



ДЕТЕКЦИЯ ЛЮДЕЙ, ПОТЕРЯВШИХСЯ В ЛЕСОПОСАДКАХ

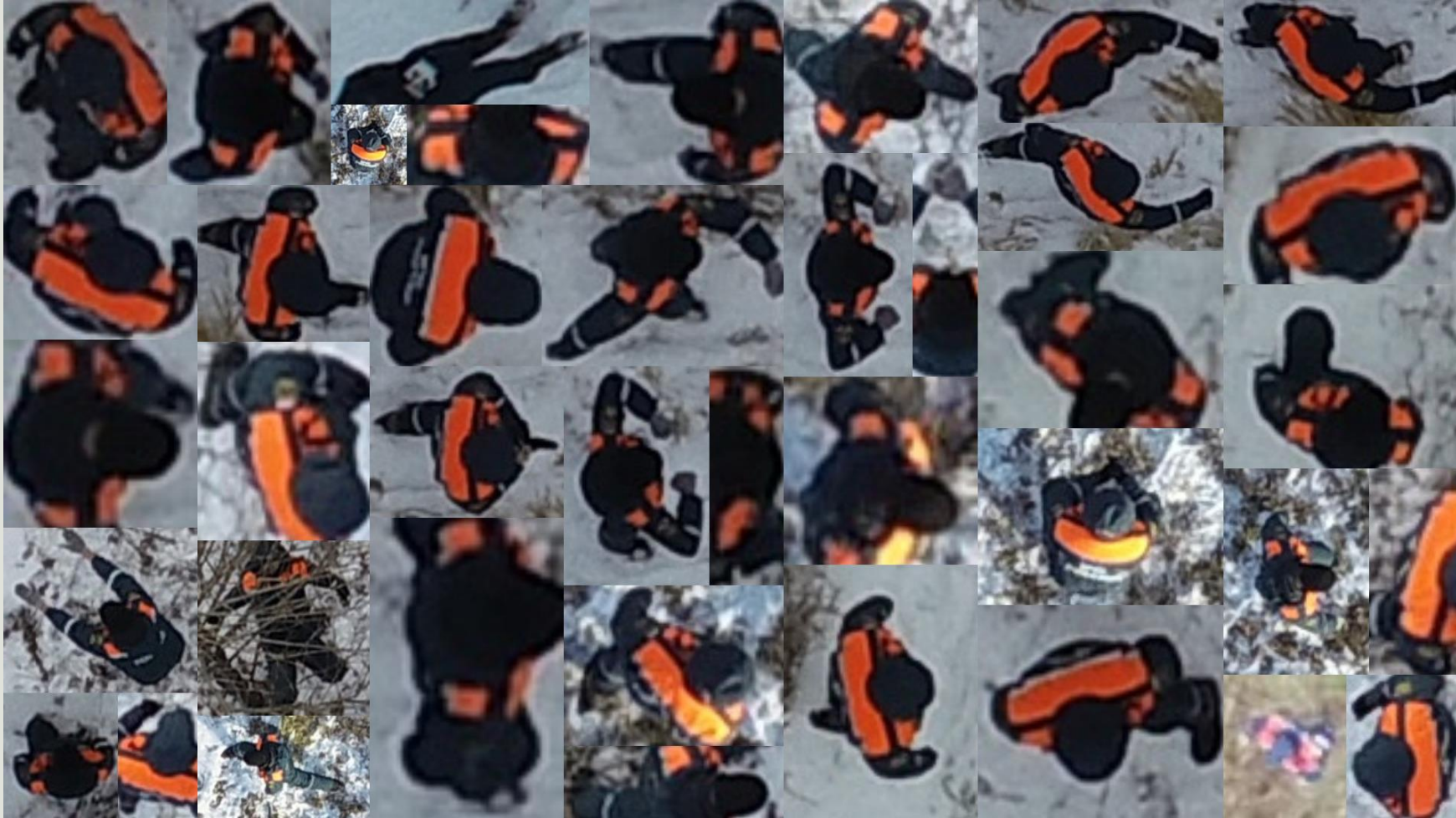
ПОСТАНОВКА ЗАДАЧИ

Дано: 1000 изображений людей, потерявшихся в лесопосадках. Каждое такое изображение сделано беспилотным летательным аппаратом и имеет разрешение 5472 на 3078 пикселей.

