



# Распознавание русского жестового языка

МОСКВА 2023



# Русский жестовой язык (РЖЯ)

РЖЯ — это национальная языковая система, имеющая уникальную лексику и грамматику, применяемая для взаимодействия среди глухих и людей с нарушениями слуха, проживающих в России

## Проблема

**>2 млн**

глухих  
и слабослышащих

**~1 тыс**

переводчиков  
с русского жестового  
языка

## Цель

**Создать сервис**

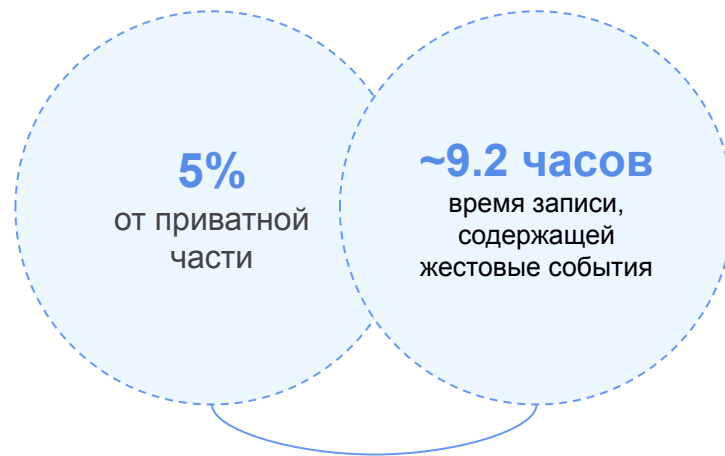
для распознавания русского жестового языка  
по видеозаписи на основе моделей глубокого обучения

