

# **Разработка средства визуализации и мониторинга криминального контента из туманной вычислительной среды**

Выполнила: Рослова Лариса Сергеевна, гр. 9304

Руководитель: Первицкий Александр Юрьевич, к.т.н., доцент

Консультант: Субботин Алексей Николаевич, ГУП "Петербургский  
метрополитен"

Санкт-Петербург, 2023

# Цель и задачи

## Актуальность:

- Количество контента в сети растёт с каждой минутой
- Нехватка человеческого ресурса и человеческий фактор
- Отсутствие доступных аналогов

**Цель:** разработать расширение для браузера, анализирующего посещаемый веб-ресурс на криминальный контент с применением средств машинного обучения и туманных вычислений.

## Задачи:

1. Обзор моделей машинного обучения
2. Обучить выбранную модель
3. Спроектировать архитектуру системы
4. Исследовать разработанную систему

# Задача 1. Обзор моделей классификации изображений

Модель	Размер (мб)	Топ-1 точность (%)	Топ-5 точность (%)	Пара- метры (млн)	Время вывода на CPU (мс)	Время вывода на GPU (мс)
<b>VGG16</b>	528	71.3	90.1	138.4	69.5	4.2
<b>InceptionV3</b>	92	77.9	93.7	23.9	42.2	6.9
<b>ResNet50</b>	98	74.9	92.1	25.6	58.2	4.6
<b>MobileNetV2</b>	14	71.3	90.1	3.5	25.9	3.8
<b>EfficientNet</b>	220	85.3	97.4	54.4	1578.9	61.6

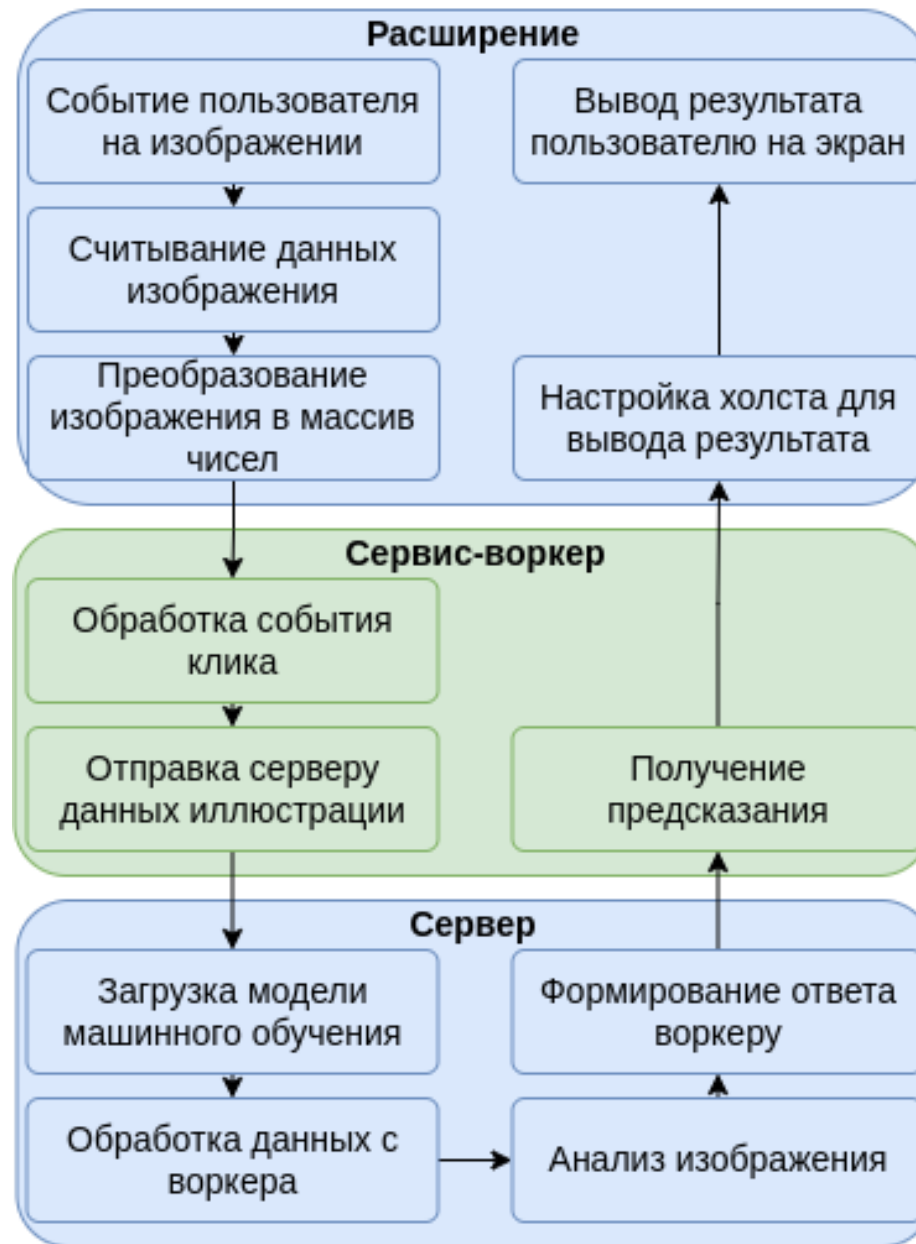
## Задача 2. Обучение модели EfficientNet