Требуется:

1. разработать алгоритм детекции людей на данных изображениях
2. разработать приложение с графическим интерфейсом для взаимодействия с алгоритмом детекции

ПОЧЕМУ МЫ ВЫБРАЛИ СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

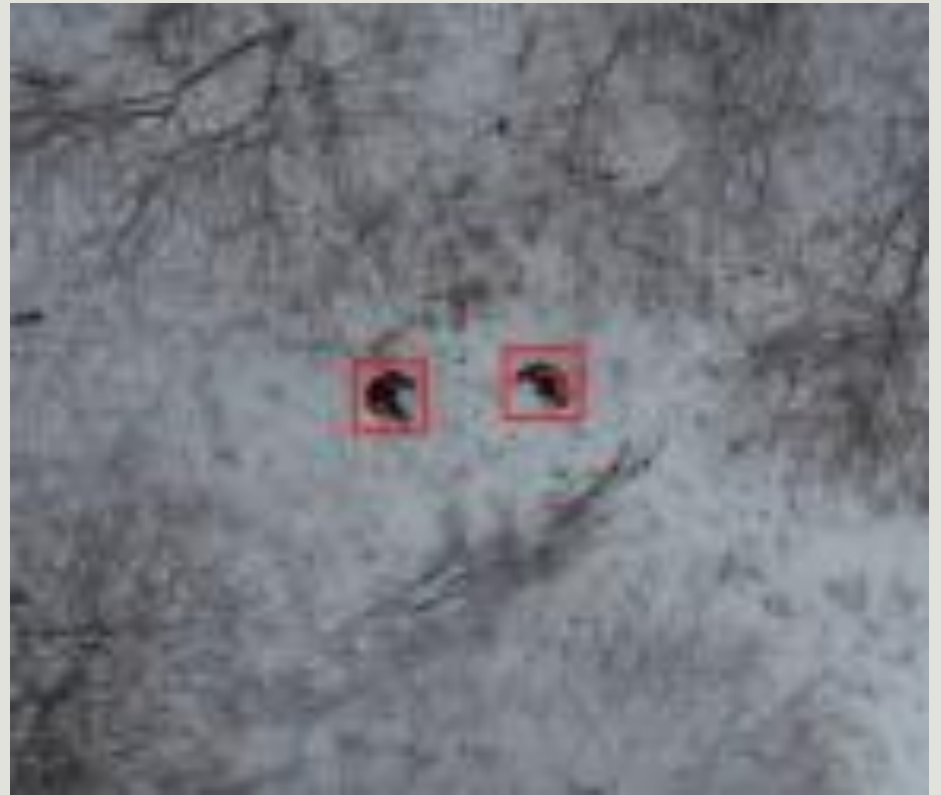


ПОДГОТОВКА ТРЕНИРОВОЧНОГО ДАТАСЕТА

Обучающая выборка составила 737 изображений (74%), валидационная – 160 (16%), тестовая – 103 (10%).



