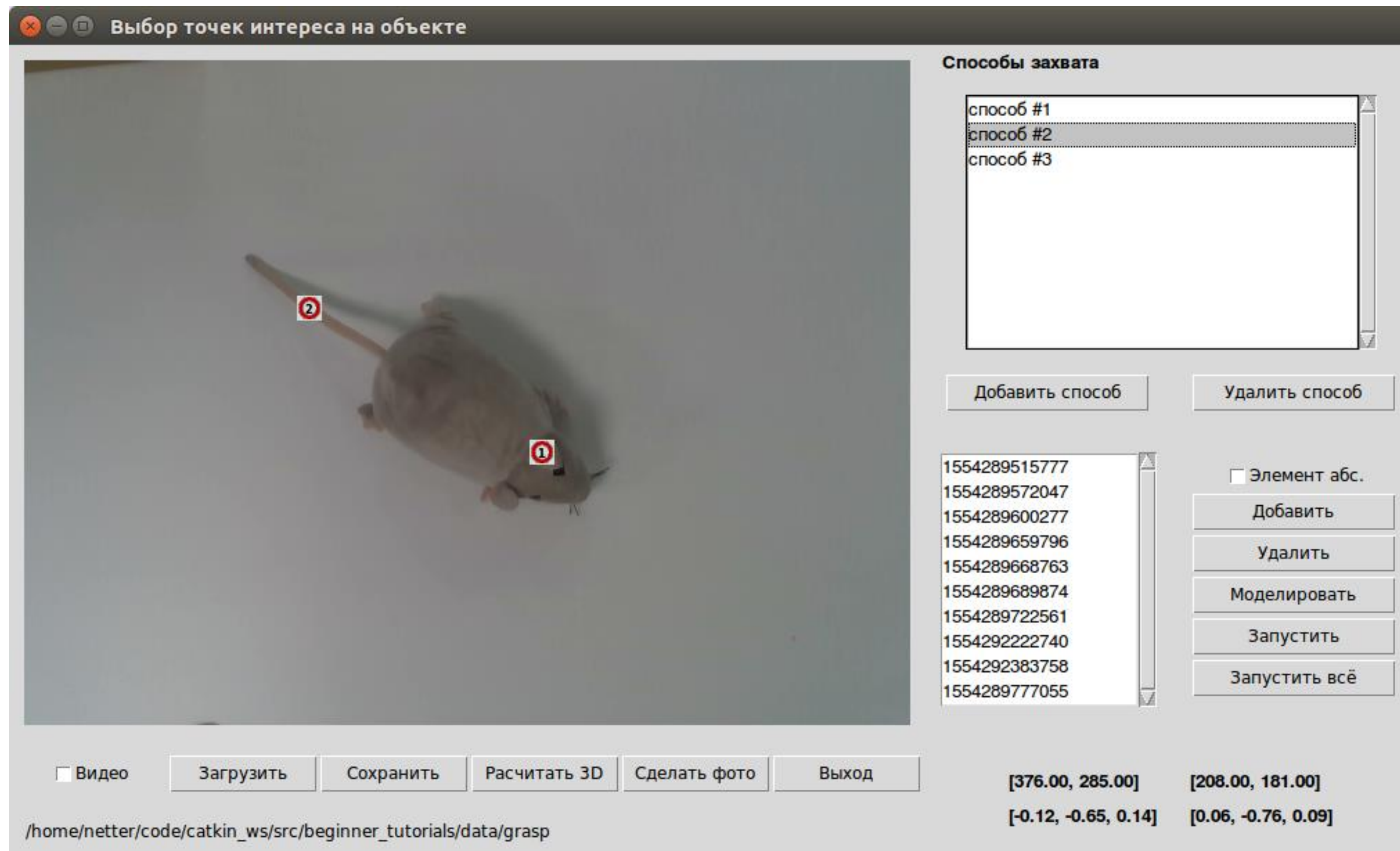
$$\text{loss}(I_1, I_2) = \frac{\sum_{n_{pos}} r(d_1(u_1, v_1), d_2(u_2, v_2))^2}{n_{pos}} + \frac{\sum_{n_{neg}} \max(0, l - r(d_1(u_1, v_1), d_2(u_2, v_2)))^2}{n_{neg}}.$$