# Slovo — дата сет от Сбер

Высококачественные размеченные данные  
русского жестового языка



**Приватный датасет**



**Публичный датасет**

# EDA



**15 тыс.**  
тренировочных  
видео



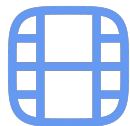
**5 тыс.**  
тестовых видео



**100**  
жестов



**20**  
экземпляров  
на каждый класс



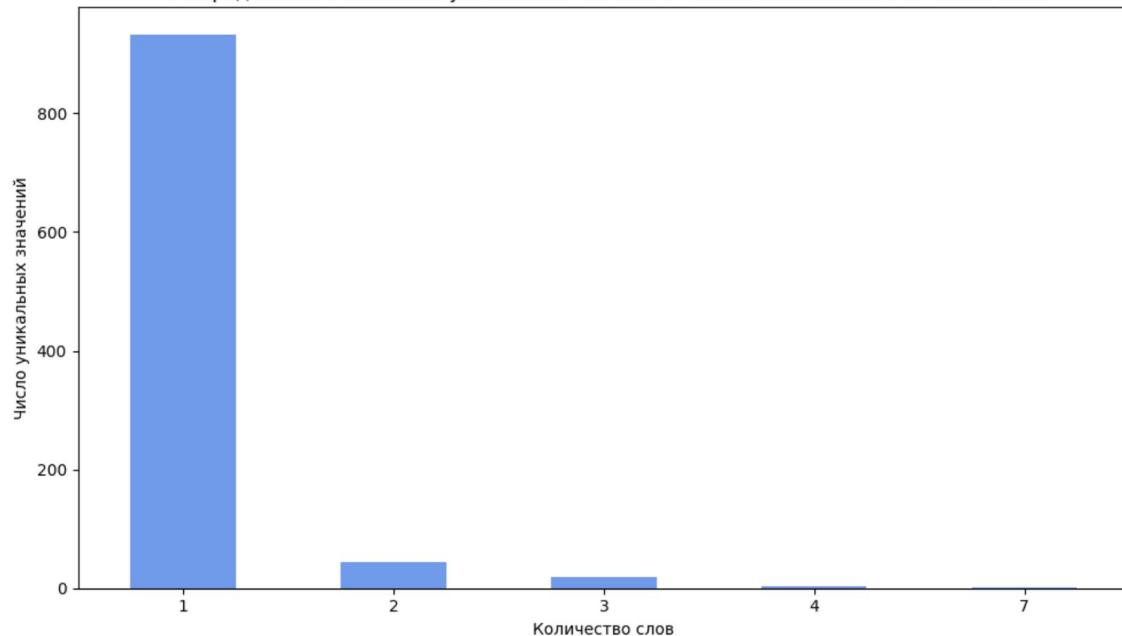
**~50**  
среднее количество  
кадров



**47**  
медианная длина  
видео

# Распределение слов в жестах

Распределение количества уникальных значений текста в зависимости от числа слов



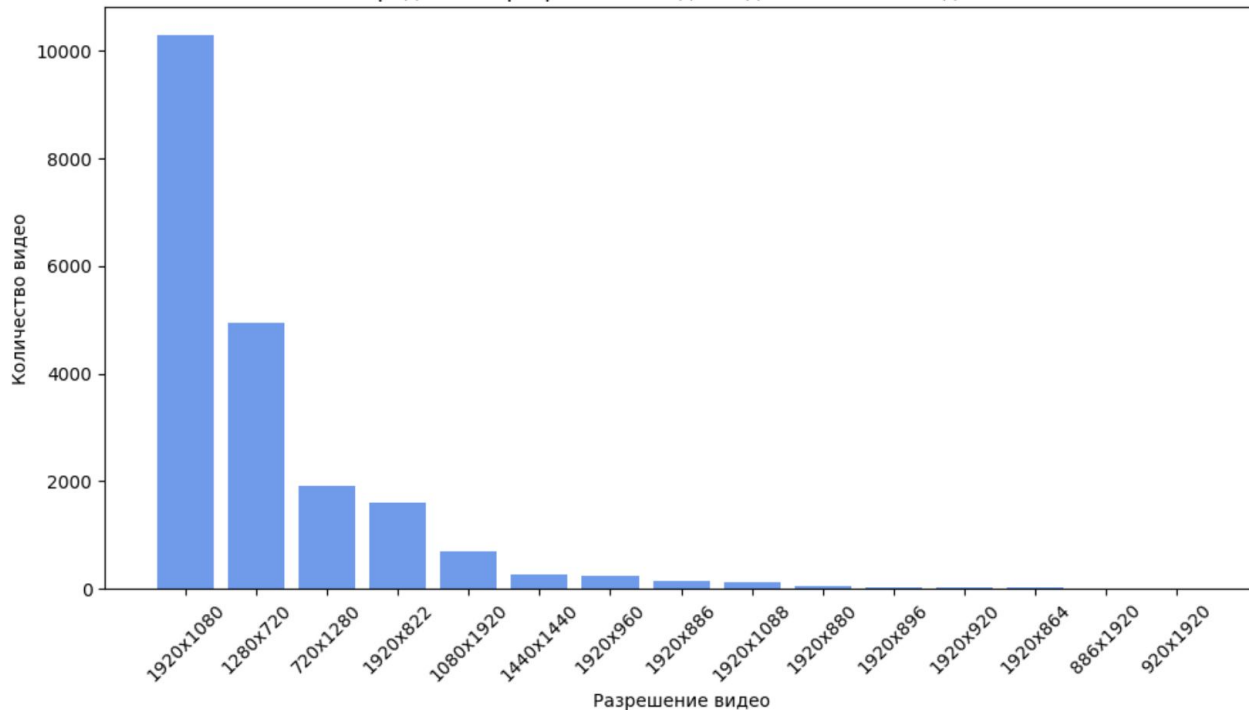
большая часть жестов  
состоит из 1-го слов

