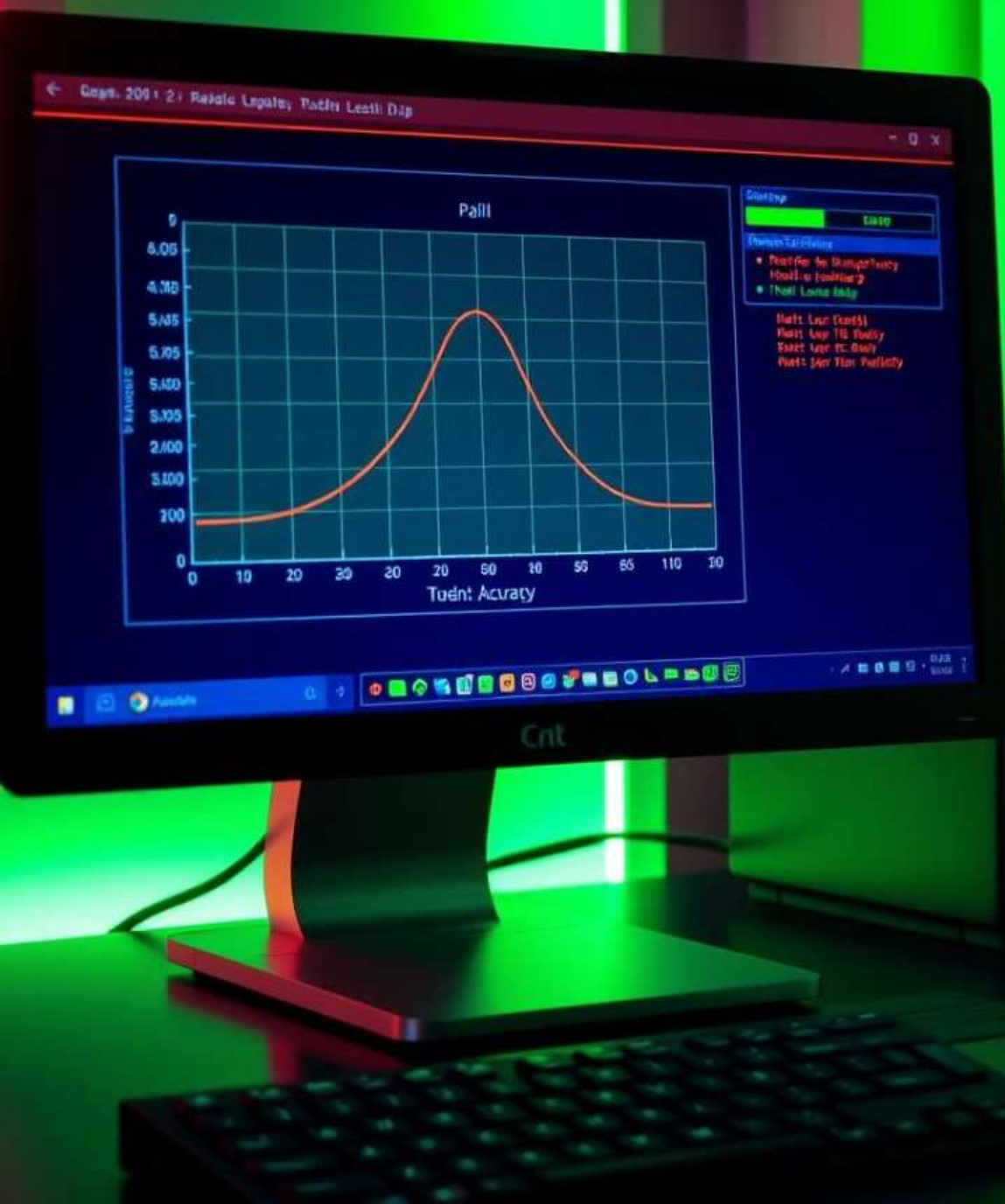


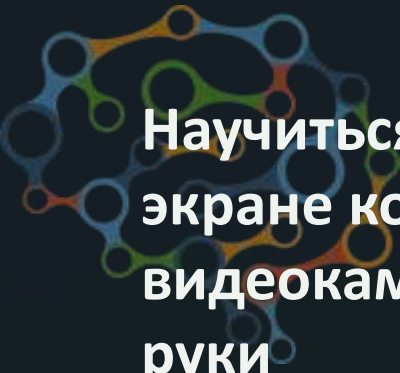
Создание высокоточной модели для распознавания жестов рук через веб- камеру и управления курсором мышки на экране