Правильно

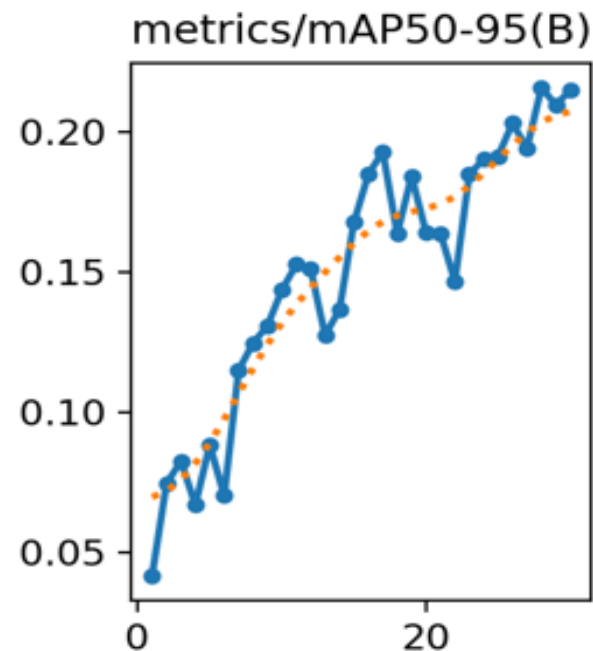
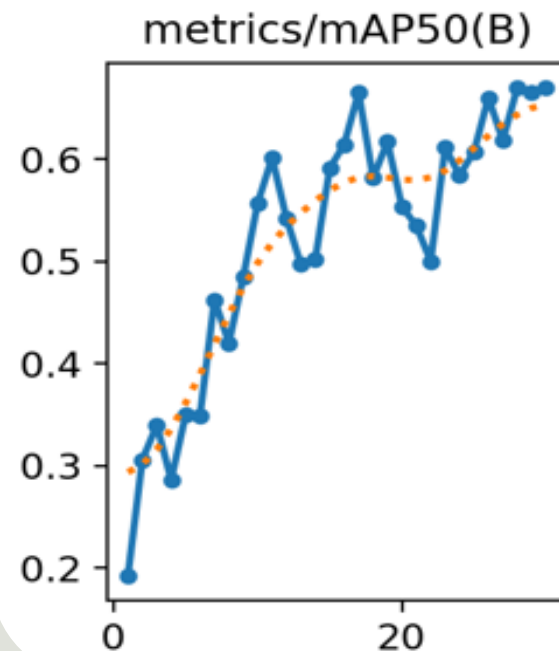
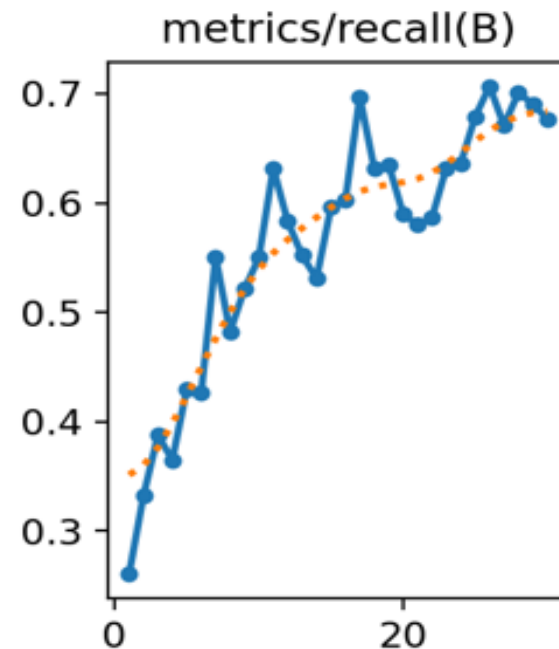
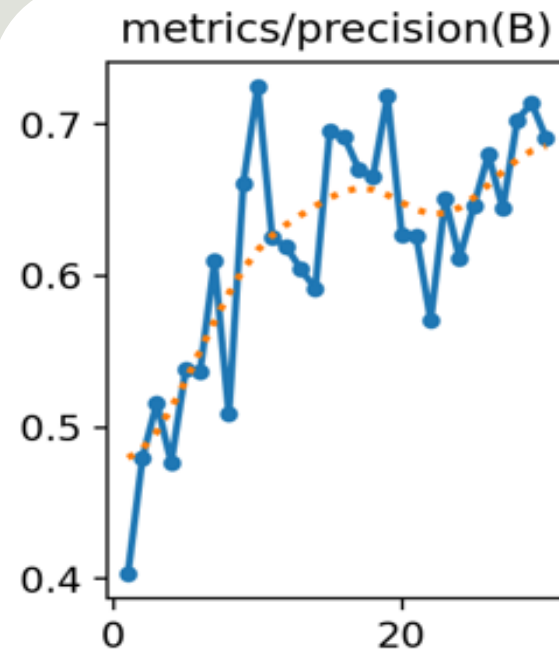


Неправильно

YOLOV5

Модель обучена на уменьшенных изображениях разрешения 480 на 480 пикселей, размер батча был равен 2, а количество эпох 30.