есть классы состоящие из  
2 и более жестов

есть жест, состоящий из 7-и  
слов

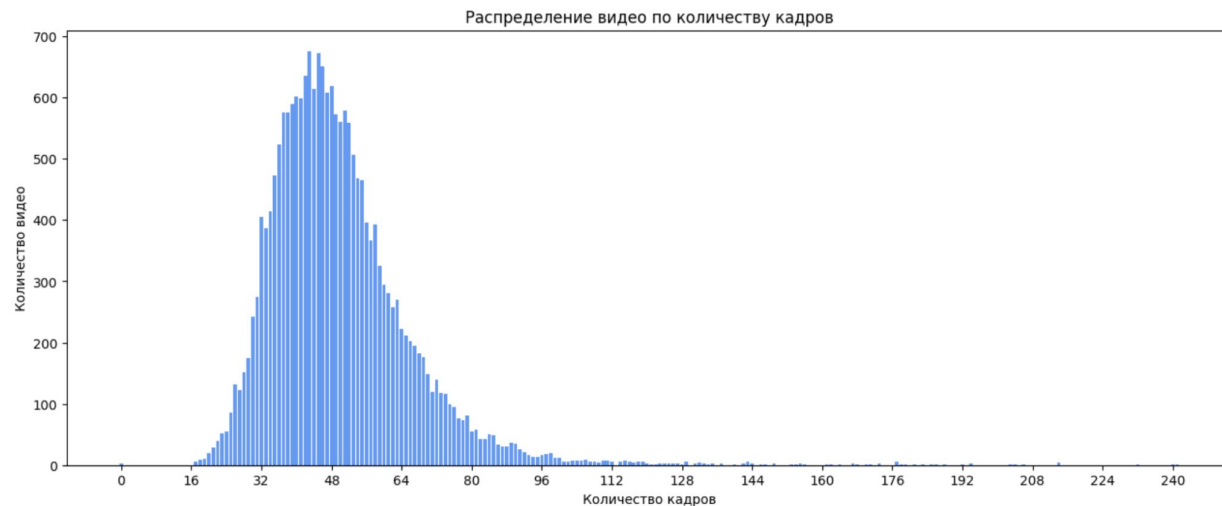
# Распределение разрешений видео

Распределение разрешений видео (где количество видео > 5)



большинство видео  
в данных сняты в  
стандартных  
разрешениях:  
1920x1080, 1280x720