### Подзадачи:

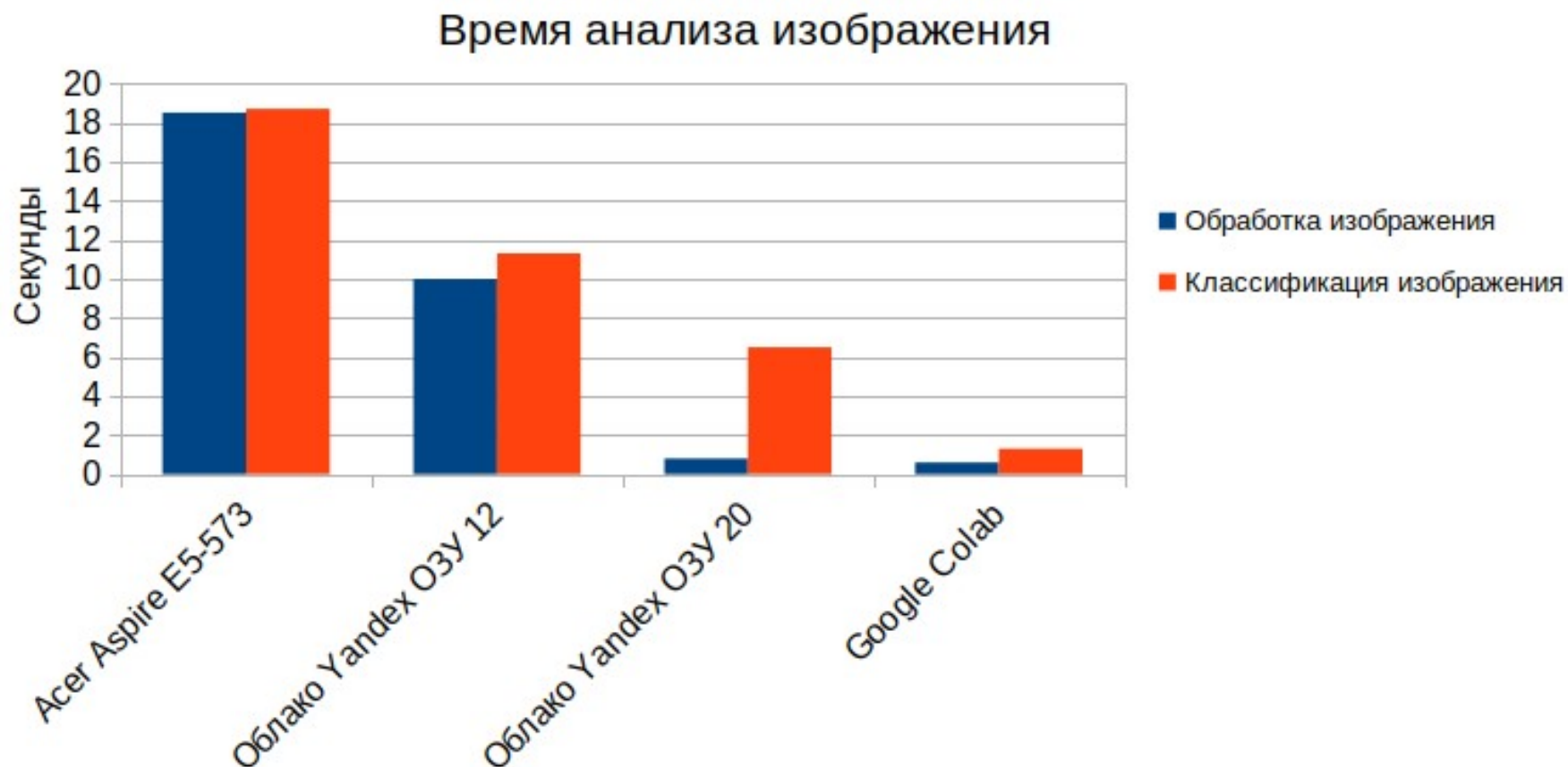
- Собрать набор данных для обучения
- Применить трансферное обучение к модели EfficientNet
- Конвертировать переобученную модель из формата HDF5 (.h5) в формат Layers (.JSON).