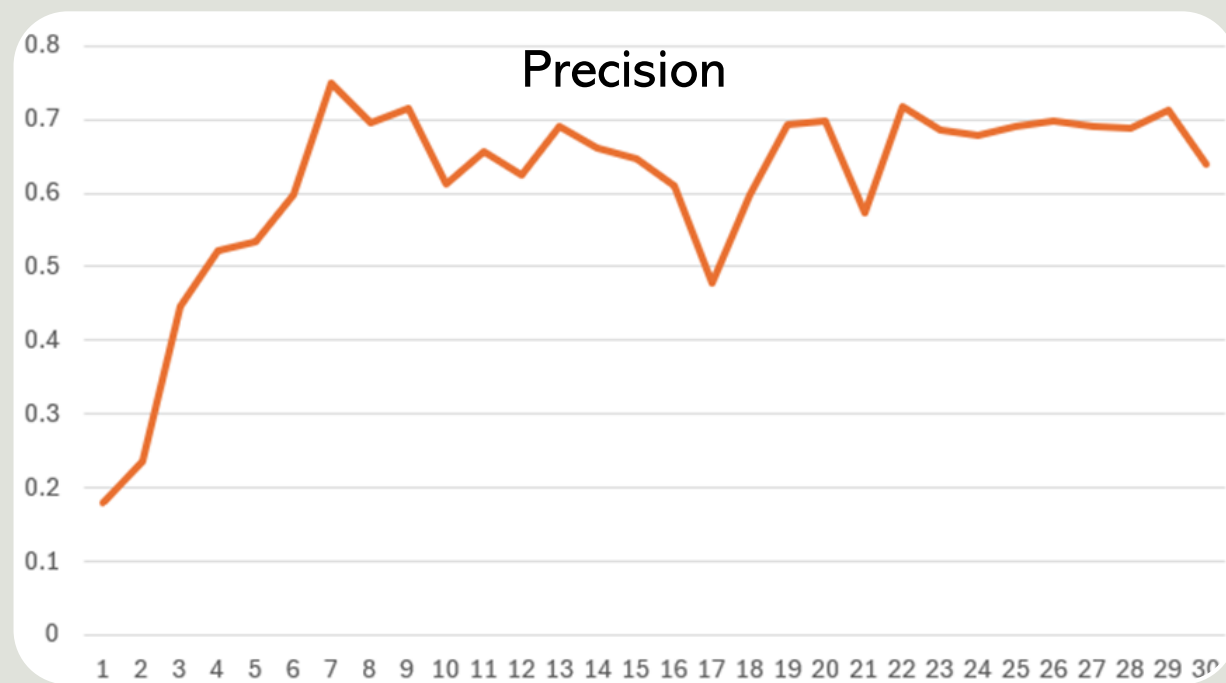
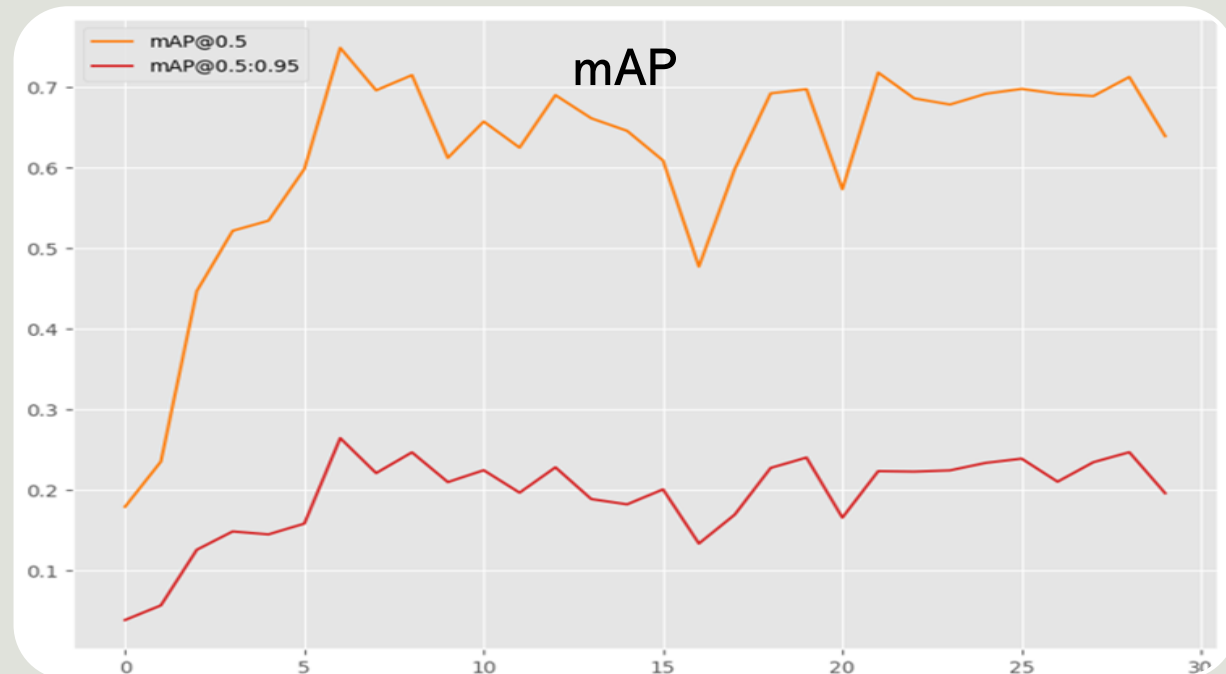
В результате 30 минут обучения precision модели стал 72%, а mAP50-95 – 22%