# Визуализация отклика нейросети



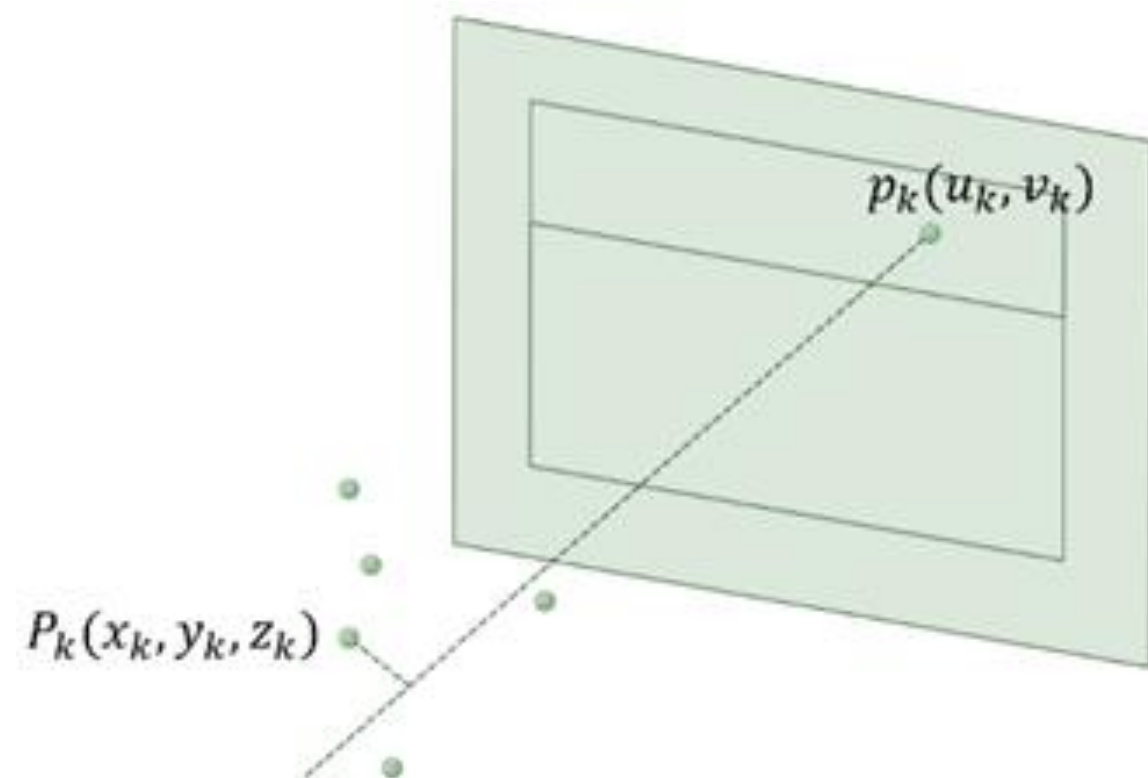
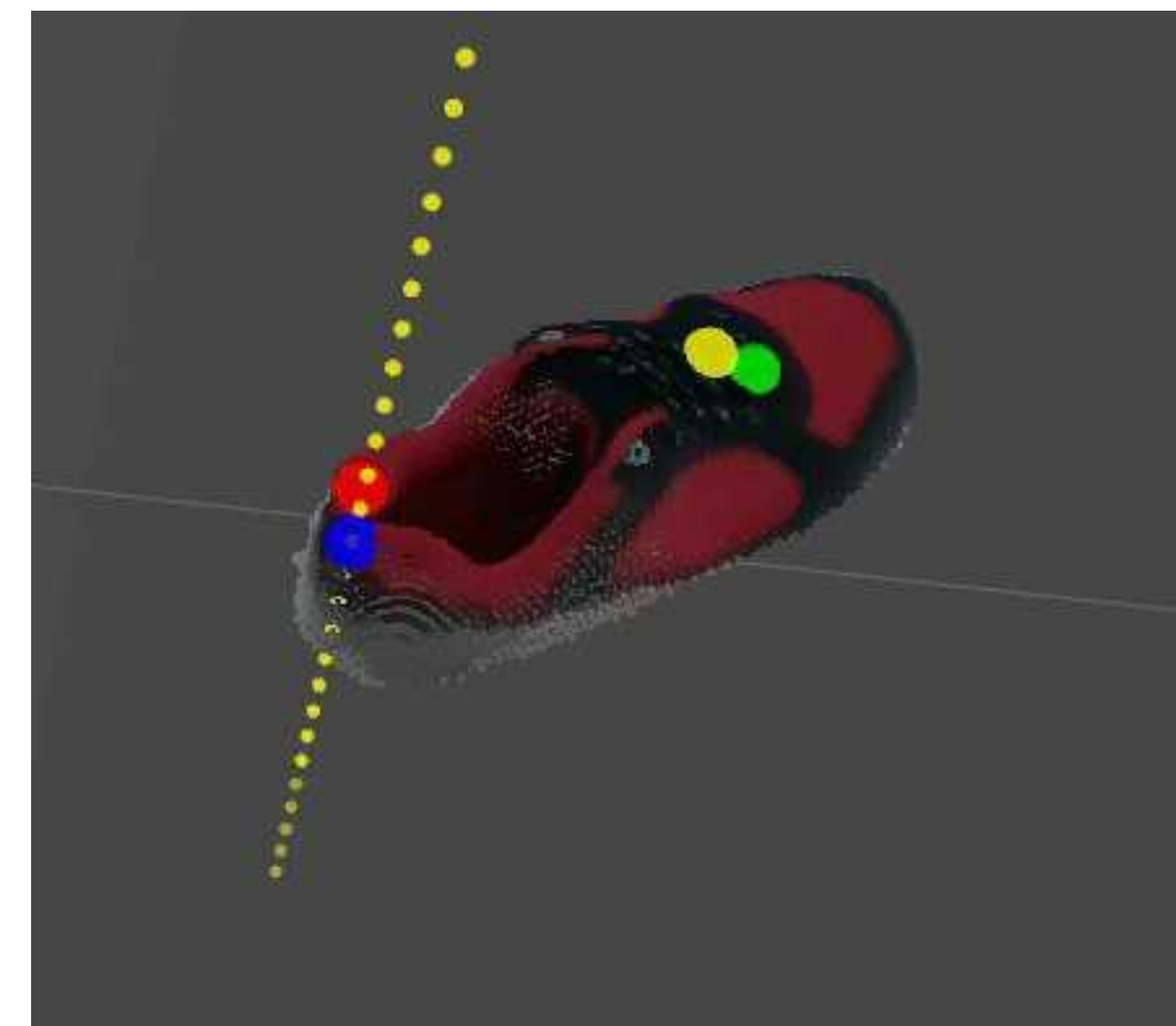