# Распределение количества кадров



количество кадров  
колеблется около  
~50 кадров

распределение  
немного скошено  
вправо

# Участники записей

194

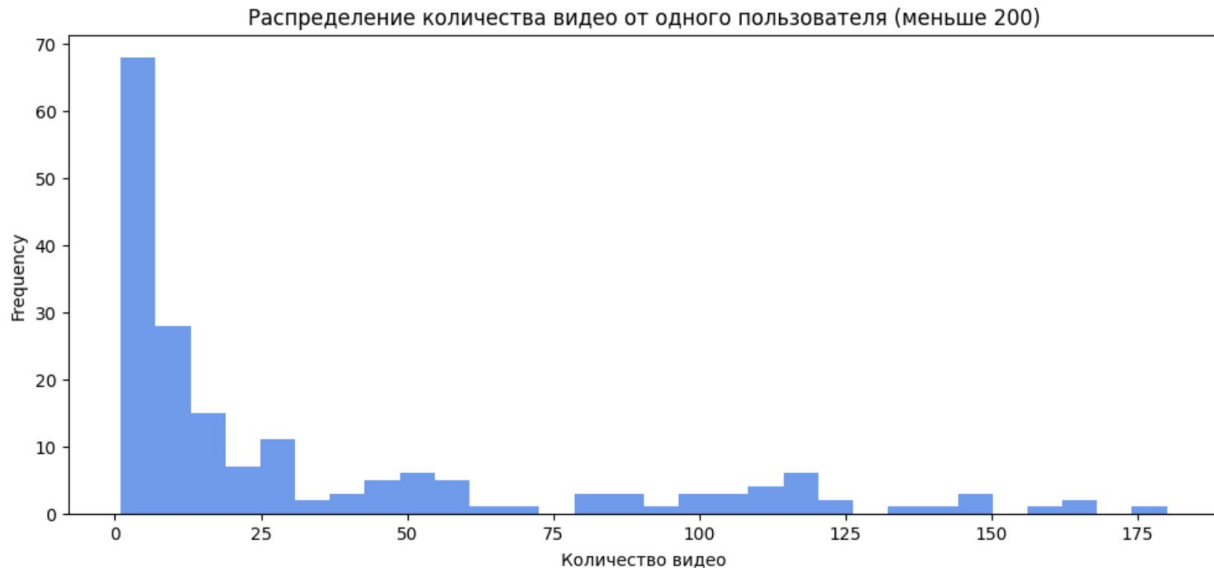
уникальных  
пользователей

13

медианное  
количество видео  
от пользователя

105

среднее  
количество видео  
от пользователя

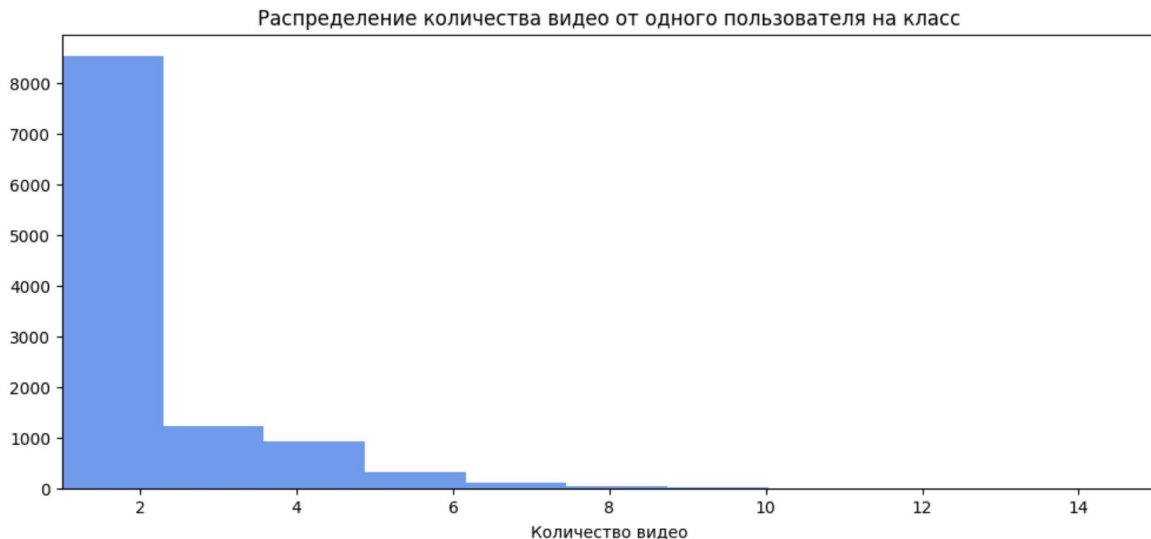




# Количество записей на класс

~1.8 среднее  
количество кадров  
записей на класс

есть пользователей с  
большим  
количеством  
попыток записи



# Модели

**Easy Sign**

Предобученная модель  
для распознавания РЖЯ

**MediaPipe Hand  
Landmarker**

Предобученная модель для  
разметки позиций рук

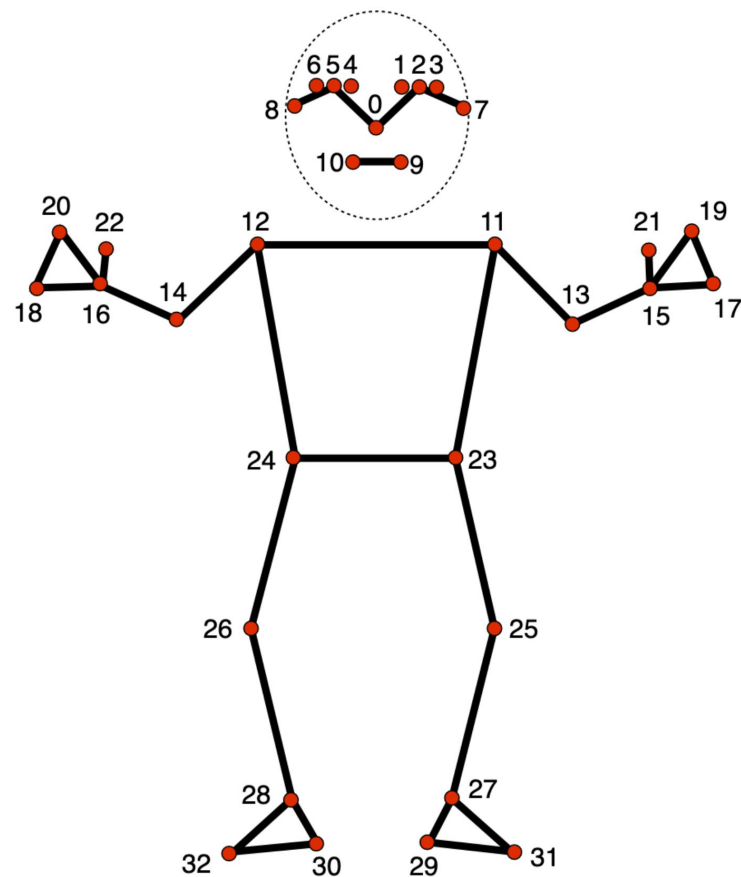