Научно-исследовательский проект

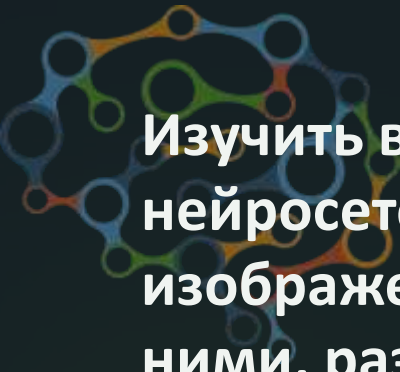
Выполнил Максим Парьев, 7 И



Идея проекта



Научиться управлять мышкой на экране компьютера через видеокамеру перемещением своей руки



Изучить возможности разных нейросетей по обучению по изображениям, научиться работать с ними, размечать изображения

Актуальность



Упрощение взаимодействия:

Использование жестов для управления курсором может значительно упростить работу с устройствами, особенно в ситуациях, когда традиционные устройства ввода, такие как мыши или клавиатуры, не удобны или недоступны.



Технологические инновации:

Проект стимулирует развитие технологий распознавания жестов, что может привести к улучшению существующих систем и созданию новых, более сложных интерфейсов.



Виртуальная и дополненная реальность:

Управление курсором жестами имеет особое значение в контексте VR и AR, где традиционное управление не всегда эффективно. Это открывает новые возможности для взаимодействия с виртуальными мирами.



Цели

1

Провести исследование возможности создания высокоточной модели для распознавания жестов рук через веб-камеру;

2

Выбрать подходящую для проекта модель;

3

Обучить модель на уникальном датасете, состоящем из фото рук с разными жестами;

4

Разработать программу управления курсором мышки с помощью рук на языке Python;

5

Исследовать возможность развития проекта и применения его в реальной жизни;



Задачи

1

Сбор и подготовка данных:

Собрать датасет из множества самодельных фото;

Разметить датасет;

Подготовить датасет к обучению;

2

Обучение модели:

Обучить модель YOLO на подготовленных данных, оптимизируя параметры модели для достижения наилучшей точности;

3

Тестирование модели:

Протестировать модели и узнать ее точность.

4

Разработка программы для управления мышкой;

5

Оптимизация производительности.

Сбор и подготовка данных



Подготовка фото

Создание большого количества фото с рукой и формирование датасета



Разметка фото

Разметка фото в программе Labellmg. Формирование XML файла с координатами прямоугольника с рукой.



Преобразование координат на Python

Координаты нормализованы под формат монитора и преобразованы в TXT файл

0	0.9375	0.30546875	0.11484375	0.48125
---	--------	------------	------------	---------