# Задание ключевых точек и способа захвата



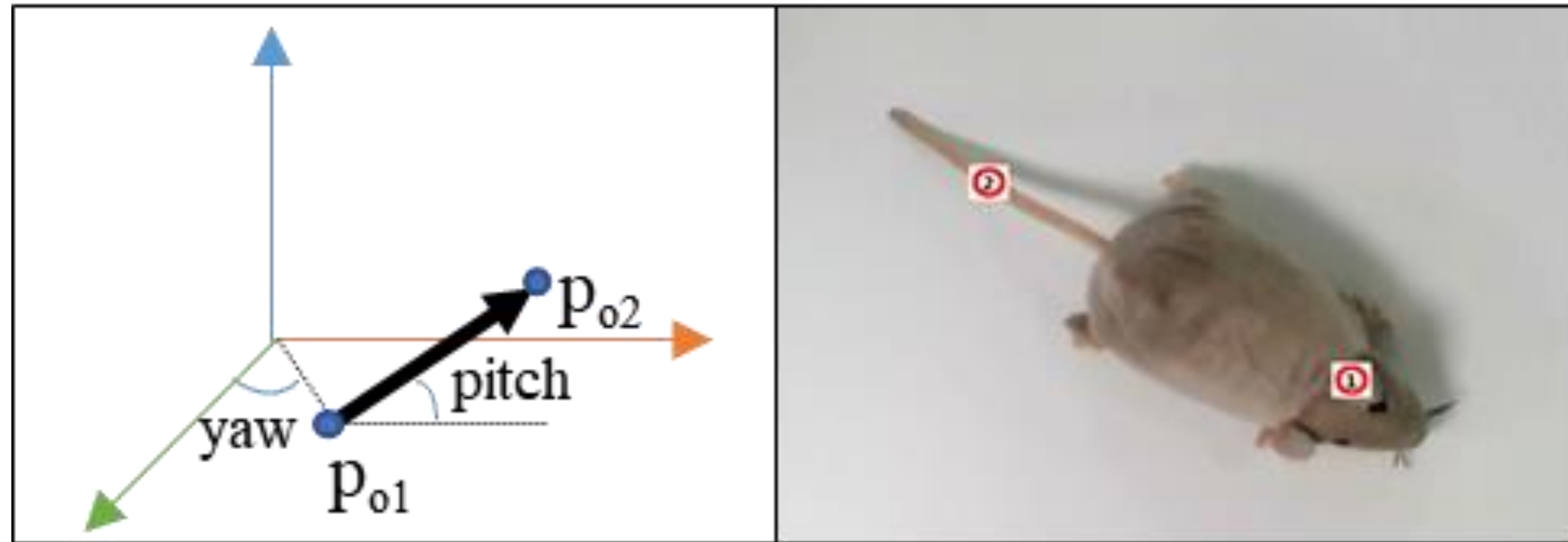
Ключевые точки задаются оператором на изображении объекта.  
Способ захвата задается оператором в режиме free-drive робота.

