



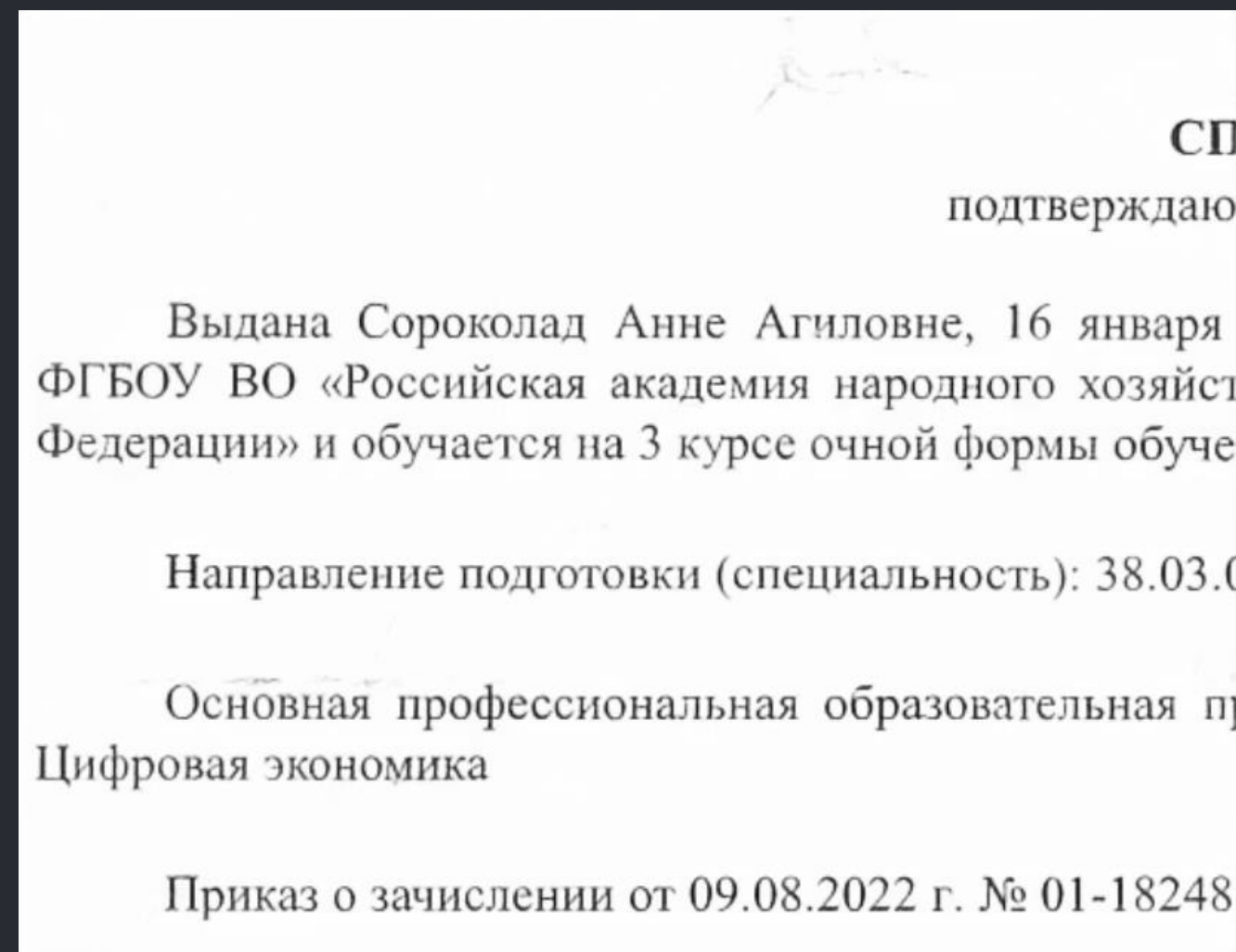
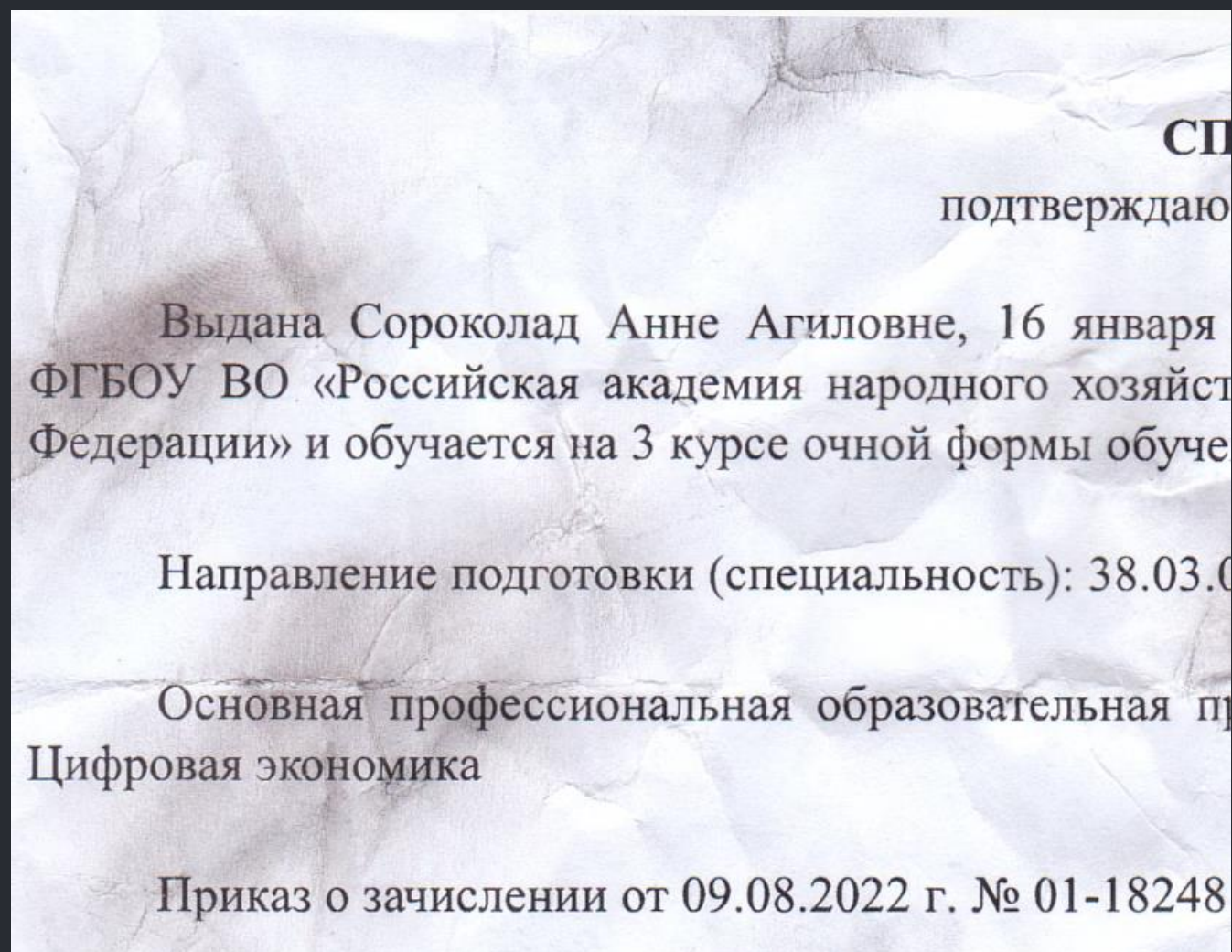
# Проект по ГО: Очистка фотографий

