

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
«Южно-Уральский государственный университет  
(национальный исследовательский университет)»  
Высшая школа электроники и компьютерных наук  
Кафедра системного программирования

# **Разработка приложения на основе нейронных сетей для оценки качества фотографий на документы**

Научный руководитель:  
ст. преподаватель кафедры СП  
Н. С. Силкина

Автор работы:  
студент группы КЭ-401  
М. А. Караев

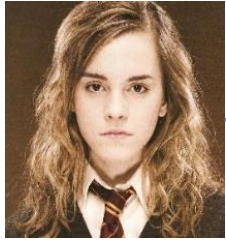
Челябинск, 2024 г.

# АКТУАЛЬНОСТЬ

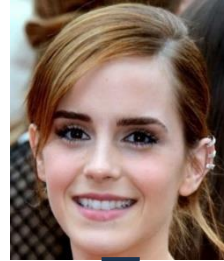
- Широкое распространение систем распознавания лиц
- Большие затраты вычислительных ресурсов на обработку системой распознавания
- Наличие легкой системы оценки качества, которая предварительно фильтрует непригодные фотографии, сохранит процессорное время

# БЕЗ ОЦЕНКИ КАЧЕСТВА

Приходящие фото



Базовое фото



Распознаватель  
(1000 мс/фото\*)

Результат  
распознавания



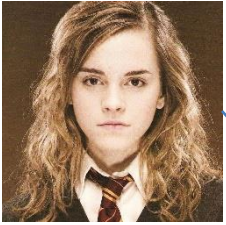
Итого: 4000 мс

\* Тестирование проводилось на ResNet400

<https://github.com/xkvma/doqual>

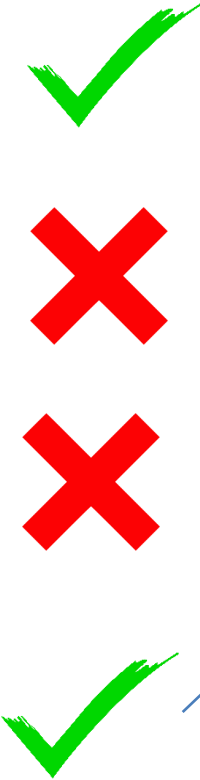
# С ОЦЕНКОЙ КАЧЕСТВА

Приходящие фото



Оценка качества

Оценщик  
Качества  
(20 мс/кадр)



Базовое фото



Распознаватель  
(1000 мс/ кадр\*)

Результат  
распознавания



Итого: 2080 мс

\* Тестирование проводилось на ResNet400

<https://github.com/xkvma/doqual>

# ЦЕЛЬ И ЗАДАЧИ

## Цель:

Разработка настольного приложения для оценки качества фотографии на документы на основе нейронных сетей

## Задачи:

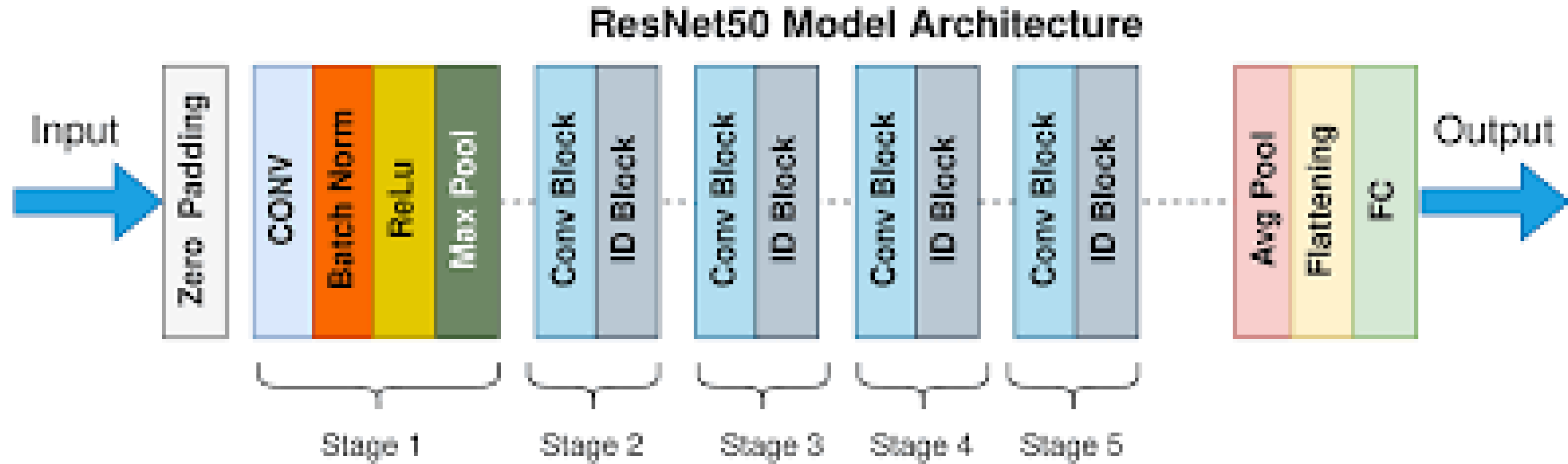
1. Изучить подходы к построению нейронной сети для оценки скаляра качества
2. Построить и обучить нейронную сеть
3. Реализовать и протестировать приложение на языке Python для оценки качества фотографии на документы