Класс

Координаты центра
bounding boxes

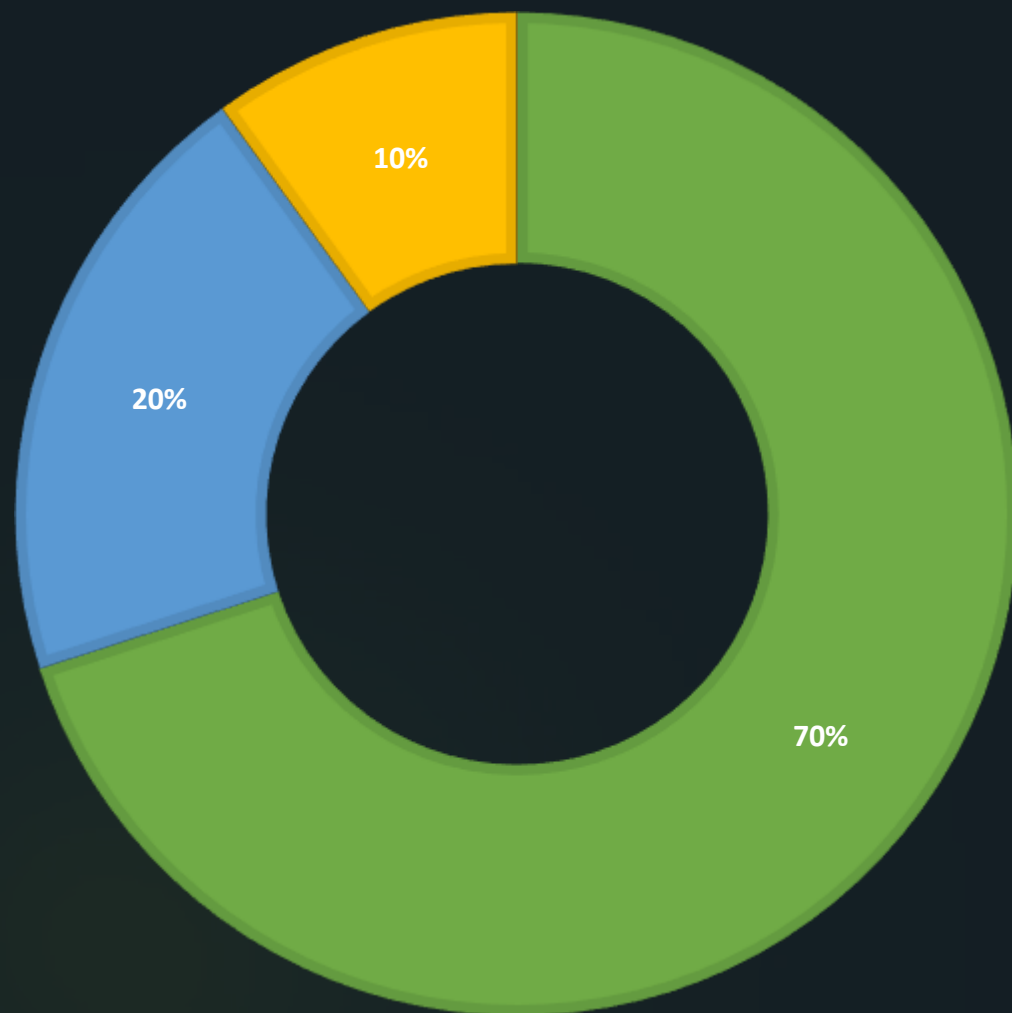
Длина ширины и высоты блока



Сбор и подготовка данных

РАЗБИЕНИЕ ДАТАСЕТОВ

■ Тренировочный ■ Валидационный ■ Скрытый (тестовый)



Пример размеченного фото



Принцип разметки



YAML файл с указанием путей к датасетам

```
train: C:/PycharmProjects/pythonProject1/datasets/train
val: C:/PycharmProjects/pythonProject1/datasets/valid
test: C:/PycharmProjects/pythonProject1/datasets/test
nc: 1
names: [hand]
```



Выбор модели для обучения

Требования к модели

1

Быстрое обучение для оперативных проверок изменений разметки датасетов и их корректировки

2

Быстрая обработка картинок для того, чтобы обеспечить минимальные задержки в обработке при захвате видео и определения координаты мышки

1

Самостоятельное создание модели

Реализовывать сложно, долго, не хватило знаний. Решил оставить в качестве развития проекта

2

Модель YOLO на Google collab

Удобная в использовании, но очень долгое обучение (1 эпоха 4 минуты), не хватало времени на переобучение и проверку настроек датасета

3

Модель ROBOFLOW

Обучалась быстро и точно (1 эпоха 1 минута), но при попытке использовать в процессе обработки видео через REST API была большая задержка по времени

4



Модель YOLO на локальном ПК

Быстро и точно обучалась (1 эпоха 1,5 минуты), быстрая обработка видео

Обучение модели

1

Выбор оптимального количества эпох

Выбрано 400 эпох – это оптимальное количество.