Презентация посвящена разработке MVP (промышленное решение) для очистки фотографий документов.



Выполнили Василенко Анна и Сороколад Анна

# MVP Очистка изображений



Показаны реальные сканы до и после очистки. Продукт уже функционирует.

# Очистка изображений

Нажмите для выбора файла или перетащите изображение сюда

Обработать изображение



# Данные и задачи очистки

There are several classic spatial filters for reducing or eliminating noise from images. The mean filter, the median filter and the closing operation are commonly used. The mean filter is a lowpass or smoothing filter that replaces each pixel with its neighborhood mean. It reduces the image noise but blurs the image. The median filter calculates the median of the pixel neighborhood for each pixel, thus removing salt and pepper noise. Finally, the opening closing filter is a mathematical morphology operation that removes the same number of erosion and dilation morphological operations to remove objects from images.

The main goal was to train a neural network in a supervised manner to denoise an image from a noisy one. In this particular case, it was much easier to train a neural network to clean an image from a clean one than to clean a subset of noisy images.

There exist several methods to design forms with fields. For instance, fields may be surrounded by bounding boxes, and guiding rulers. These methods specify where to write and avoid the effect of skew and overlapping with other parts of the form. They can be located on a separate sheet of paper that is located next to the form or they can be printed directly on the form. The use of guiding rulers is much better from the point of view of the quality of the output. It requires giving more instructions and, more importantly, it is more commonly used for this reason. Light rectangles are easier to remove with filters than dark lines whenever the handwriting is dark.

