**Исследовательский документ на тему**

**«Восстановление цвета черно-белых изображений средствами машинного обучения»**

**Команда: «Supertos Industries»**

**Глава команды: Жаворонков Никита Дмитриевич**

**Telegram: @supertosofficial**

**Ментор: Лобанов Захар Олегович**

**Telegram: @zak\_1**

**1. Обзор проекта**

Настоящее исследование посвящено разработке веб-платформы для автоматизированного восстановления цветовой информации в монохромных фотографиях с применением методов машинного обучения. Платформа имеет гибридную архитектуру, в которой динамическая обработка запросов пользователей, интеграция с разработанной нейросетевой моделью, а также управление данными реализованы с использованием FastAPI. Статические ресурсы, включая мультимедийный контент и интерфейсные компоненты, обслуживаются отдельной инфраструктурой, включающей высокопроизводительные веб-серверы и систему управления контентом (Headless CMS). Такой подход нацелен на оптимизацию производительности системы, разделение ответственности компонентов и обеспечение масштабируемости.

**2. Анализ конкурентов и существующих технических решений**

На рынке технологий обработки изображений представлен ряд решений, предлагающих автоматическое восстановление цветовой информации в монохромных изображениях. Для проведения сравнительного анализа рассмотрим основные конкурирующие продукты, структурированные по следующим параметрам: описание, сильные стороны, слабые стороны и целевая аудитория.

****2.1. DeOldify****

DeOldify — инструмент автоматической раскраски черно-белых изображений и видео. Построен на основе методов глубокого обучения, включая архитектуру генеративно-состязательных сетей (GANs), в частности, U-Net, модифицированной для улучшенной работы с изображениями.

DeOldify использует модель NoGAN — гибридную архитектуру, объединяющую элементы традиционного GAN и методов предобучения. Основной архитектурой является U-Net с генератором на базе ResNet.

Модель обучена на больших наборах данных, включающих миллионы цветных изображений из открытых источников. Для повышения качества раскраски использовался подход самосупервизии, где цветные изображения переводились в черно-белые и затем раскрашивались обратно для вычисления ошибки. Финальная донастройка выполнялась на данных с историческими изображениями.

-Imagenet для предобучения генератора.

-Наборы данных с историческими черно-белыми изображениями для дообучения и тестирования.

DeOldify предоставляет пользователям следующие возможности:

-Раскраска статичных изображений.

-Обработка видео, включая поддержку высококачественных кадров.

-Настройка интенсивности и стиля раскраски.

-Пакетная обработка.

Для работы с изображениями высокого разрешения используется подход разбиения изображения на плитки. Каждая плитка обрабатывается отдельно, а затем объединяется в итоговое изображение. Этот метод позволяет обойти ограничения памяти GPU. Параметры модели остаются неизменными для каждой плитки, обеспечивая целостность цветового стиля.

****2.2. Colourise.sg****

Colourise.sg — веб-сервис для автоматической раскраски черно-белых фотографий. Разработан на основе нейросетевой архитектуры, использующей глубокие сверточные сети (CNNs) и методы машинного обучения, оптимизированные для работы с изображениями.

Основой Colourise.sg является сверточная нейронная сеть, модифицированная для задач раскраски изображений. Архитектура сети напоминает стандартную U-Net с интеграцией элементов ResNet для улучшения передачи информации между слоями. Модель дополнительно включает слои для обработки текстур и контекстной информации, чтобы создавать естественные цвета.

Модель обучалась на больших наборах данных цветных изображений, которые использовались для симуляции черно-белых кадров. Обучение проводилось с использованием функции потерь, учитывающей структурные и цветовые различия, чтобы минимизировать цветовые искажения. Особое внимание уделялось адаптации модели для исторических изображений.

-COCO (Common Objects in Context) для обучения и проверки точности модели.

-Специализированные исторические датасеты с высококачественными черно-белыми изображениями для тонкой настройки.