Меньше – будет недостаточная точность

Больше – может возникнуть эффект «переобучения модели»

2

Выбор координат разметки

При первой попытке разметки на «левый верхний + правый нижний угол» точность была небольшая (около 40%)

При изменении координат на «левый нижний + правый верхний угол» точность возросла до >98%

3

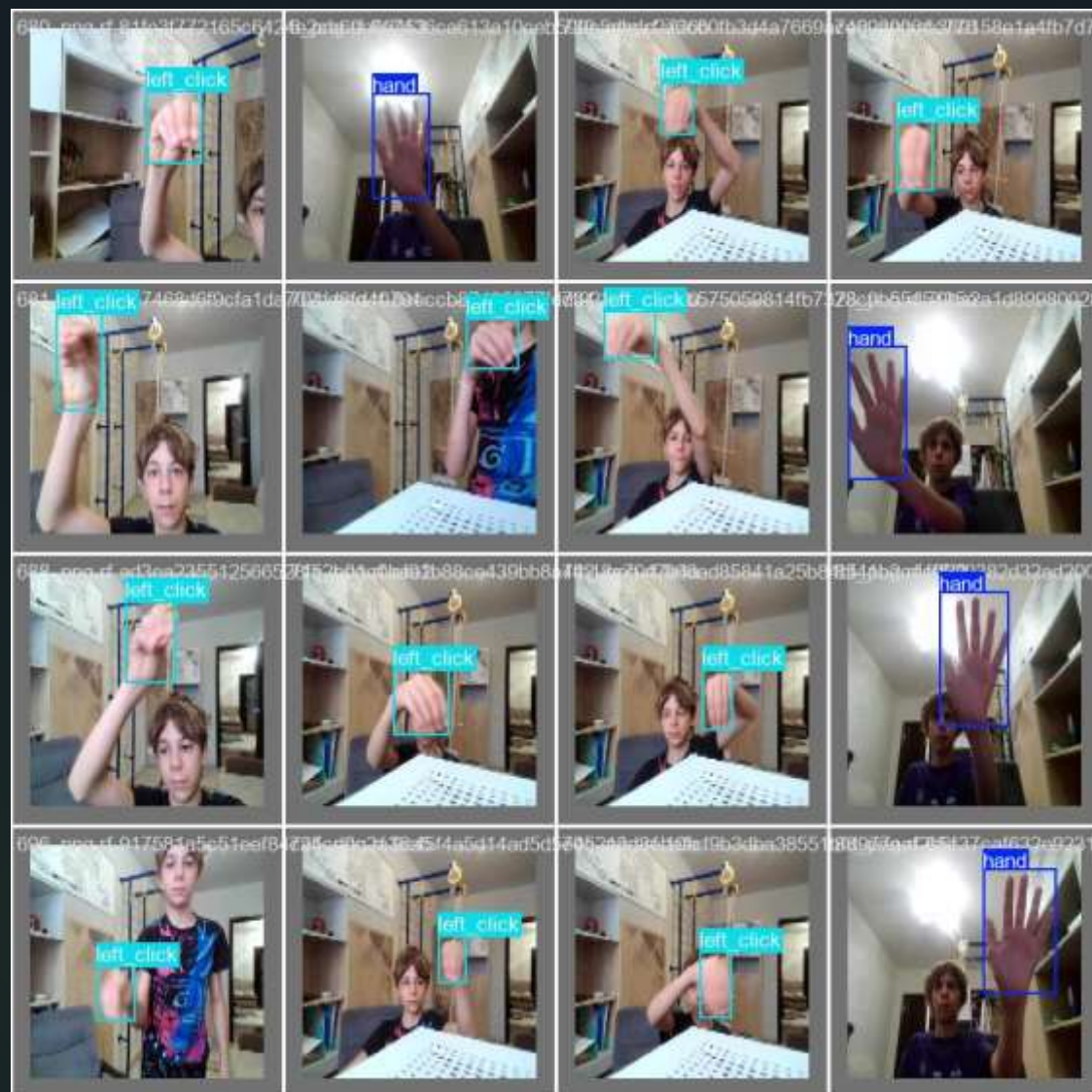
Запуск обучения

Для запуска обучения была создана программа на Python, обучения было с использованием библиотеки yolo