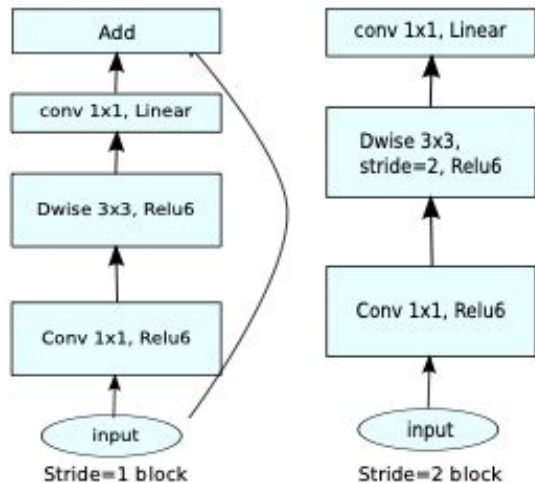
**MediaPipe Body  
Landmarker**

Предобученная модель для  
разметки позиций тела

# Разметка позы

## MediaPipe Pose Landmarker

CNN: MobileNetV2-like



# Разметка позы (пример)



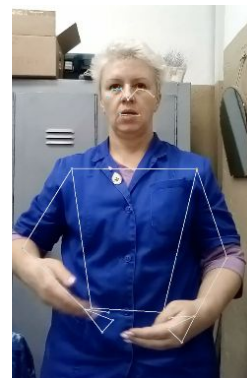
Model

Pose Landmarker

Json

```
{
  "pose_landmarks": [
    {
      "x": 0.5004,
      "y": 0.3441,
      "z": -1.1202
      // ...
    },
    // ...
  ],
  "pose_world_landmarks": [
    {
      "x": -0.0210,
      "y": -0.5436,
      "z": -0.4230
      // ...
    },
    // ...
  ],
  "segmentation_masks": null
}
```

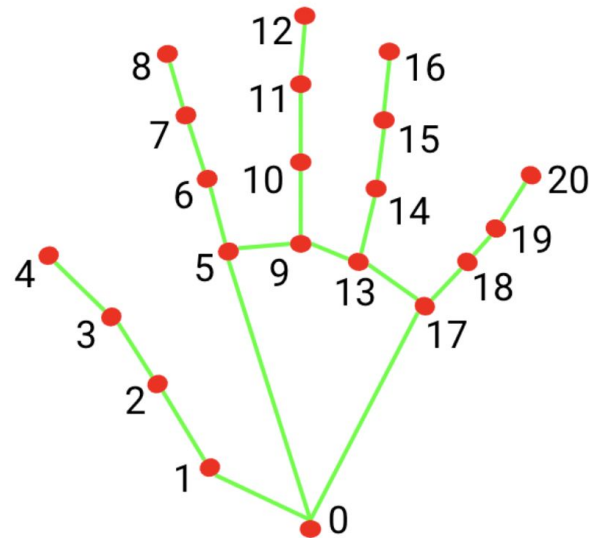
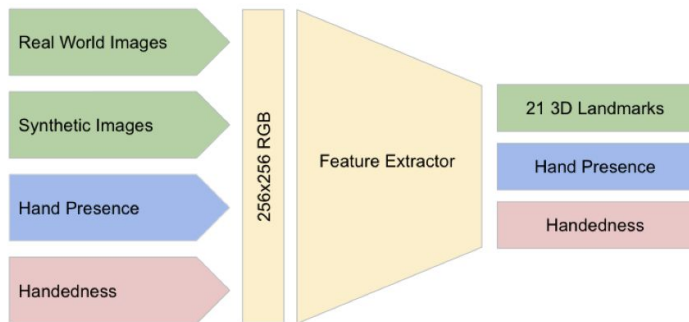
Наложенная  
разметка