## Задача 1

Очищать мятые сканы документов

## Задача 2

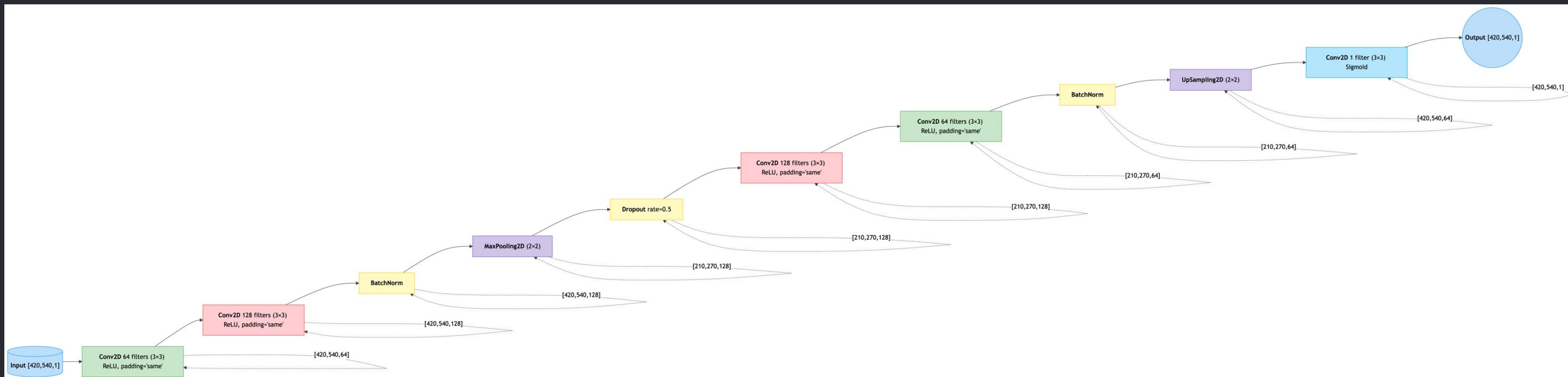
Удалять следы грязи и пятен с документов

## Особенности

Обработка фото любых размеров.

## Источник данных

Набор с конкурса Kaggle: [denoising dirty documents](#)



# Технические параметры и архитектура модели

## Параметры обучения

80 эпох, batch size 12,  
оптимизатор Adam, скорость  
0.001

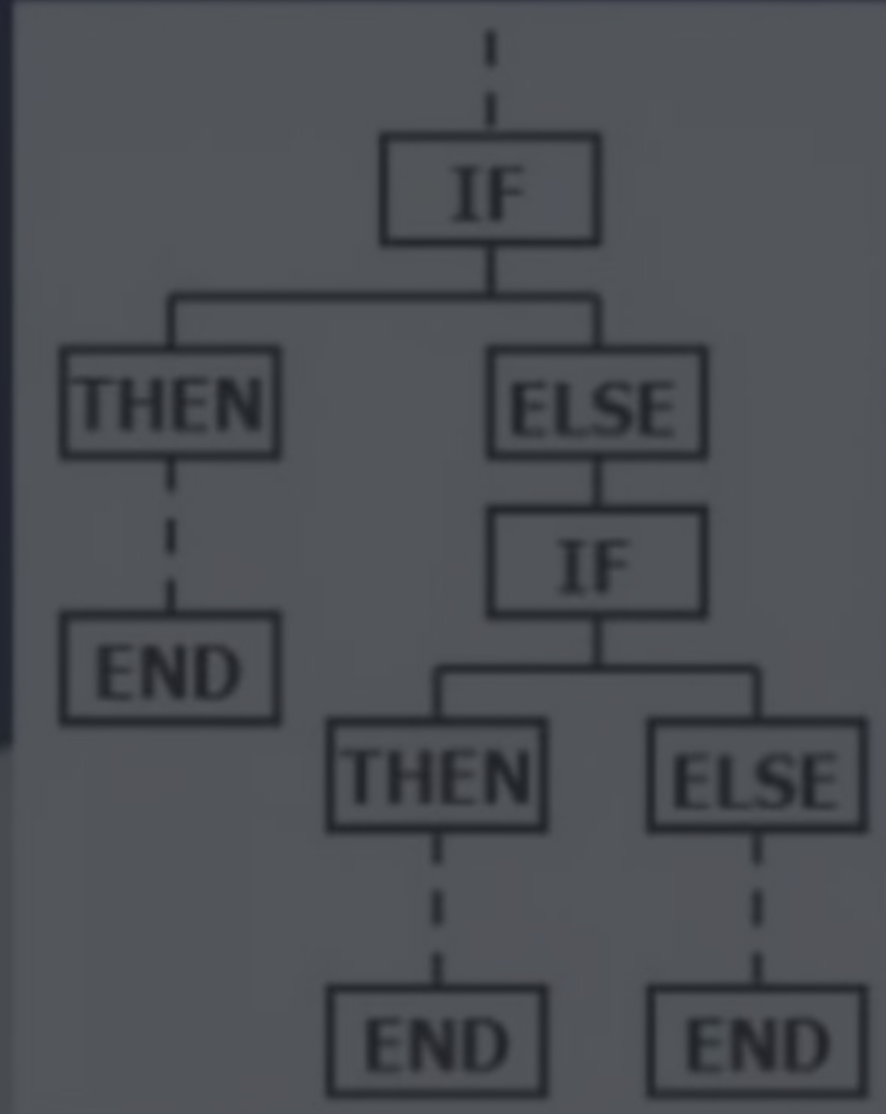
## Метрики и данные

Используется MSE, MAE для  
оценки, 85% данных обучение,  
15% валидация

## Особенности

Встроенная предобработка с  
reshape

[Ссылка](#) на статью, откуда взята архитектура автоэнкодера



**AI**

Спасибо за внимание

Готовы ответить на вопросы и обсудить дальнейшее развитие проекта.