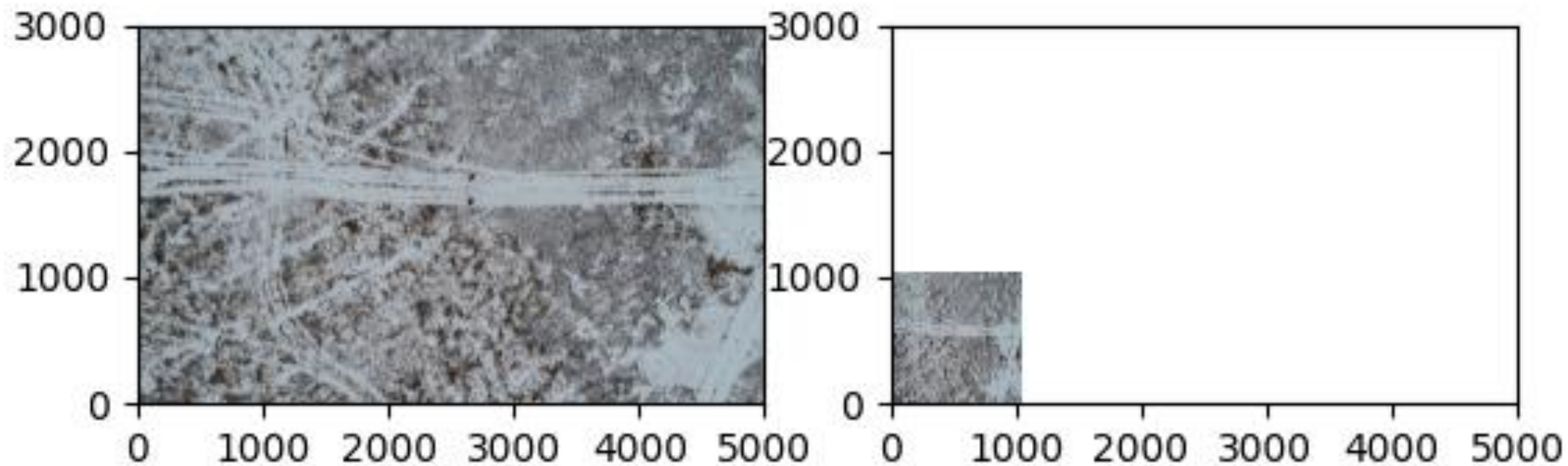
RETINANET

Модель обучена на уменьшенных изображениях разрешения 480 на 480 пикселей, размер батча был равен 2, а количество эпох 30.