# Поиск и локализация ключевых точек

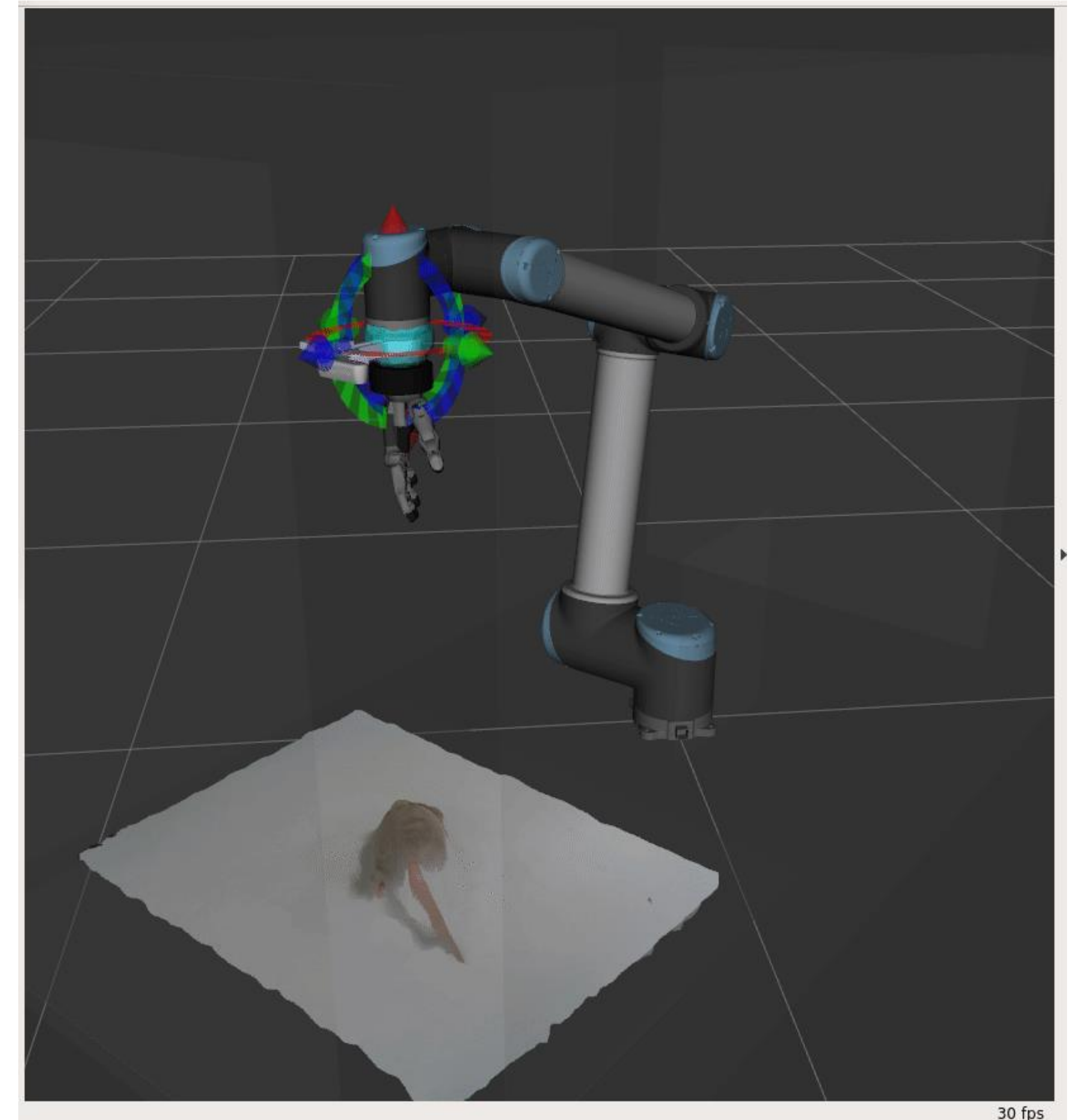


$$r_{P_{k,l}} = \min_{P_i \in C} r_{P_i,l}.$$

# Автоматический способ захвата объекта



$$J_r = [p_r, q_r], J_o = [p_{o2}, q_o];$$
$$q_o = f_{e2q}(0, pitch, yaw);$$
$$p_{ro} = (T_o^W)^{-1} * p_r, q_{ro} = \frac{q_r}{q_o};$$
$$J_{o_{new}} = [p_{o_{new}}, q_{o_{new}}];$$
$$p_{r_{new}} = T_{o_{new}}^W * p_{ro};$$
$$q_{r_{new}} = q_{ro} * q_{o_{new}}$$



# Эксперименты: мышь



Ограничения:

- использование симметричных фигур приводят к ошибкам при поиске ключевых точек;
- допущение нулевого значения угла крена приводит к не различению некоторых ситуаций;
- все ключевые точки должны быть выше опорной поверхности, чтобы избежать столкновения захватного устройства.

**Обучение проводилось только на одном экземпляре объекта.**

**Из серии 20 попыток ошибка в среднем 2 раза, успешный захват с первой попытки более 90%**

# Эксперименты: дрель



## Ограничения:

- объект очень тяжелый для хватного устройства и может смещаться во время и после захвата;
- из-за жесткости и слабой устойчивости на столе любая небольшая ошибка в захвате вызывает падение объекта;
- большое количество недостижимых позиций для расчета инверсной кинематики в силу ограничений рабочего пространства робота и габаритов объекта.

# Эксперименты: коробка



Основные ограничения:

- в силу симметричности объекта позиции найденных ключевых точек могут совпадать;
- жесткие объекты ограничивают количество приемлемых конфигураций для антропоморфной руки