# МЕТРИКА EFFICIENCY

$$\eta(r) = \frac{1}{r} \left( \frac{FNMR(0) - FNMR(r)}{FNMR(0)} \right)$$

$FNMR(r)$  – процент ложно распознанных людей при выкидывании  $r$  изображений с худшим качеством из датасета

# ИСХОДНЫЙ РАСПОЗНАВАТЕЛЬ



# НАБОРЫ ДАННЫХ

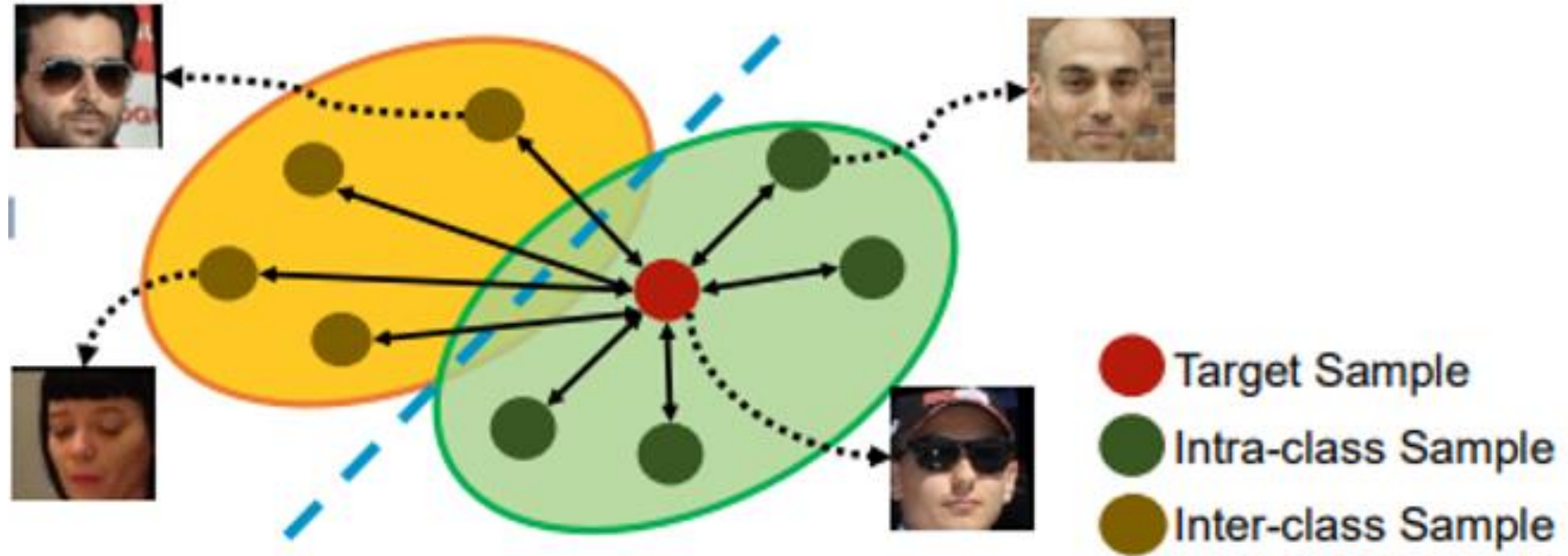
Обучение: Glint360k  
Тестирование: LFW





# ГЕНЕРАЦИЯ КАЧЕСТВА

Способ генерации качества из метода SDD-FIQA\*



\* Fu-Zhao Ou, Xingyu Chen, Ruixin Zhang, Yuge Huang, Shaoxin Li, Jilin Li, Yong Li, Liujuan Cao, Yuan-Gen Wang  
SDD-FIQA: Unsupervised Face Image Quality Assessment with Similarity Distribution Distance  
<https://doi.org/10.48550/arXiv.2103.05977>