В результате 10 часов обучения precision модели стал 64%, а mAP50-95 – 20%



ЛЮДИ НА АЭРОФОТОСНИМКАХ СЛИШКОМ МАЛЕНЬКИЕ



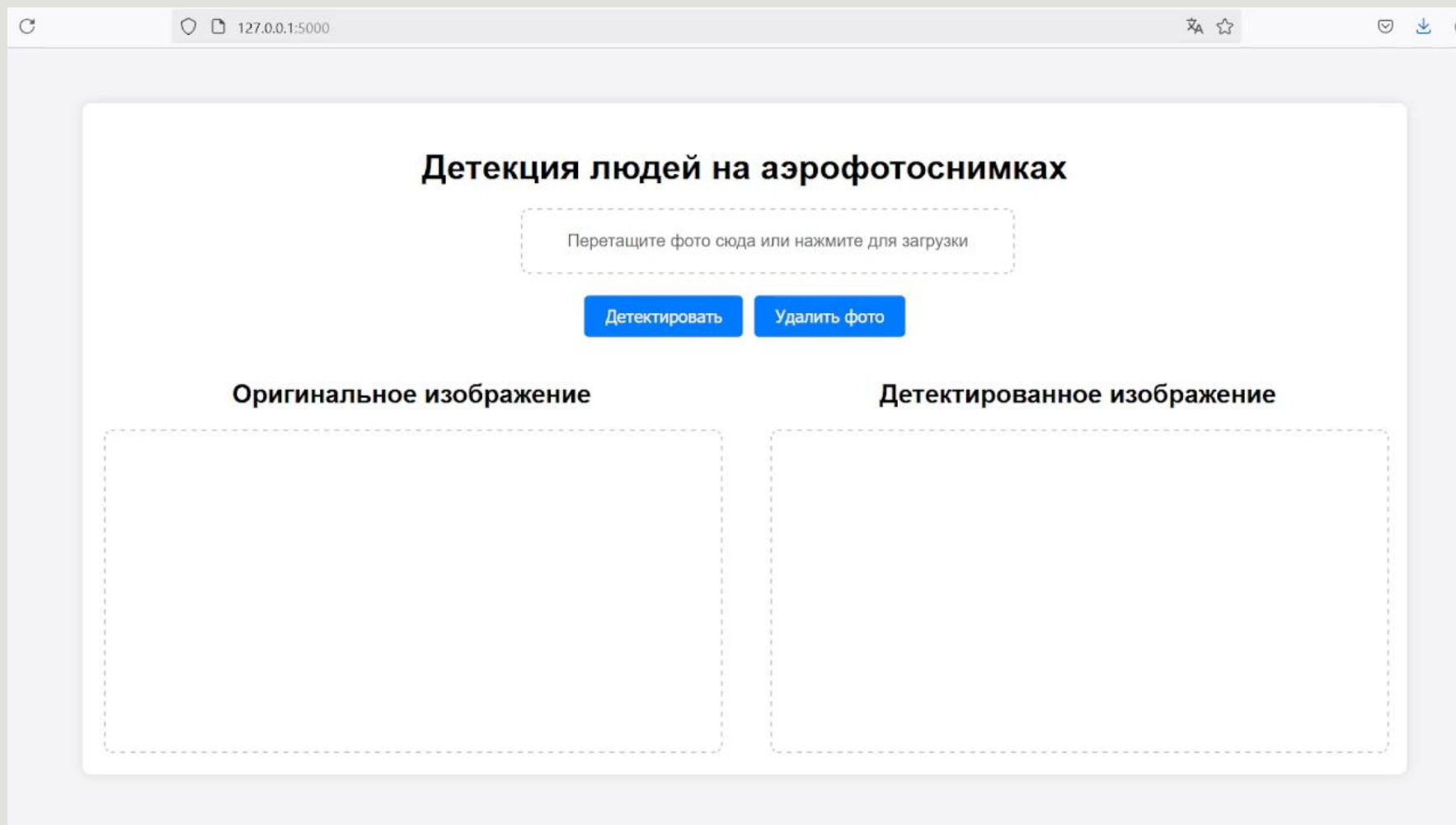
ОБЩАЯ СХЕМА АЛГОРИТМА

1. Разбить изображение на равные квадратные фрагменты
2. Для каждого фрагмента получить результаты детекции
3. Объединить фрагменты с полученными результатами детекции