**Результат:** после обучения на собранном датасете нейронная сеть показала точность в **92%** на тестовых данных.

# Задача 3. Архитектура системы



## Задача 4. Интеграция с облачными технологиями



С самой дорогой конфигурацией вычислительных мощностей можно уменьшить задержки в 10 раз по сравнению с персональным компьютером.

# Заключение

- Были изучены технологии машинного обучения для анализа изображений.
- На основе сравнения была выбрана и обучена нейросеть EfficientNet, которая показала 92% точности определения криминального контента на изображениях.
- Спроектирована архитектура расширения для браузера, анализирующего изображения на криминальное содержание.
- Тестирование и оптимизация работы продукта показали, что при использовании мощных конфигураций облачных сервисов можно добиться ускорения анализа контента в 10 раз.

# Способы развития продукта

- Будет полезным внедрить возможности анализа текстового контента и мониторинг нескольких изображений за единицу времени.
- Улучшение точности и производительности.
- Расширение поддерживаемых платформ и браузеров.
- Улучшение пользовательского интерфейса.
- Регулярное обновление моделей машинного обучения и базы данных негативного контента поможет улучшить работу расширения и его способность обнаруживать новые типы негативного содержания.



# Апробация работы

Исходный код разработанного продукта доступен на GitHub:  
[https://github.com/LRoslova/criminalContent\\_monitor](https://github.com/LRoslova/criminalContent_monitor)



В описании репозитория есть подробная инструкция по установке и работе программы.

Запасные слайды

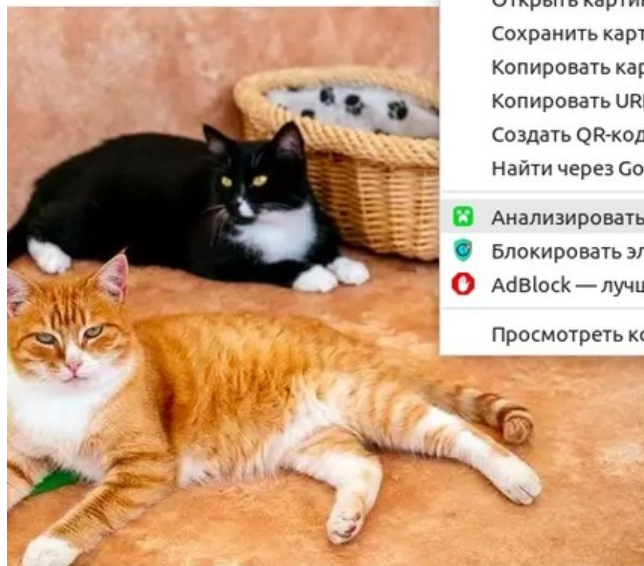
# Параметры тестирования

Используемые вычислительные мощности:

- Acer Aspire E5-573, ОЗУ 8 Гб, CPU 4;
- Облако на платформе Yandex, ОЗУ 20 Гб, CPU 20;
- Облако на платформе Yandex, ОЗУ 12 Гб, CPU 14;
- Платформа Google Colab ОЗУ 12.7, GPU T4;

# Интерфейс расширения

к «Новый дом»



- Открыть картинку в новой вкладке
- Сохранить картинку как...
- Копировать картинку
- Копировать URL картинки
- Создать QR-код для этого изображения
- Найти через Google
- Анализировать изображение
- Блокировать элемент...
- AdBlock — лучший блокировщик рекламы ▶
- Просмотреть код



Ученые: у владельцев собак  
более крепкое сердце - BBC...

[Перейти](#)