# НОРМАЛИЗАЦИЯ БАТЧА

$$\text{Normalizing } x'_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

$$\text{Scaling and shifting } y_i = \gamma x'_i + \beta$$

$\mu_B$  – мат ожидание признаков во время обучения

$\sigma_B^2$  – дисперсия во время обучения

$x'_i$  – признаки после нормализации

# ДИВЕРГЕНЦИЯ КУЛЬБАКА ЛЕЙБЛЕРА

$$D_{sKL}(F||BN) = \frac{1}{C} \sum_{c=1}^C [D_{KL}(F_c||BN_c) + D_{KL}(BN_c||F_c)]$$

$$KL(P||Q) = - \sum P(X) \log \frac{Q(x)}{P(x)}$$

$D_sKL$  – симметричная Дивергенция Кульбака-Лейблера

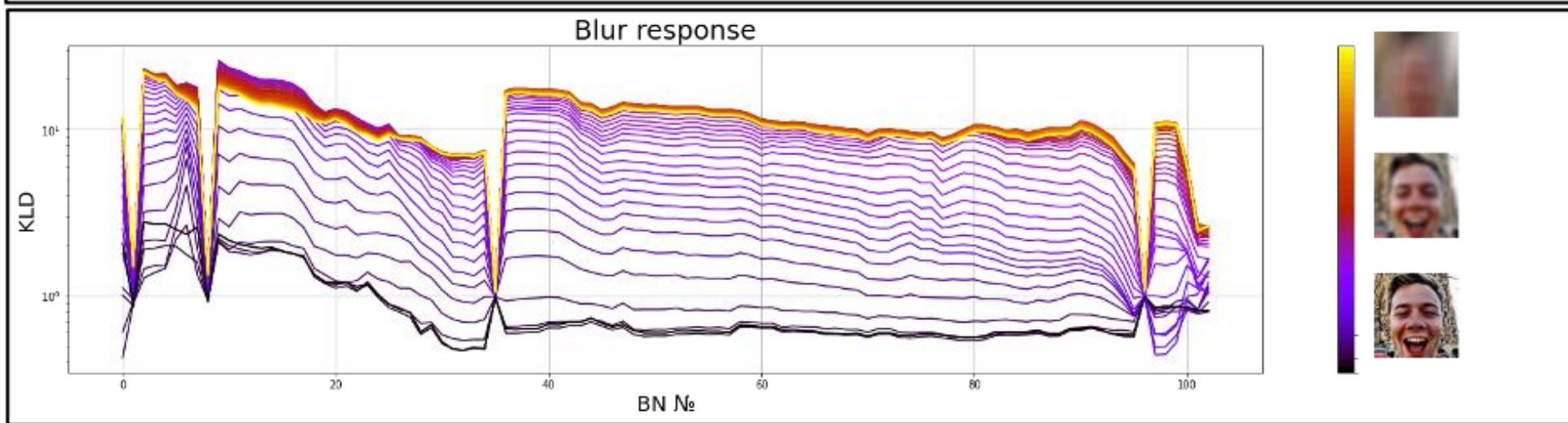
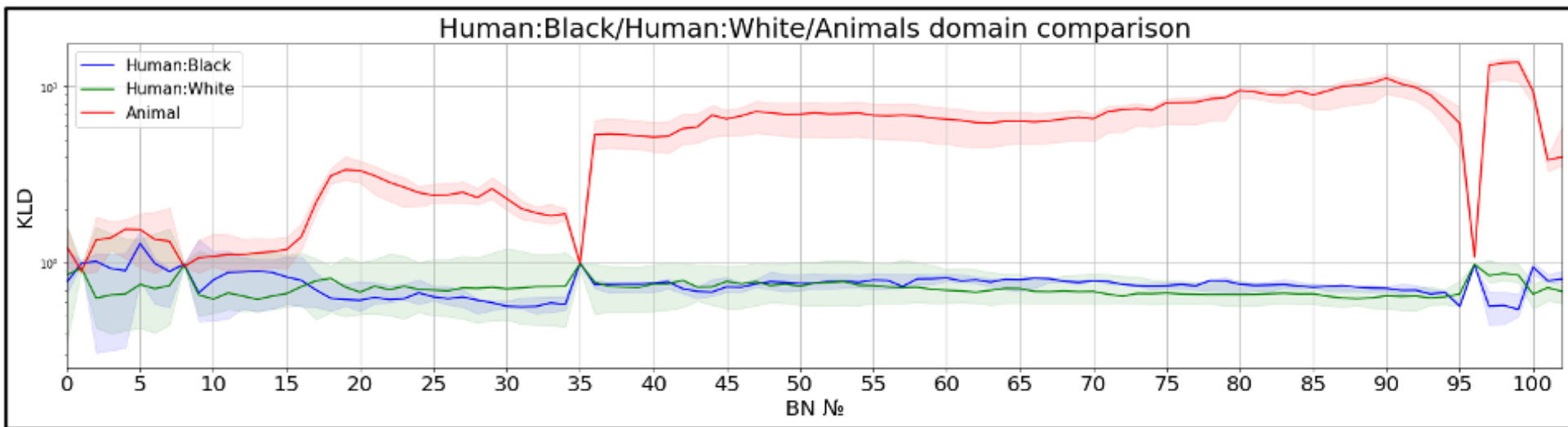
$D_{KL}$  – дивергенция Кульбака-Лейблера