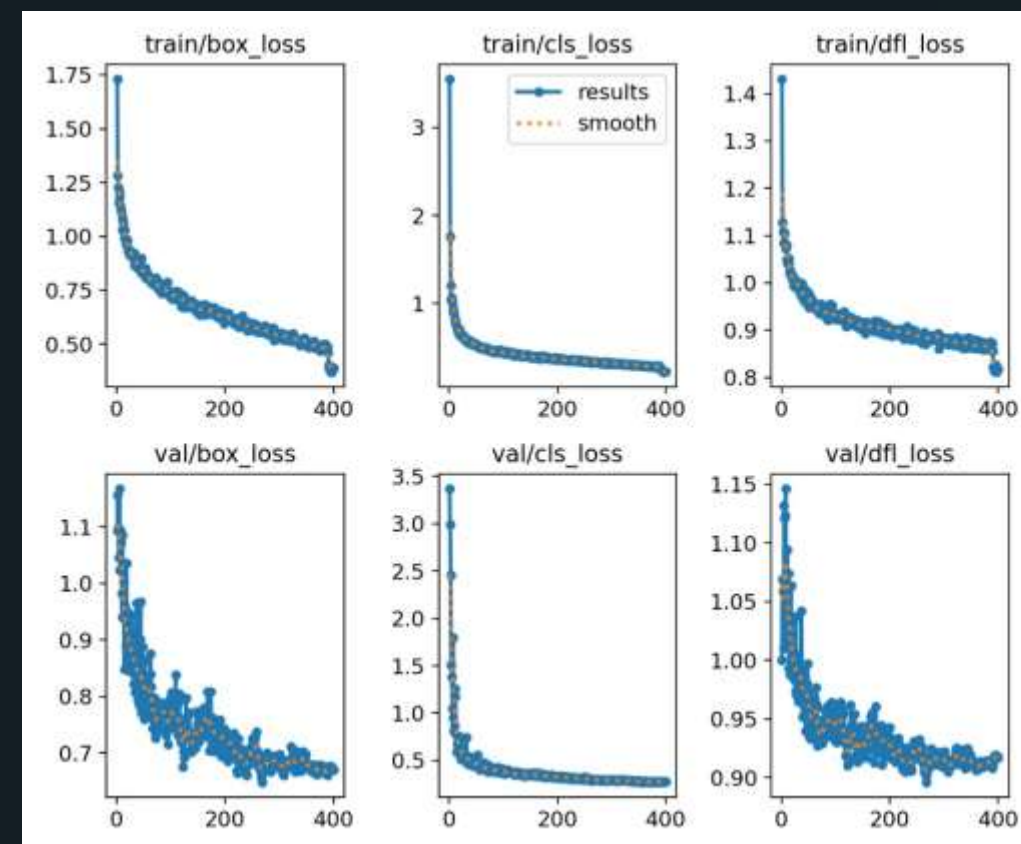


Обучение модели. Результаты и графики

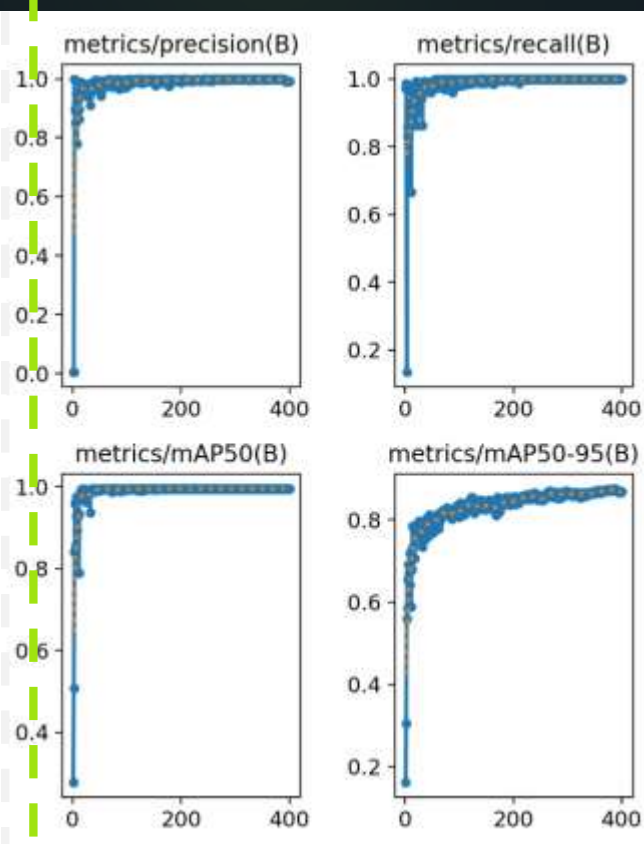
Размеченные изображения
на тестовой выборке