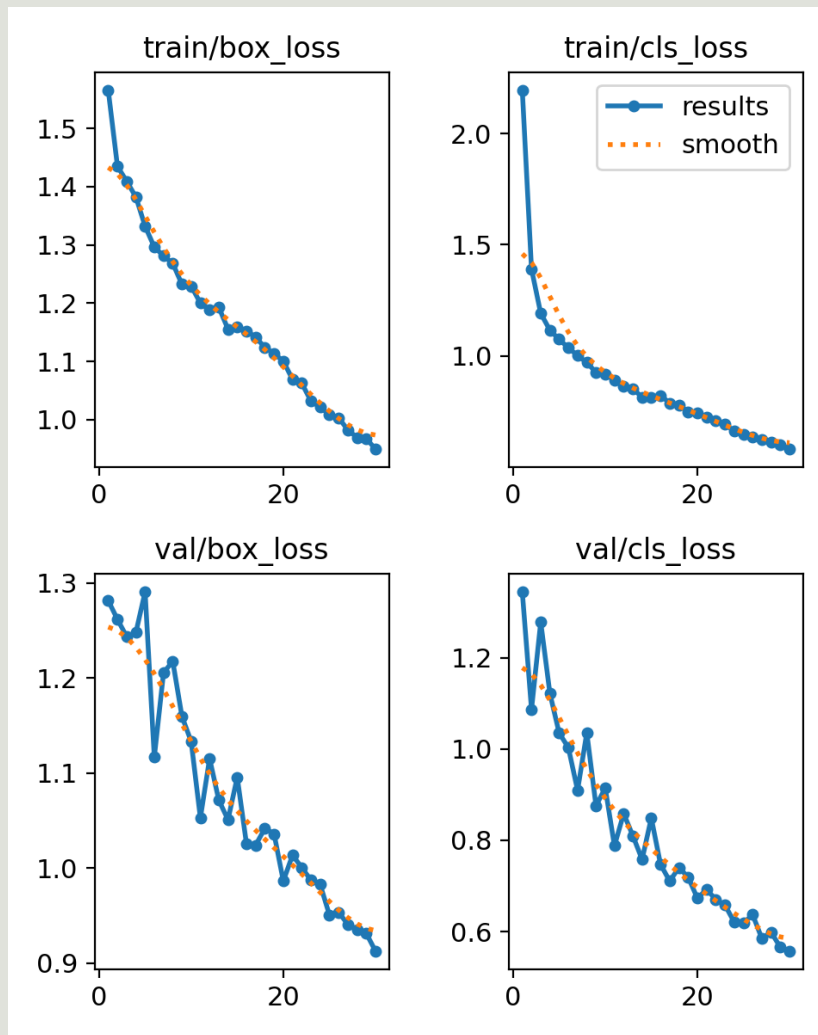
НАГЛЯДНАЯ РАБОТА АЛГОРИТМА



ДЕМОНСТРАЦИЯ РАБОТЫ



МЕТРИКИ ОБУЧЕННОЙ МОДЕЛИ YOLOV5



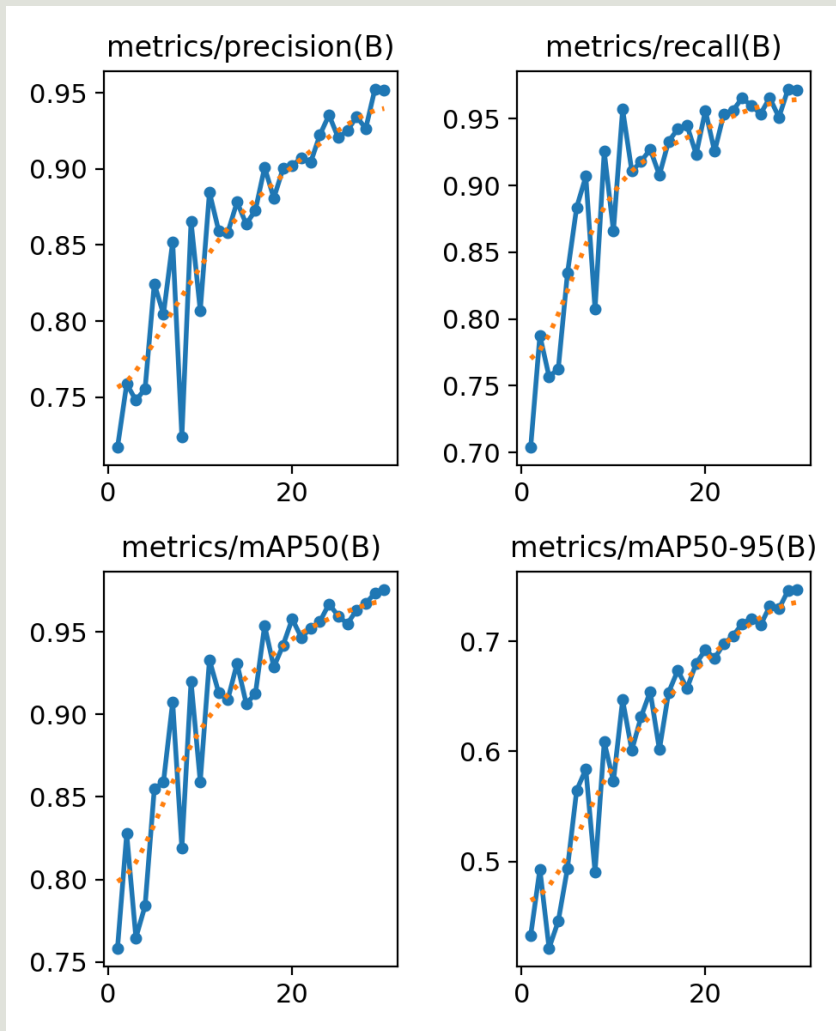
1.train/box_loss — ошибка предсказания координат ограничивающих прямоугольников на тренировочных данных. Значение на графике убывает, значит модель определяет местоположение объектов точнее.

2.train/cls_loss — ошибка классификации объектов на тренировочных данных. Значение на графике убывает, значит модель лучше различает классы объектов.

3.val/box_loss — ошибка предсказания координат ограничивающих прямоугольников на валидационных данных. Значение на графике убывает, значит модель хорошо обобщает на новых данных.

4.val/cls_loss — ошибка классификации объектов на валидационных данных. Значение на графике убывает, значит модель корректно классифицирует объекты на данных, не участвовавших в обучении.