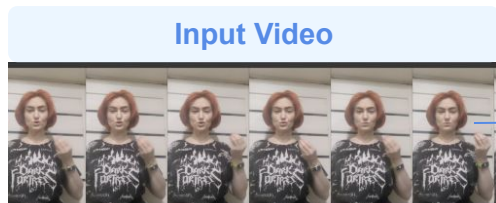
# Разметка рук

## MediaPipe Hand Landmarker

Двухэтапный пайплайн, состоящий из нейронной сети с one-shot детектором и последующей регрессионной моделью, работающей на обрезанной области



# Разметка рук (пример)



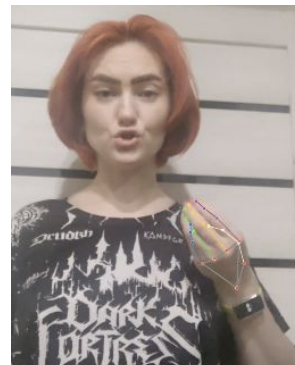
Model

Hand Landmarker

Json

```
{
  "handedness": [
    {
      "index": 1,
      "score": 0.9969,
      "display_name": "Left"
    }
  ],
  "hand_landmarks": [
    {
      "x": 0.8520,
      "y": 0.6402,
      "z": -0.0000003
    }, // ... ]
  ]
}
```

Наложенная  
разметка



# Easy Sign модель

## S3D

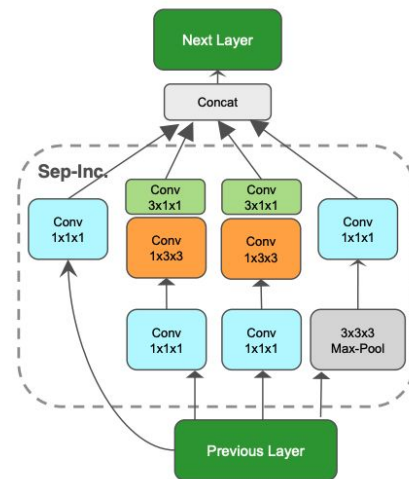
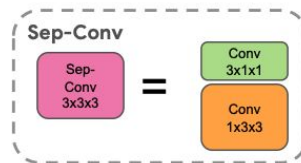
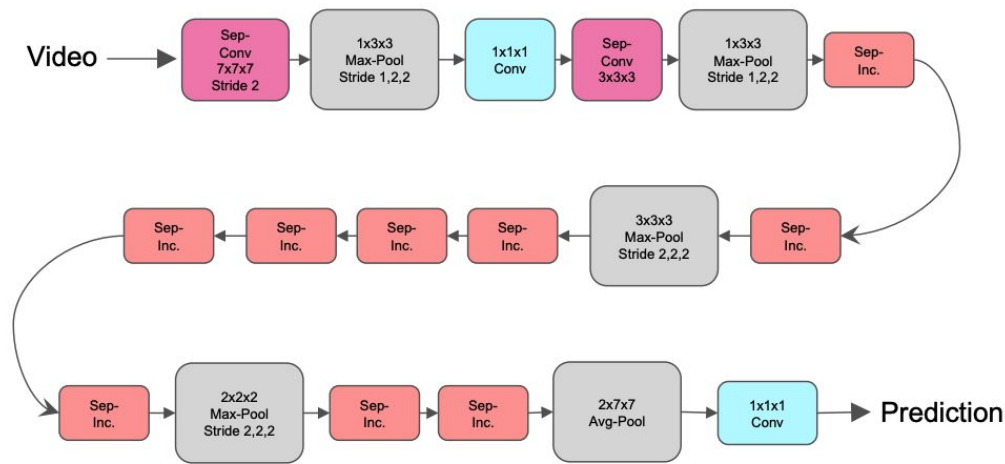
В основе лежит S3D предобученная на датасете Kinetics400\*