$F_c$  – признаки одного канала обрабатываемого изображения

$C$  – число каналов слоя BatchNorm

$BN_c$  – признаки одного канала заученные слоем BatchNorm

# СРАВНЕНИЕ ПРИЗНАКОВ



# ИТОГОВЫЙ МЕТОД

- 1) Генерируем истинное качество как схожесть эмбединга с центроидом эмбедингов этого класса
- 2) Подаем фотографию в нейронную сеть, обученную на распознавание лица (легкую)
- 3) Извлекаем дивергенцию KL между распределениями, заученными слоями батч нормализации и признаками, извлеченными из обработанного фото
- 4) Обучаем регрессор качества на основе извлеченных признаков

# ВЫБОР ТЕХНОЛОГИЙ



# КОНФИГУРАЦИЯ ОБУЧЕНИЯ

– Nvidia GeForce RTX 3070

– CatBoostRegressor:

- depth = 10
- l2\_leaf\_reg = 2
- iterations = 2500
- learning\_rate = 0.03

# ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

| Описание метода  | Efficiency 1%  | Efficiency 5% | Efficiency 10% | Время выполнения, мс |
|--|----------------|---------------|----------------|----------------------|
| ResNet18 среднее + дисперсия <u>catboost</u>                                   | 10,7491        | 6,8820        | 4,3800         | 20                   |
| ResNet18 KLD <u>пока-</u><br><u>нальное catboost</u>                           | 10,7670        | 8,3167        | <b>6,5721</b>  | 60                   |
| ResNet100 норма эмбединга  | 17,1890        | 9,0890        | 4,4490         | 50                   |
| ResNet200 среднее + дисперсия (10 последних слоев) <u>cat-</u><br><u>boost</u> | <b>17,7337</b> | <b>9,2522</b> | 5,9580         | 160                  |