# Эксперименты: кроссовки

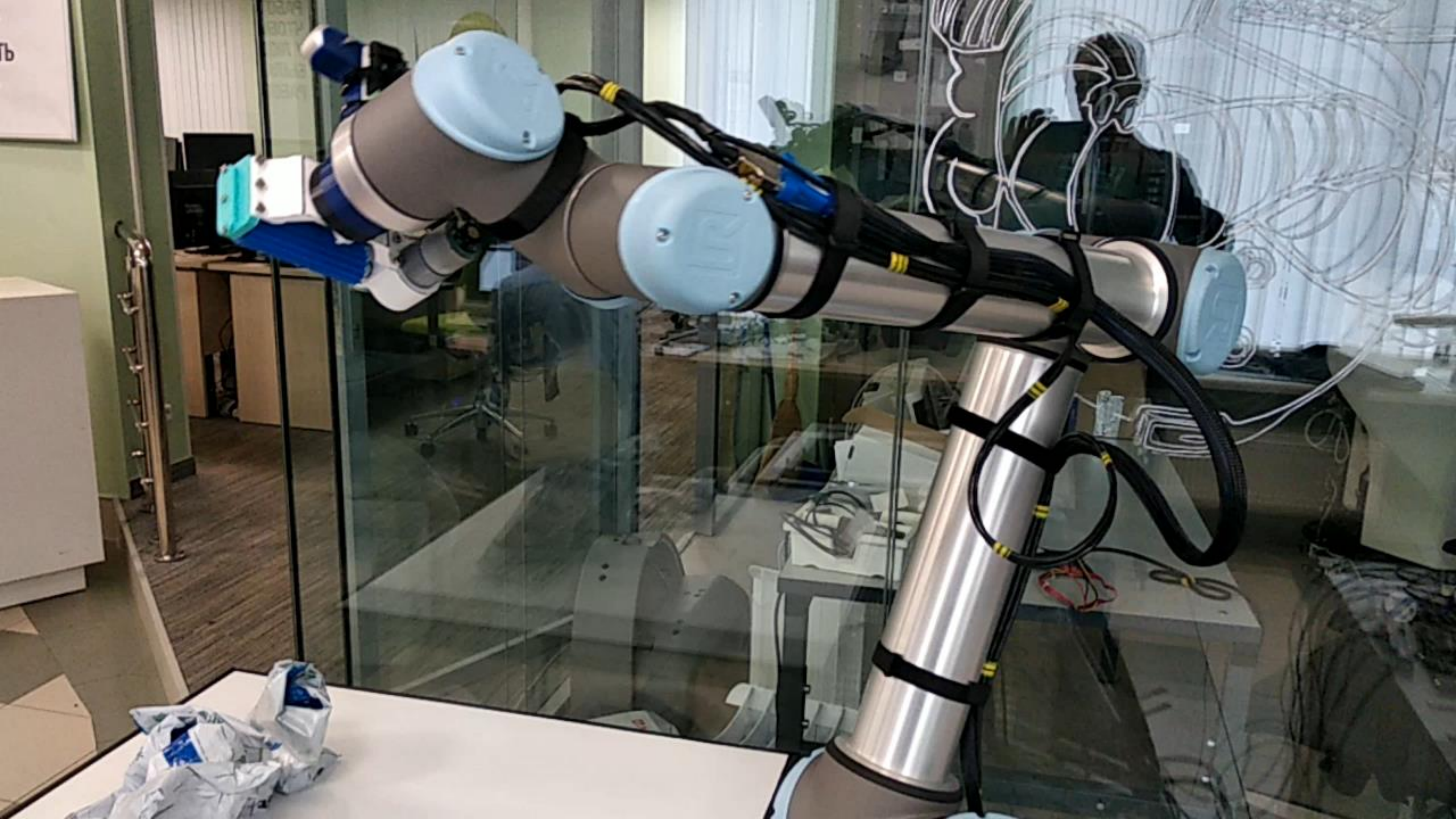


Ограничения:

- При выборе ключевых точек необходимо учитывать ландшафт поверхности объекта;

Обучение проводилось только на одном экземпляре объекта.

Успешный захват с первой попытки более 98%.





# Проект: CoinBot – практическое применение

## Эксперимент:

- Использование компьютерного зрения и машинного обучения с подкреплением
- Поиск новых методов и алгоритмов для манипулирования объектами переменной формы и веса
- Автоматизация выгрузки сейф-пакетов с монетами из подвижных тележек на стол оператора
- Обеспечение не менее 90% успешных захватов сейф-пакетов

## Intel® RealSense™ 435

## Партнерство с Microsoft:

- Совместные исследование по разработке, обучению с подкреплением и развертыванию модели BONSAI BRAIN для управления коллаборативным роботом
- Совместное исследование и разработка технологии multi-joint в симуляторе Microsoft Airsim для моделирования коллаборативных роботов манипуляторов в симуляторах в задачах обучения с подкреплением



Sberbank

Robotics

Laboratory



# Eurobot

Ведущие европейские соревнования в области робототехники для инженеров до 30 лет.

- Для участия требуется создание коллаборативных интеллектуальных логистических роботов
- Это отвечает направлениям работы Лаборатории и является идеальным полигоном для поиска кадров и отладки идей.

В системе управления командой роботов используются **Intel RealSense, Intel NUC** и Open CV:

- Распознавание положения своих и конкурирующих роботов на игровом поле
- Предотвращение столкновений и обход препятствий
- Распознавание игровых объектов и манипуляция ими





# Как мы себя видим?



«Смотри, какую прикольную штуку они сделали!»



«Мы не понимаем, как они это сделали!»