Сервис Colourise.sg предлагает простой и интуитивно понятный интерфейс с минимальными настройками. Пользователи могут:

-Загружать изображения для автоматической раскраски.

-Получать результат в течение нескольких секунд.

-Использовать сервис для одиночных и пакетных изображений.

Для обработки больших изображений используется техника рескейлинга, при которой изображения временно уменьшаются до стандартного разрешения, проходят через нейросеть, а затем масштабируются обратно с дополнительной обработкой для сохранения качества. Этот подход снижает нагрузку на GPU и ускоряет обработку.

****2.3. MyHeritage InColor****

MyHeritage InColor — инструмент, интегрированный в платформу MyHeritage, предназначенный для автоматической раскраски черно-белых фотографий. Базируется на технологиях глубокого обучения и использует специализированные нейросетевые алгоритмы, адаптированные для исторических изображений.

Сервис MyHeritage InColor построен на основе архитектуры глубоких сверточных сетей (CNNs), оптимизированных для обработки изображений. Алгоритм дополнен модулями, использующими принципы генеративных моделей и контекстного анализа, чтобы воспроизводить исторически достоверные цвета. Точные детали архитектуры не раскрыты, но упоминается, что используются улучшенные подходы на основе GAN.

#### Обучение модели

Модель была обучена на больших наборах данных, включающих как современные, так и исторические изображения. Использовались методы самосупервизии, где модель предсказывала цвета для черно-белых изображений, обучаясь на результатах сравнения с оригинальными цветными данными. Дополнительно применялись техники аугментации данных и специальные исторические фильтры.

-Крупные библиотеки цифровых изображений, доступные через партнерские соглашения.

-Специализированные коллекции исторических фотографий, включающие этническое и культурное разнообразие для повышения точности раскраски.

MyHeritage InColor предоставляет:

-Раскраску черно-белых фотографий с возможностью их дальнейшего редактирования.

-Поддержку улучшения качества изображений через дополнительный инструмент Enhance.

-Интуитивно понятный интерфейс для загрузки и обработки фотографий.

-Функцию сравнения “до” и “после” для визуализации результата.

Сервис использует комбинированный подход: изображения масштабируются до оптимального разрешения для обработки, а после раскраски применяются алгоритмы улучшения качества. Это позволяет обрабатывать даже крупные фотографии с сохранением деталей и структуры.

В сервисах DeOldify, Colourise.sg и MyHeritage InColor используются подходы глубокого обучения, направленные на автоматическую раскраску изображений. Основные технологии включают:

* **Сверточные нейронные сети (CNNs)**: Все три сервиса используют архитектуры, оптимизированные для обработки изображений, такие как U-Net и ResNet.
* **Генеративные сети (GANs)**: DeOldify и MyHeritage InColor применяют модифицированные GAN для повышения реалистичности раскраски и минимизации артефактов.
* **Методы предобучения и самосупервизии**: Используются для увеличения точности раскраски, включая тренировку на цветных изображениях, переводимых в черно-белые, и последующее восстановление цвета.
* **Контекстная обработка**: У всех сервисов реализованы механизмы анализа текстуры и контекста, что улучшает реалистичность итогового результата.

Для обучения и тестирования моделей использовались обширные и разнообразные наборы данных:

* **COCO и ImageNet**: Обширные общедоступные датасеты, использованные для обучения базовых моделей.
* **Исторические изображения**: Специализированные наборы данных с черно-белыми фотографиями, позволяющие настроить модели для работы с архивными изображениями.
* **Культурно-историческое разнообразие**: Данные, охватывающие различные эпохи, этнические группы и стили фотографий, используются для повышения универсальности моделей.

Все три сервиса используют стратегии для адаптации к ограничениям ресурсов, включая масштабирование изображений, обработку плитками или уменьшение разрешения для повышения производительности. Это позволяет обеспечить стабильное качество работы, даже с большими и сложными изображениями.