Графики
ошибок



Графики
ТОЧНОСТИ



Преобразование видео

Разбиение видео на отдельные кадры

- 1 Написал код на Python для разбиения видео на отдельные кадры-изображения
Для работы с видео использовал библиотеку cv2
В программе запущен бесконечный цикл для считывания видео и кадрирования
- 2 Для работы с видео использовал библиотеку cv2
- 3 В программе запущен бесконечный цикл для считывания видео и кадрирования





Сборка проекта

Модули системы

1

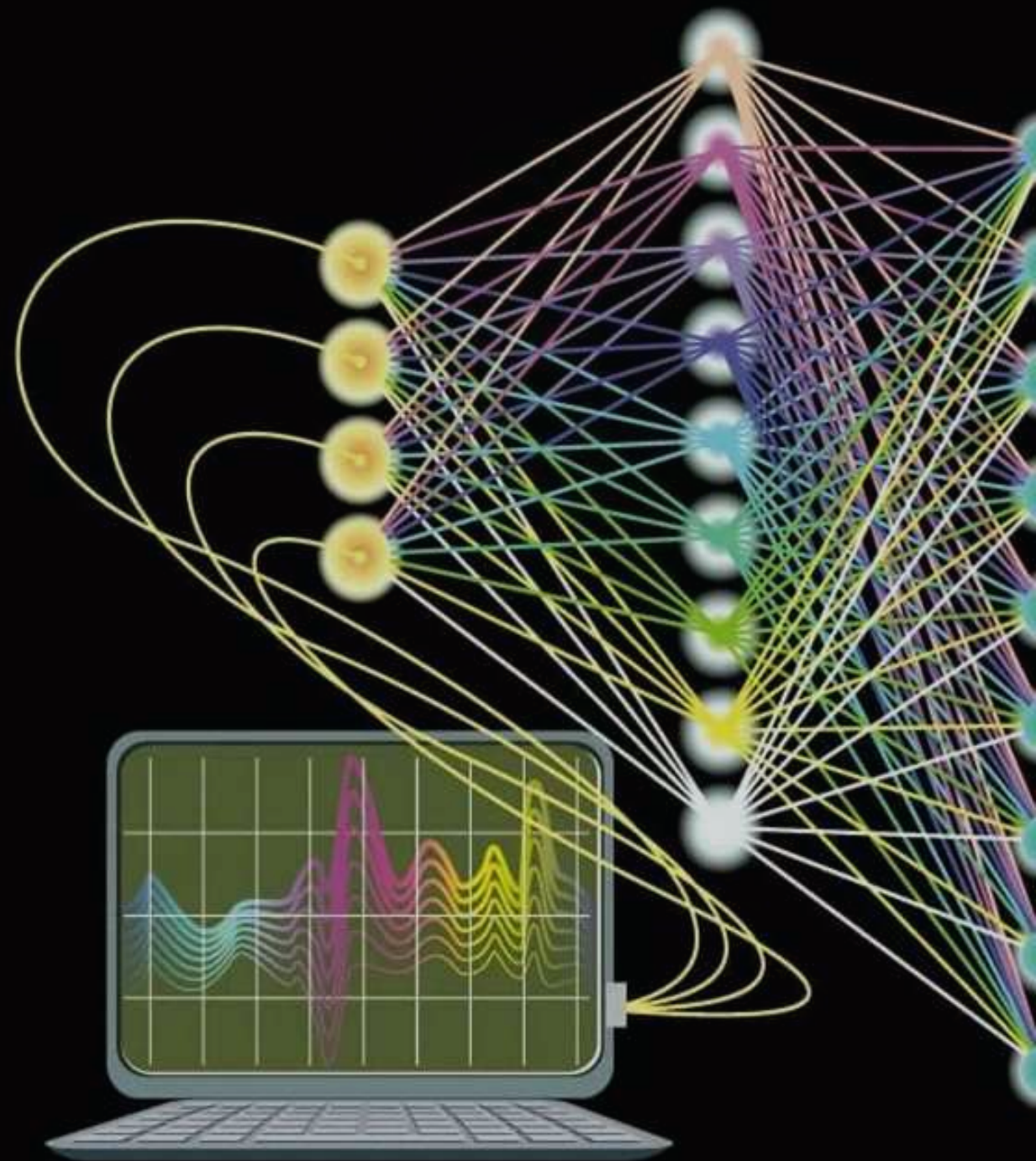
Получение изображений
с камеры

2

Отправка изображения в
обученную
нейросеть

3

Распознавание изображение
Определение класса объекта
Обработка координат
Перемещение курсора
Нажатие/отпускание клавиши мыши