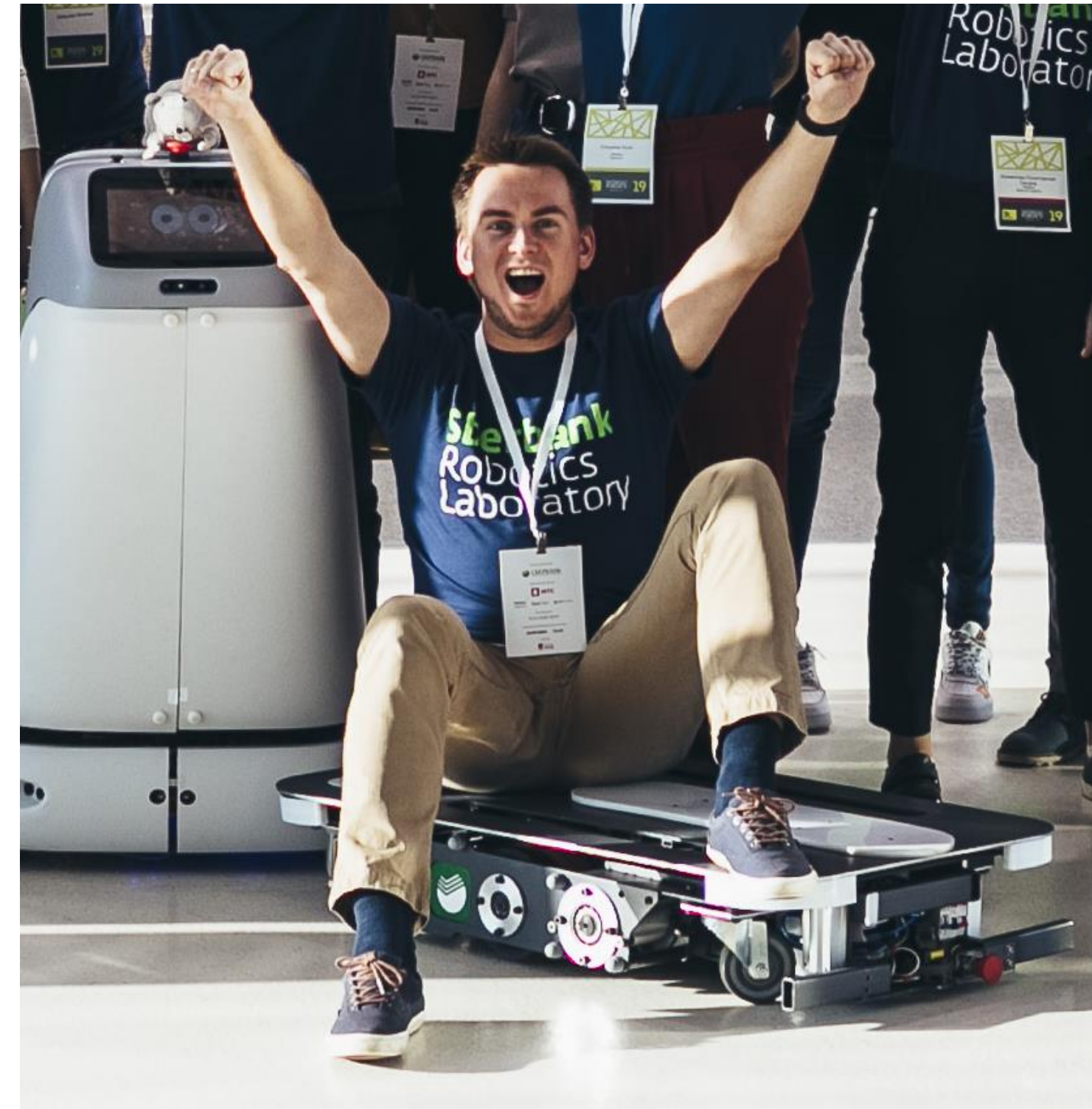
# ...Или так?