МЕТРИКИ ОБУЧЕННОЙ МОДЕЛИ YOLOV5



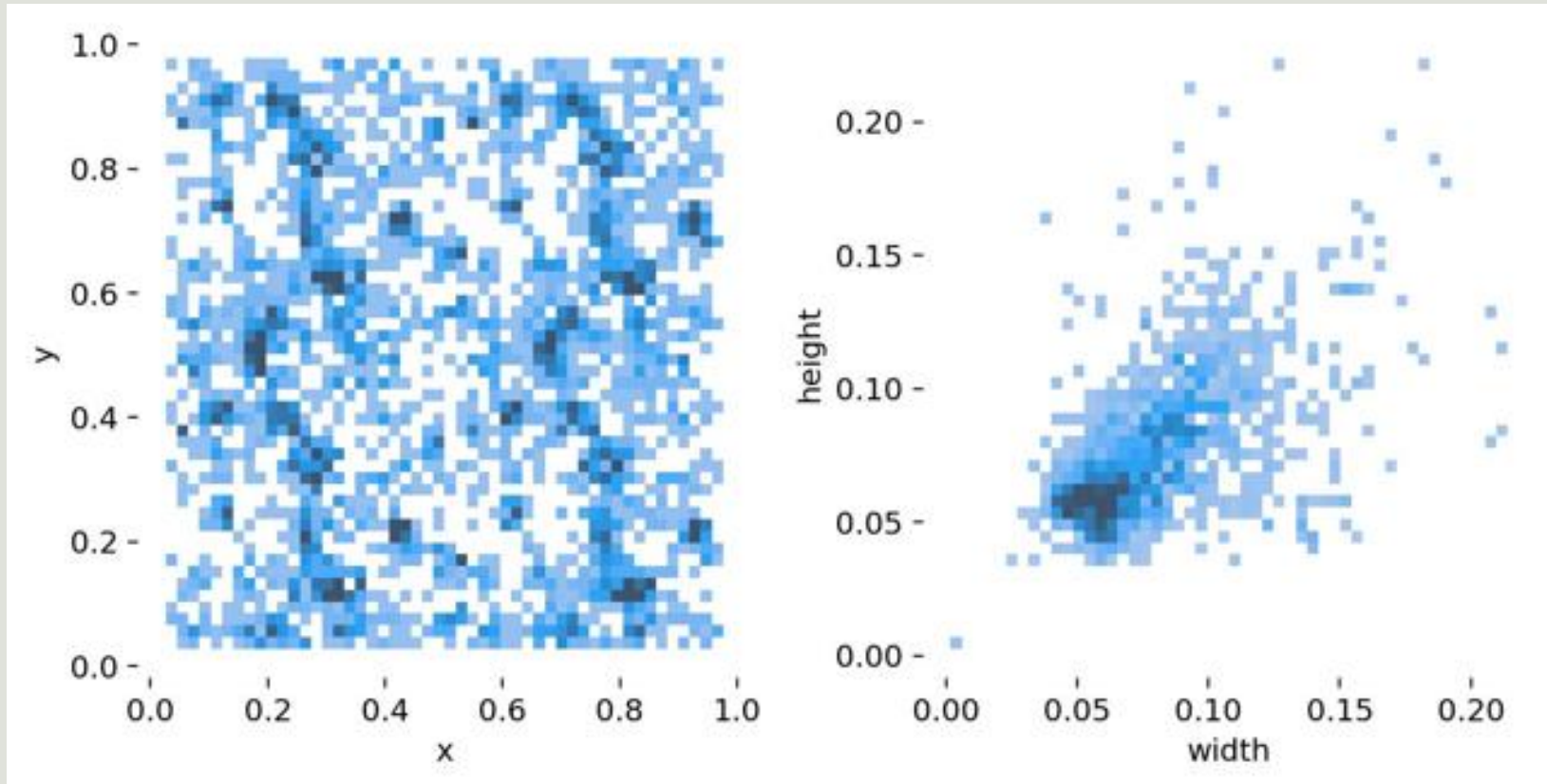
1.Precision — показывает, насколько модель "точна" в своих предсказаниях. Было получено значение **0.95**.

2.Recall — отражает способность модели находить все объекты. Было получено значение **0.97**.

3.mAP50 — средняя точность по всем классам, где объект считается правильно обнаруженным. Было получено значение **0.98**.

4.mAP50-95 — более строгая метрика по сравнению с mAP50, учитывающая качество локализации объектов. Было получено значение **0.75**.

МЕТРИКИ ОБУЧЕННОЙ МОДЕЛИ YOLOV5



На левом графике показана плотность расположений людей на тренировочной и валидационной выборке. На правом графике отражено соотношение ширины и длины человека на изображениях.

Как можно видеть, оно имеет линейную зависимость.

СРАВНЕНИЕ С ДРУГИМИ РЕШЕНИЯМИ

Название	Precision	mAP50-95
YOLO5 640px с разбиением на фрагменты	0.95	0.75
YOLO5 640px	0.86	0.36
YOLO5 480px	0.72	0.22
YOLO11 480px	0.70	0.22
RetinaNet 480px	0.64	0.20

ЧТО СДЕЛАЛИ МЫ, А ЧТО БИБЛИОТЕКИ?

Что сделала команда	Что сделали использованные инструменты
Подбор и разметка данных	Label Studio – преобразование размеченных изображений в нужный формат
Написание кода для кадрирования изображений	PIL – склеивание изображений
	OpenCV – чтение и сохранение фото
Подбор параметров для обучения нейронной сети	Yolo – обучение, возврат лучших весов обученной нейронной сети, отображение, получившихся метрик качества
	RetinaNet – обучение, возврат лучших весов обученной нейронной сети, отображение, получившихся метрик качества
	Pytorch – предоставление инструментов для обучения нейронных сетей
Написание кода для приложения с пользовательским интерфейсом	Flask – разворачивание сервера с работающим веб-приложением

ВЫВОДЫ

Таким образом был разработан алгоритм детекции людей в лесопосадках с показателем precision равным 95% и mAP 75%. При этом не наблюдается переобучения, что показывает детекция на отложенных тестовых данных. Разработано приложение с пользовательским интерфейсом для взаимодействия с алгоритмом. Представленный механизм детекции имеет перспективы улучшения (уменьшение времени обработки изображений и времени обучения). Кроме того, в разработанное приложение могут быть добавлены новые функции (загрузка новых размеченных данных, изменение параметров детекции и многое другое).