Запуск и проверка

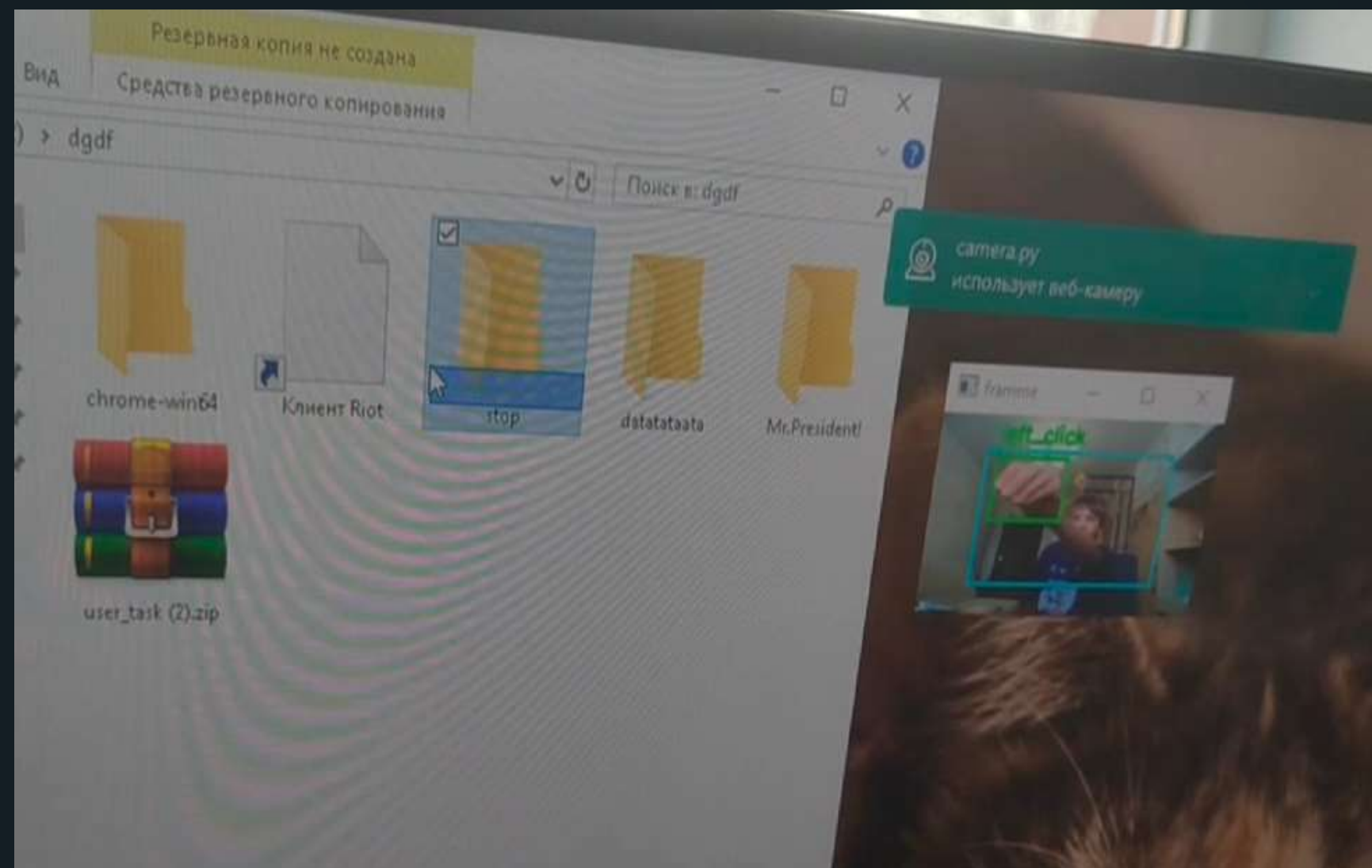
На видео видно, как программа распознает на кадрах видео руку, обводит ее в рамку и указывает найденный класс (hand)