Таким образом, анализ показывает схожесть подходов в архитектуре и обучении, при этом каждая из систем находит свои уникальные решения для обработки и адаптации результата под потребности пользователей.

**3. Целевая аудитория и ее потребности**

Целевая аудитория проекта включает следующие категории пользователей и их потребности:

****1. Музеи и архивы****

-Визуальное восстановление исторических документов и фотографий.

-Сохранение исторической точности и контекста при раскрашивании.

****2. Обычные пользователи и генеалоги****

-Простой и доступный инструмент для раскрашивания семейных архивных фотографий.

-Быстрая обработка изображений.

-Возможность загружать и скачивать результаты через веб-интерфейс.

****3. Творческие профессионалы (дизайнеры, художники)****

-Возможность настройки результата для использования в художественных проектах.

Проект ориентирован на различные категории пользователей, предлагая качественное и простое решение для автоматического раскрашивания изображений

**4. Описание технологий, используемых в проекте**

1. **Frontend**
   * **HTML/CSS/JavaScript**: Для создания пользовательского интерфейса, обеспечивающего взаимодействие с пользователем.
2. **Backend**
   * **FastAPI**: Современный фреймворк для разработки высокопроизводительных API с асинхронной обработкой запросов.
   * **Uvicorn**: ASGI-сервер для запуска приложения на FastAPI.
   * **PostgreSQL**: Реляционная база данных для хранения информации о загруженных изображениях и результатах их обработки.
3. **Машинное обучение**
   * **TensorFlow/Keras**: Для создания и обучения модели машинного обучения, которая выполняет раскраску изображений.
   * **OpenCV**: Для предварительной обработки изображений перед их передачей в модель.
   * **NumPy**: Для работы с данными изображений и выполнения математических операций.
4. **Асинхронная обработка**
   * **RabbitMQ**: Брокер сообщений, который используется для организации обмена сообщениями между различными сервисами и компонентов системы.
   * **Celery**: Система для асинхронного выполнения задач, использующая RabbitMQ для управления очередями и обработки запросов на разных машинах.
5. **DevOps и развертывание**
   * **Docker**: Контейнеризация приложения, что упрощает его развертывание и управление зависимостями.
   * **Nginx**: Веб-сервер для обработки входящих запросов, балансировки нагрузки и раздачи статических файлов.
   * **CI/CD**: Автоматизация процессов тестирования и развертывания с использованием GitHub Actions для упрощения интеграции и деплоя.

**5. Описание и источники наборов данных для обучения**

В процессе обучения модели машинного обучения будут задействованы следующие источники данных:

****5.1. ImageNet[1]****

Обширный набор изображений, предназначенный для обучения моделей машинного обучения. В основном используется для нейросетей распознавания, но может быть адаптирован для потребностей проекта

****5.2. Flickr Historical Images[2]****

Коллекция старых монохромных фотографий с открытой лицензией. Используется для тестирования.

5.3 **CelebA[3]**

**Набор изображений для определения лиц. Может быть адаптирован для потребностей проекта**

**5.4 DIV2K [4]**

**Набор изображений, часто использующийся для задач исправления дефектов и увеличения разрешений изображений.**

Применение комбинированного набора данных позволит достичь высокой точности и качества раскрашивания изображений с сохранением исторического контекста и деталей.

**6. Цели проекта**

1. Полностью функциональный веб-сайт с простым и интуитивно понятным интерфейсом.
2. Интегрированная модель машинного обучения для автоматической раскраски изображений.
3. REST API для взаимодействия с сервером и обработки запросов.
4. База данных для хранения информации о запросах, их статусе и результатах обработки.
5. Автоматизированное развертывание с использованием Docker и CI/CD-пайплайнов.

**8. Источники**

[1] - <https://www.image-net.org/>

[2] - <https://www.flickr.com/>

[3] - <https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>

[4] - <https://data.vision.ee.ethz.ch/cvl/DIV2K/>