# ПРИМЕР ХУДШЕГО КАЧЕСТВА

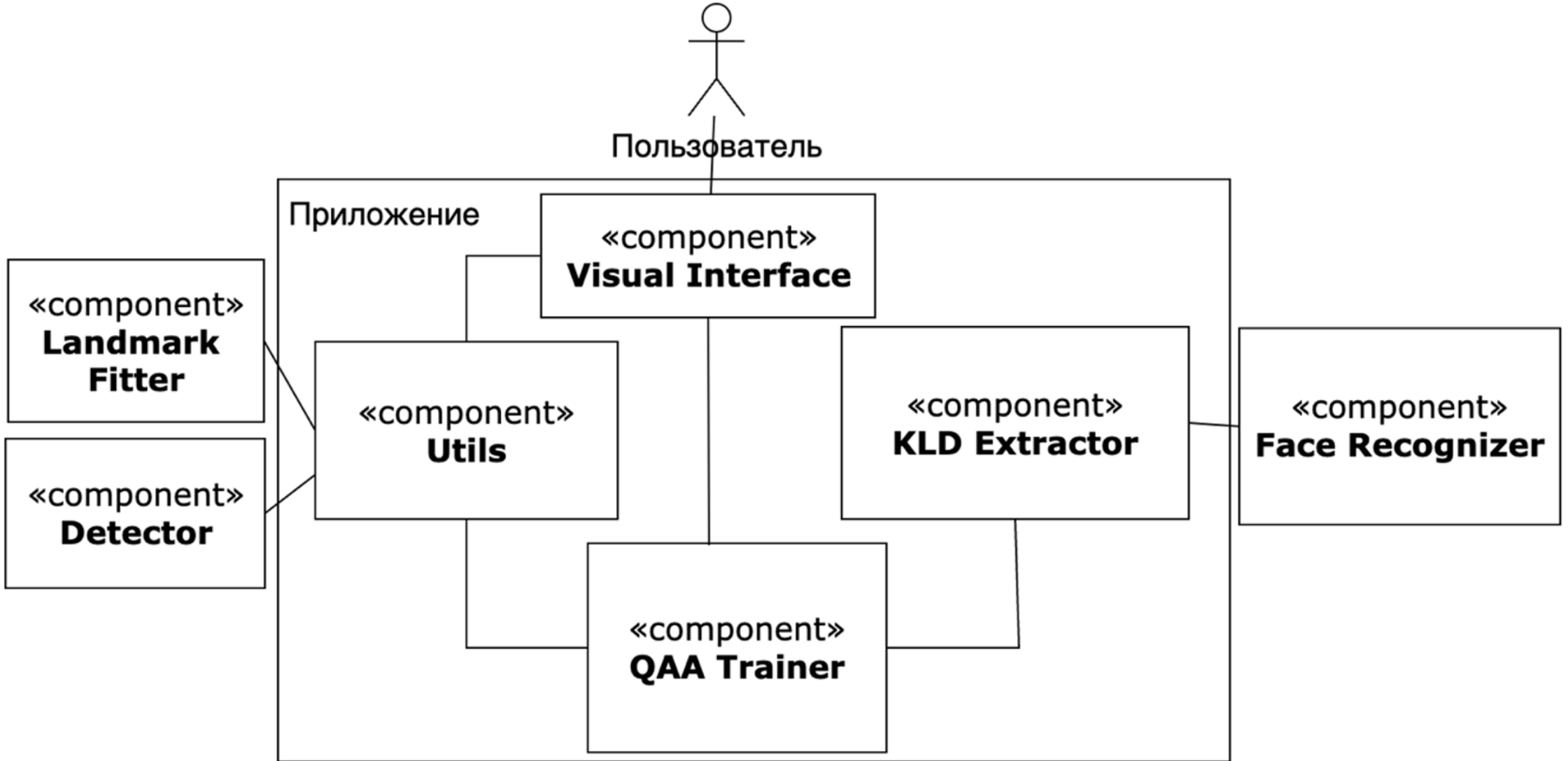


\* Q – оценка качества разработанным методом

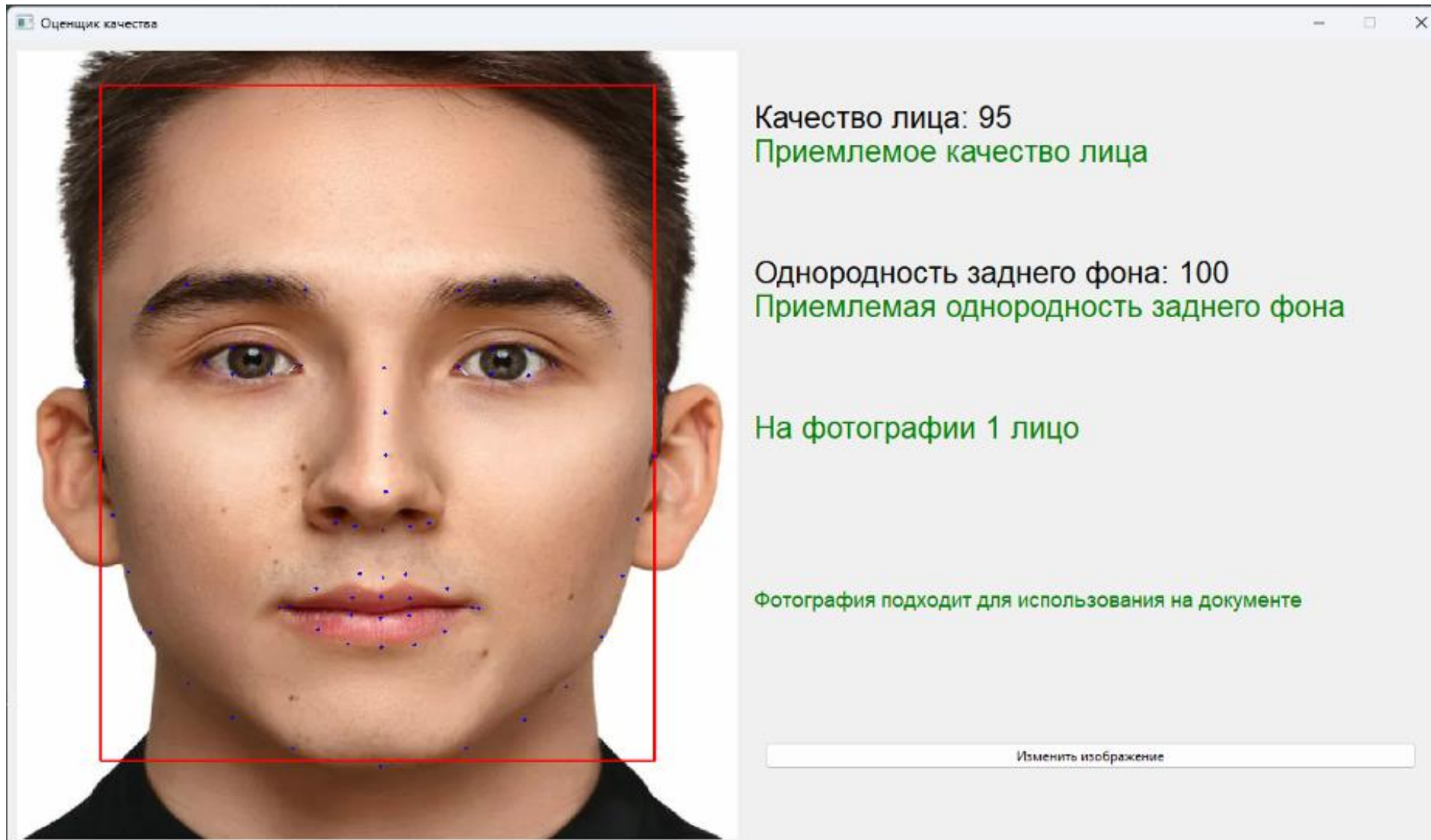
# ФУНКЦИОНАЛЬНЫЕ ТРЕБОВАНИЯ

- 1) Система должна оценивать скаляр качества фотографии лица
- 2) Система должна оценивать однородность заднего фона на фотографии лица
- 3) Система должна иметь интерфейс для взаимодействия с пользователем

# КОМПОНЕНТЫ СИСТЕМЫ



# ПРИМЕР РАБОТЫ



# ТЕСТИРОВАНИЕ

| № | Название теста                  | Шаги  | Ожидаемый результат   |
|---|---------------------------------|---|---|
| 1 | Загрузить фотографию            | В стартовом окне нажать кнопку “загрузить фотографию” и указать путь к фотографии.                                      | На экран выведется фотография и оцененные параметры   |
| 2 | Изменить загруженную фотографию | В окне с оцененными параметрами фото нажать кнопку «изменить фотографию».<br>Выбрать новую фотографию с лицом человека. | Фотография на экране заменится новой, будут переподсчитаны оценка качества лица и однородность фона |
| 3 | Загрузить фотографию без лиц    | Загрузить на обработку фотографию без лиц   | На экран выведется сообщение об отсутствии лиц на фотографии  |

# РЕЗУЛЬТАТЫ

1. Изучены подходы к построению нейронной сети для оценки качества лица.
2. Построена и обучена нейросеть для оценки качества фотографии лица на документы.
3. На основе обученной нейронной сети было реализовано и протестировано приложение на языке Python для оценки качества фотографий на документы.