## Slovo

Дообучена на датасете Slovo

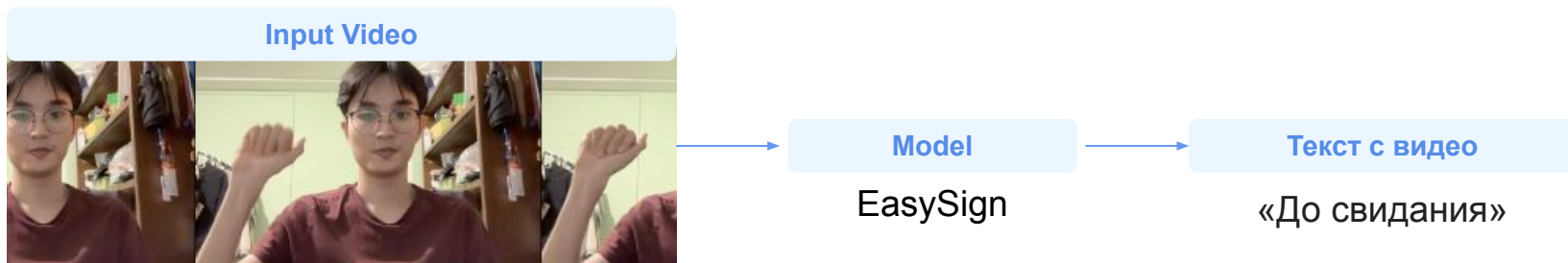
## Скорость

На CPU - 2-2.5 PPS



# Распознавание РЖЯ

(пример)