Далее рассчитываются координаты центра прямоугольника, в который вписана рука и туда перемещается курсор мыши

Также видно, что задержки видео практически нет, курсор следует за рукой, даже при быстром движении руки

Стоп-кадр видео (полное видео работы размещено на диске

https://disk.yandex.ru/i/gT2zj9MRpGJMcQ_):





Планы по развитию проекта

1

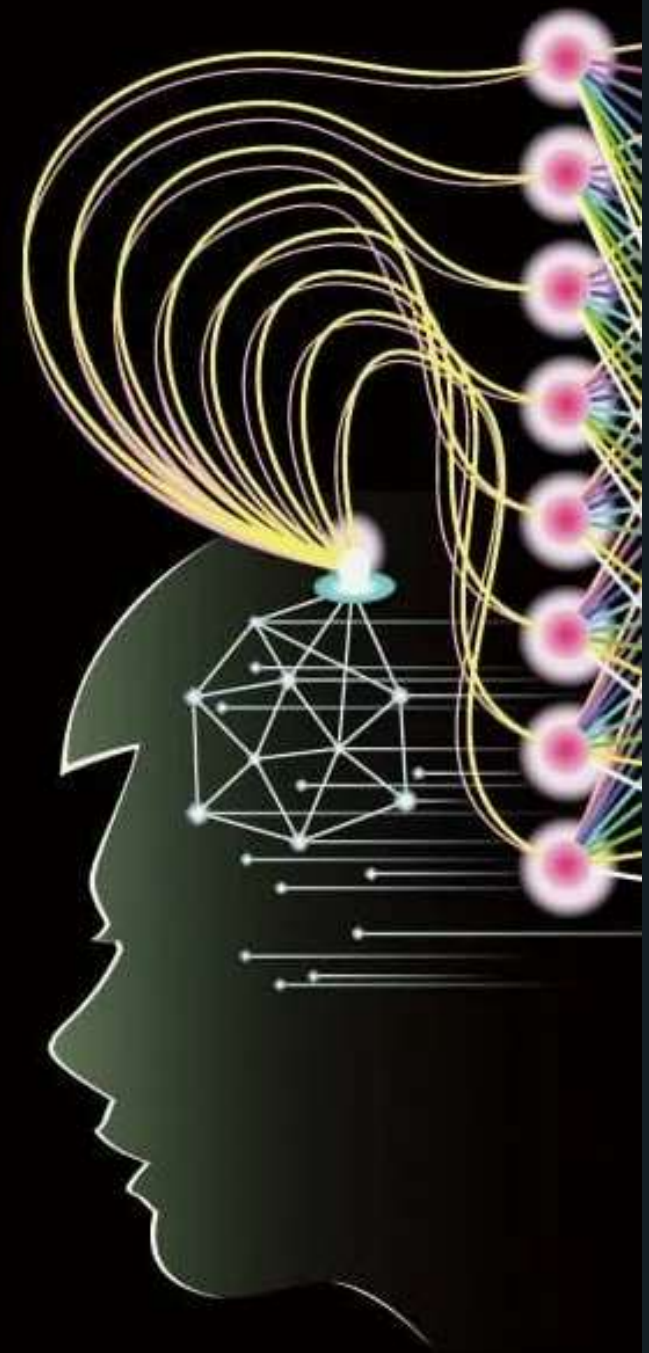
Оформить проект на GitHub и выложить свой код;

2

Создать свою модель и обучить ее по примеру YOLO;

3

Подумать над применимостью проекта в реальной жизни.



Выводы

1

Цели и задачи проекта достигнуты. Исследования в рамках проекта завершились положительно

2

Значительно увеличились знания по моделям, принципам их обучения и использования в Python

3

Есть четкое понимание развития проекта в будущем





Источники и материалы

Ultralytics YOLO Docs

https://docs.ultralytics.com/ru/yolov5/tutorials/train_custom_d22-create-labels

HABR

<https://habr.com/ru/articles/821971/>

GitHub

<https://github.com/boppreh/mouse#api>

YOUTUBE

<https://www.youtube.com/>

Алгоритмы и методы обучения нейросетей

<https://data-light.ru/blog/obucheniye-nejrosetej>