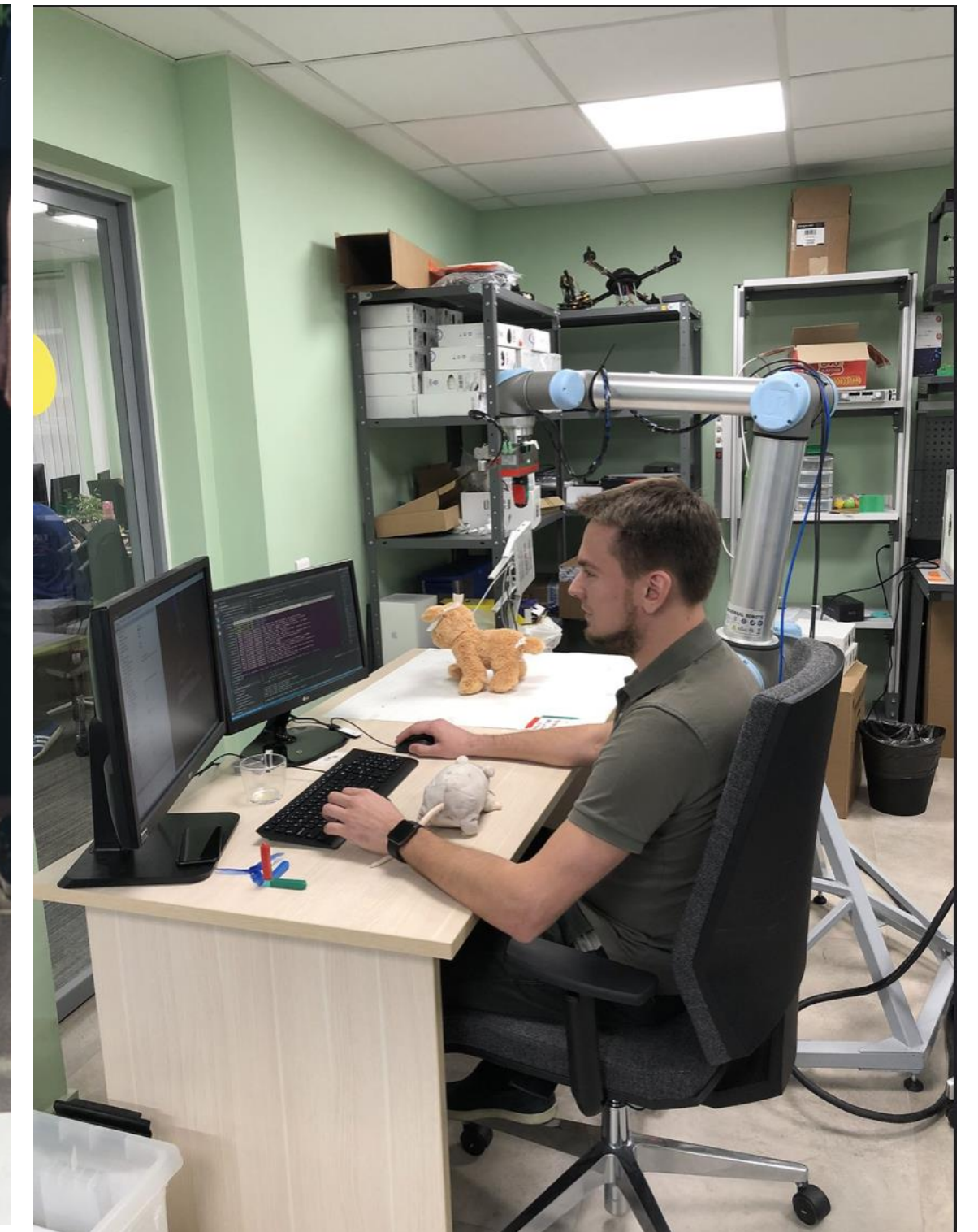
Как видят мою работу  
мои родители



Как видит мою работу моя  
жена



Как видят мою работу мои друзья

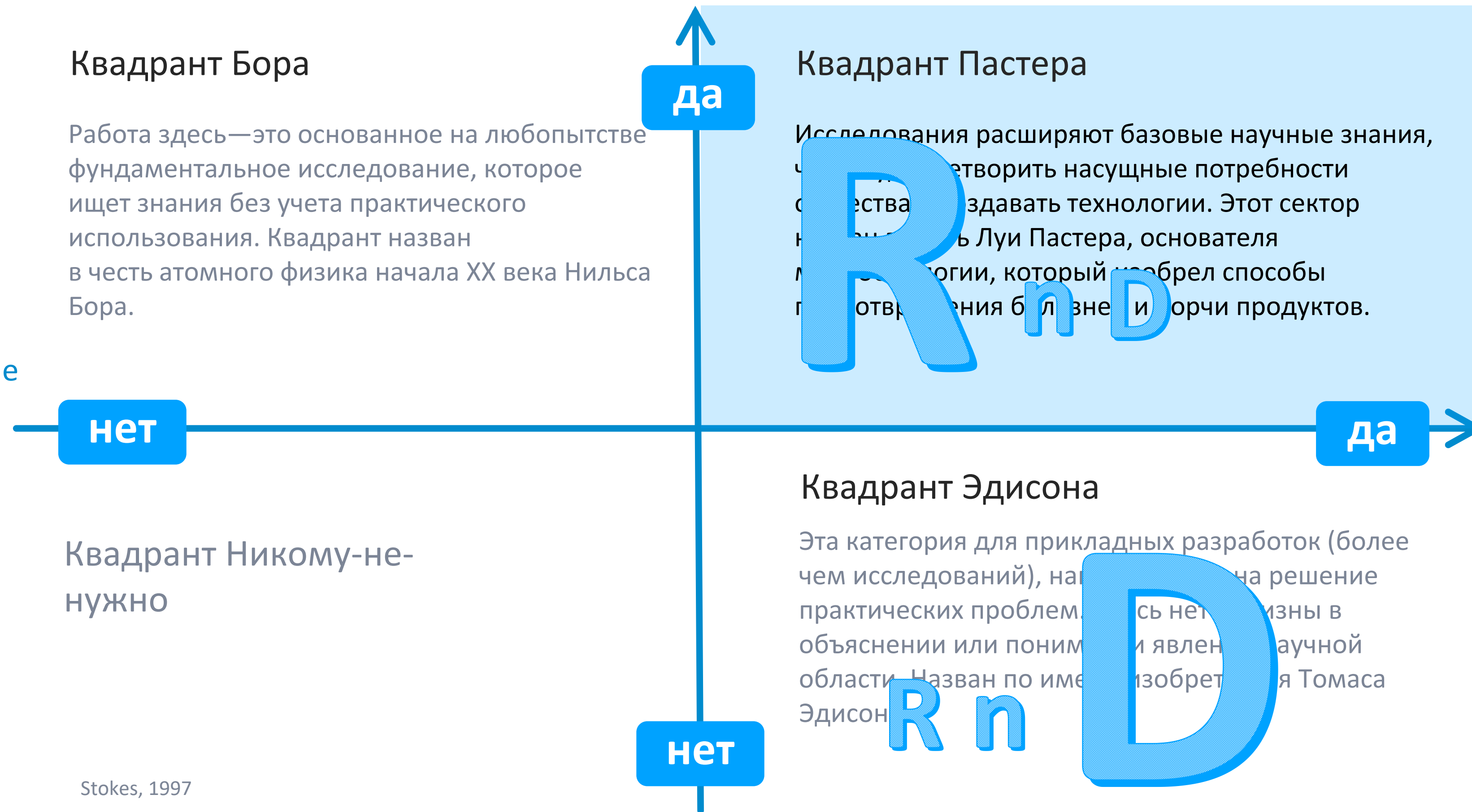


Как на самом деле

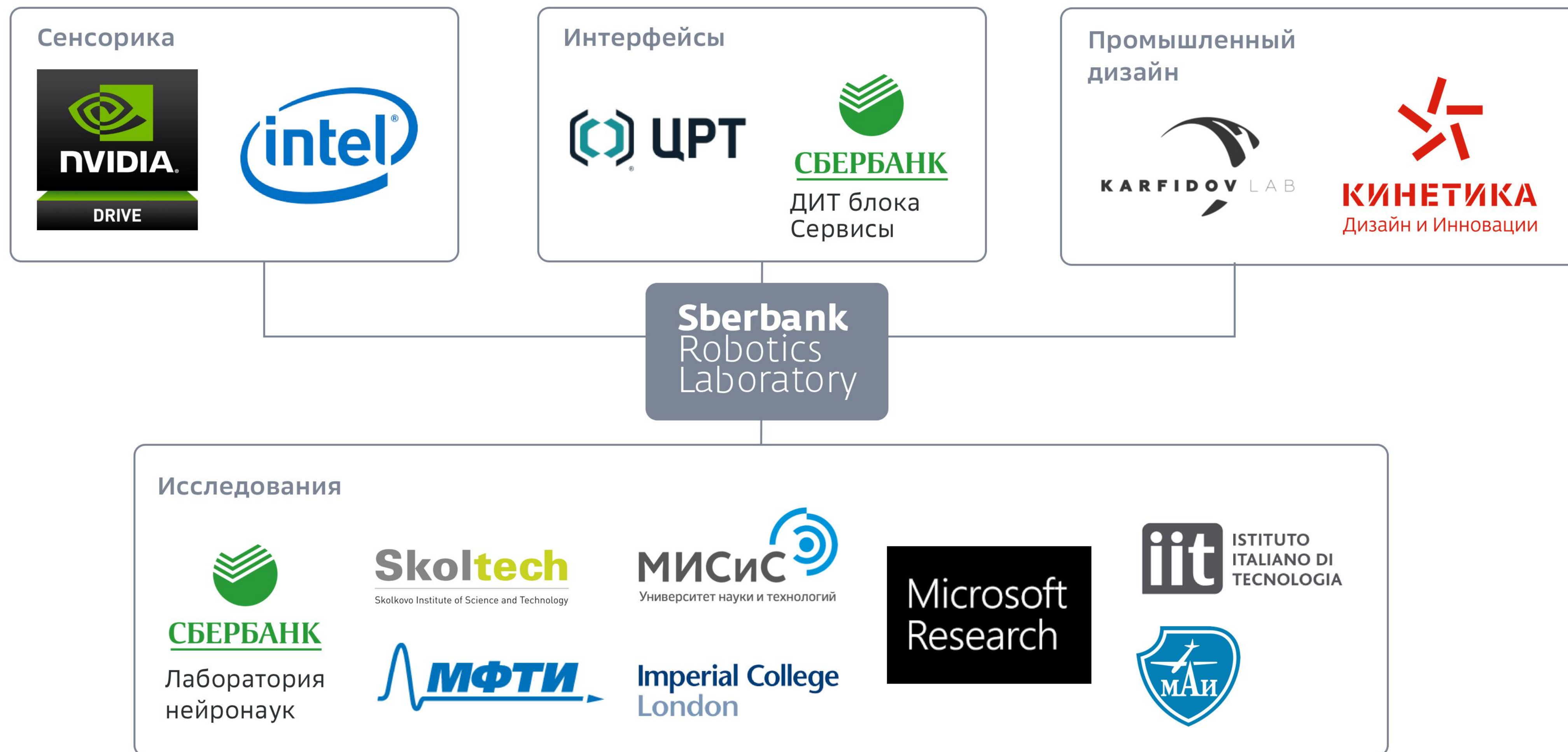
# Наша цель — полезные исследования!

Возможно ли немедленно использовать результаты?

Содержит нерешенные фундаментальные вопросы?



# Какого робота мы можем сделать вместе с ...?





**Спасибо за внимание**

**Алексей Гонноченко**  
СТО

[robotics@sberbank.ru](mailto:robotics@sberbank.ru)