## Метрика

$$mAcc = \frac{1}{N} \sum_{i=0}^{N-1} \frac{TP_i}{TN_i + FN_i}$$

**mean accuracy**

макро усреднение  
точности по всем классам

## Качество

Количество кадров	mAcc, %
32	44.22
48	52.28
64	55.86

# Сервис распознавания РЖЯ

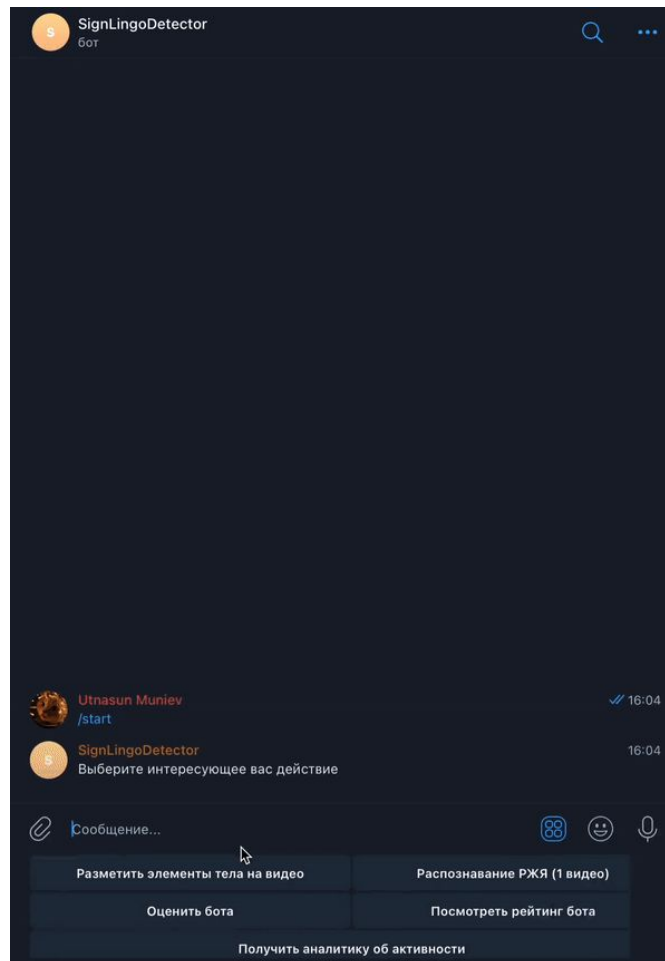


# Распознавание РЖЯ

Распознавание и интерпретация русского жестового языка из загруженных пользователями видео

## Результат:

— сообщение с распознанным текстом РЖЯ с видео



# Определение ключевых точек поз

Анализ видео на предмет определения позы тела, предоставляя информацию о положении тела и движениях

## Результат:

- видео с размеченными точками позы;
- файл с разметкой в формате json

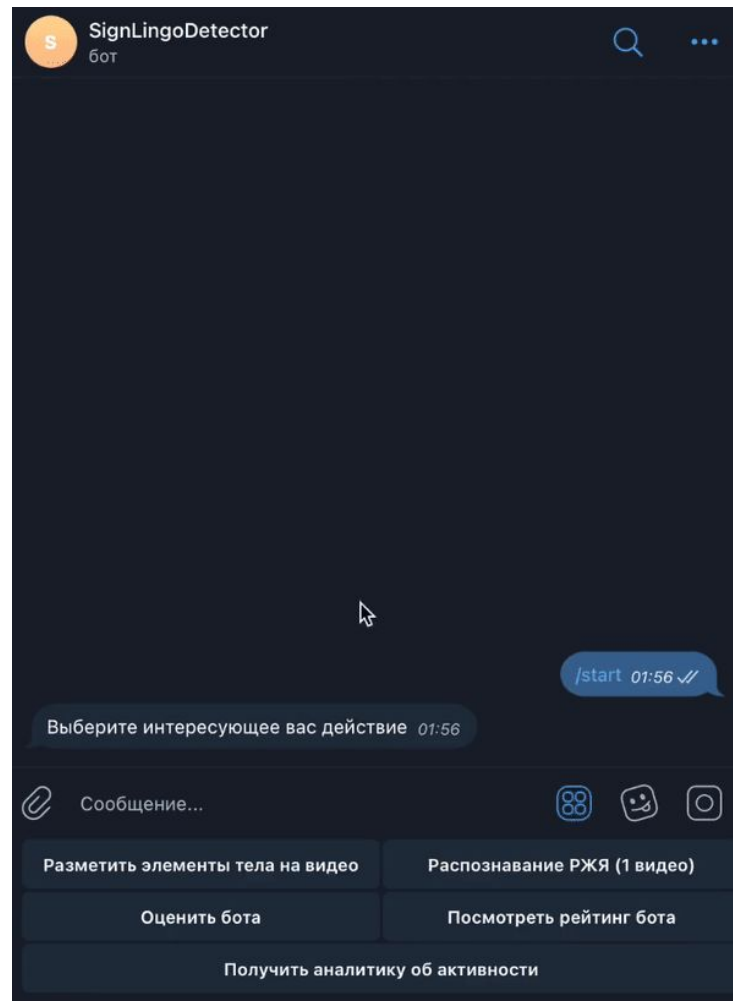


# Обнаружение ключевых точек рук

Анализ положения рук,  
предоставление информации  
о положение рук

## Результат:

- видео с размеченными точками рук;
- файл с разметкой в формате json



# Структура проекта

**data**

Папка для данных

**hse\_slr**

Модули распознавания объектов

**notebooks**

Исследования

**slr\_bot**

Telegram-бот

**tests**

Юнит-тесты для бота