

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет информатики, математики и компьютерных наук

**Программа подготовки бакалавров по направлению
09.03.04 Программная инженерия**

Овсянников Артём Сергеевич

Шейх Руслан Халедович

Мельников Андрей Александрович

КУРСОВАЯ РАБОТА

Сервис генерации изображений с помощью диффузионных моделей.
Архитектура и разработка сервиса

Научный руководитель
старший преподаватель НИУ
ВШЭ - НН

Саратовцев Артем Романович

Нижний Новгород, 2025 г.

Содержание

1	Постановка задачи	2
2	Актуальность	3
3	Обзор существующих решений	4
3.1	Создание датасетов	4
3.1.1	Краудсорсинговые сервисы и открытые датасеты	4
3.1.2	Сервисы для создания синтетических данных	4
3.2	Генерация постаревших лиц	5
3.2.1	FaceApp	5
3.2.2	AILabTools Editor	6
4	Заключение	7

1 Постановка задачи

В последние годы наблюдается значительный прогресс в области генеративных моделей, что стимулирует развитие новых методов визуализации процессов, ранее недоступных традиционными подходами. Особенно интересными являются методы, позволяющие моделировать сложную динамику изменений возраста человека. В рамках проекта предлагается разработка веб-сервиса, способного генерировать динамическое представление старения лица.

Цель данной курсовой работы заключается в создании веб-сервиса, который получает на вход фотографию лица и автоматически генерирует серию изображений, демонстрирующих постепенное старение. Проект использует модель ALAE для генерации лиц и модель LightSB для обучения вероятностного отображения между распределениями данных в структурированном латентном пространстве. Такой подход позволяет быстро и эффективно генерировать изображения постаревших людей на фото.

Реализация данного проекта опирается на достижения современной области генеративных моделей, что позволяет обеспечить непрерывность и плавность переходов между различными этапами старения. В результате пользователь получает уникальную возможность увидеть, как может измениться его внешность с течением времени, что может иметь применение как в научных исследованиях, так и в развлекательных или образовательных продуктах.

2 Актуальность

Технологии искусственного интеллекта стремительно развиваются в последние годы, особенно в области генеративных моделей, способных создавать реалистичные изображения. Эти модели находят применение в различных сферах, включая искусство, дизайн, медицину и развлекательную индустрию.

Создание реалистичных визуализаций изменений внешности человека с течением времени может быть полезным в различных контекстах: от образовательных программ до медицинских исследований. Например, такие визуализации могут помочь в обучении студентов-медиков, предоставляя наглядные примеры возрастных изменений, или использоваться в психологии для изучения восприятия собственного старения.

Разработка веб-сервиса, способного генерировать динамику старения лица на основе загруженного изображения, представляет собой актуальную задачу. Такой сервис может быть востребован широкой аудиторией, включая пользователей, интересующихся своим будущим обликом, а также специалистов, работающих в области геронтологии, косметологии и социальной психологии. Кроме того, подобные технологии могут найти применение в индустрии развлечений, например, в создании визуальных эффектов для кино и игр. Также, такой сервис может быть применим в области криминалистики или даже для автоматической генерации датасетов лиц.

В рамках данного проекта планируется использование алгоритма Light Schrödinger Bridge (LightSB), который обеспечивает эффективное обучение вероятностных отображений между распределениями данных в структурированном латентном пространстве. Это позволяет сохранять индивидуальные особенности лица при генерации изображений, что особенно важно для достижения реалистичности и персонализации результатов.

3 Обзор существующих решений

3.1 Создание датасетов

Так как данный сервис может применяться в том числе для генерации датасетов, были рассмотрены сервисы для выполнения данной задачи. Для создания датасетов могут применяться как краудсорсинговые сервисы и открытые датасеты, так и сервисы для создания синтетических данных.

3.1.1 Краудсорсинговые сервисы и открытые датасеты

1. Датасет **CelebA**, содержащий 202 599 изображений знаменитостей с 40 атрибутами (пол, возраст, форма носа и т.д.) и часто применяемый для генеративных моделей.
2. Датасет **FFHQ**, содержащий 70 000 изображений людей, изначально созданный в качестве бенчмарка для оценивания качества работы GAN.
3. Датасет **UTKFace**, состоящий из более чем 20 000 изображений лиц с аннотациями возраста, пола и этнической принадлежности.

Достоинства данного подхода:

- Высокое качество изображений
- Наличие фотографий, сделанных в разных условиях (различный климат, наличие/отсутствие головных уборов, эмоции и т.д.)

Недостатки:

- Сложность и медленность сбора датасета
- Риски нарушения конфиденциальности личных данных
- Зависимость от социальных факторов и сложность создания репрезентативной выборки

3.1.2 Сервисы для создания синтетических данных

1. Сайт <https://thispersondoesnotexist.com/>, создающий новое изображение человека при каждом заходе
2. Датасет **Generated Photos** от icons8, содержащий миллионы сгенерированных лиц

Достоинства данного подхода:

- Возможность создания очень большого датасета
- Балансировка количества изображений по лейблам
- Низкая стоимость создания

Недостатки:

- Слабая реалистичность сгенерированных фотографий
- Необходимость обучения моделей генерации
- Требование к сбору обучающей выборки

3.2 Генерация постаревших лиц

Для генерации постаревших лиц могут использоваться следующие сервисы:

3.2.1 FaceApp

Достоинства:

- Высокое качество преобразований (ИИ-фильтры)
- Множество функций и фильтров
- Простой интерфейс
- Быстрая обработка

Недостатки:

- Платная подписка для большинства фильтров
- Споры о конфиденциальности данных
- Работа только на сертифицированных устройствах Google
- Невозможность выбора конкретного возраста

3.2.2 AILabTools Editor

Достоинства:

- Реалистичная обработка изображений
- Изменение пола, возраста и фона
- Простой интерфейс
- Указание конкретного возраста

Недостатки:

- Ограниченная бесплатная версия (5 кредитов/месяц)
- Генерация занимает 5 секунд
- Требуется сертифицированные устройства Google
- Нет выбора конкретного возраста
- Отсутствие пакетной обработки

4 Заключение

В курсовой работе был разработан полноценный веб-сервис, демонстрирующий наработки российских учёных в области генеративного искусственного интеллекта для быстрого обучения моделей для решения задачи трансляции представлений (image-to-image translation). Наш сервис использует LightSB для преобразования данных между распределениями молодых и старых людей, чтобы делать человека на фото старше, чем он есть на самом деле.

Актуальность проекта демонстрируется наличием веб-сервиса и MLOps пайплайна, позволяющими наглядно показать результаты работы LightSB и ALAE с помощью генерации лиц постаревших людей благодаря простому интерфейсу и сбору статистики для модели.

Данный проект может быть полезен для генерации датасетов лиц людей и позволяет в удобном формате создавать новые изображения.