

Министерство науки и высшего образования Российской Федерации  
федеральное государственное автономное  
образовательное учреждение высшего образования  
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники  
Направление подготовки: 09.04.01 Информатика и вычислительная техника  
ООП/ОПОП: Искусственный интеллект и машинное обучение  
Отделение школы (НОЦ): Отделение информационных технологий

### **ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА МАГИСТРАНТА**

Тема работы
<b>Разработка нейросетевого алгоритма восстановления изображений лиц</b> УДК 004.032.26:004.932

Обучающийся

Группа	ФИО	Подпись	Дата
8ВМ13	Куулар Чингис Тимурович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Иванова Ю.А.	К.Т.Н.		

### **КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:**

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Былкова Т.В.	К.Э.Н.		

По разделу «Социальная ответственность»

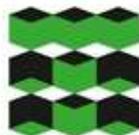
Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ООД ШБИП	Федорчук Ю.М.	Д.Т.Н.		

### **ДОПУСТИТЬ К ЗАЩИТЕ:**

Руководитель ООП, должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын В.Г.	Д.Т.Н.		

## ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ООП

Код компетенции	Наименование компетенции
<b>Универсальные компетенции</b>	
<b>УК(У)-1</b>	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий
<b>УК(У)-2</b>	Способен управлять проектом на всех этапах его жизненного цикла
<b>УК(У)-3</b>	Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели
<b>УК(У)-4</b>	Способен применять современные коммуникативные технологии, в том числе на иностранном (-ых) языке (-ах), для академического и профессионального взаимодействия
<b>УК(У)-5</b>	Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия
<b>УК(У)-6</b>	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки
<b>Общепрофессиональные компетенции</b>	
<b>ОПК(У)-1</b>	Способен самостоятельно приобретать, развивать и применять математические, естественно-научные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
<b>ОПК(У)-2</b>	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач
<b>ОПК(У)-3</b>	Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями
<b>ОПК(У)-4</b>	Способен применять на практике новые научные принципы и методы исследований
<b>ОПК(У)-5</b>	Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
<b>ОПК(У)-6</b>	Способен разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования
<b>ОПК(У)-7</b>	Способен адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий
<b>ОПК(У)-8</b>	Способен осуществлять эффективное управление разработкой программных средств и проектов
<b>Профессиональные компетенции</b>	
<b>ПК(У)-1</b>	Способен разрабатывать и администрировать системы управления базами данных
<b>ПК(У)-2</b>	Способен проектировать сложные пользовательские интерфейсы
<b>ПК(У)-3</b>	Способен управлять процессами и проектами по созданию (модификации) информационных ресурсов
<b>ПК(У)-4</b>	Способен осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ
<b>ПК(У)-5</b>	Способен проектировать и организовывать учебный процесс по образовательным программам с использованием современных образовательных технологий



Министерство науки и высшего образования Российской Федерации  
федеральное государственное автономное  
образовательное учреждение высшего образования  
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники  
Направление подготовки: 09.04.01 Информатика и вычислительная техника  
ООП/ОПОП: Искусственный интеллект и машинное обучение  
Отделение школы (НОЦ): Отделение информационных технологий

УТВЕРЖДАЮ:  
Руководитель ООП/ОПОП  
Спицын В.Г.  
(Подпись) (Дата) (ФИО)

**ЗАДАНИЕ  
на выполнение выпускной квалификационной работы**

Обучающийся:

Группа	ФИО
8ВМ13	Куулар Чингис Тимурович

Тема работы:

Разработка нейросетевого алгоритма восстановления изображений лиц	
<i>Утверждена приказом директора (дата, номер)</i>	от 09.02.2023 №40-57/с

Срок сдачи обучающимся выполненной работы:

**ТЕХНИЧЕСКОЕ ЗАДАНИЕ:**

<b>Исходные данные к работе</b>	Объектом исследования является алгоритм восстановления лиц на изображениях при помощи генеративно-сопоставительной сети.
<b>Перечень разделов пояснительной записки подлежащих исследованию, проектированию и разработке</b>	<ol style="list-style-type: none"> <li>1. Восстановление изображений при помощи генеративно-сопоставительной сети;</li> <li>2. Реализация генеративно-сопоставительной сети;</li> <li>3. Обучение модели и результаты;</li> <li>4. Раздел ВКР «Финансовый менеджмент ресурсоэффективность и ресурсосбережение».</li> <li>5. Раздел ВКР «Социальная ответственность».</li> <li>6. Раздел ВКР на английском языке.</li> </ol>
<b>Перечень графического материала</b> <i>(с точным указанием обязательных чертежей)</i>	Презентация в формате *.pptx
<b>Консультанты по разделам выпускной квалификационной работы</b>	

<i>(с указанием разделов)</i>	
<b>Раздел</b>	<b>Консультант</b>
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Доцент ОСГН, ШБИП к.э.н. Былкова Т.В.
Социальная ответственность	Профессор ООД, ШБИП д.т.н. Федорчук Ю.М.
Раздел на иностранном языке	Доцент ОИЯ, ШБИП к.п.н. Сидоренко Т.В.
<b>Названия разделов, которые должны быть написаны на иностранном языке:</b>	
Обзор методов восстановления изображений	
Обзор реализуемой модели	

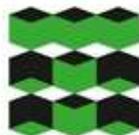
<b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b>	
---	--

**Задание выдал руководитель / консультант (при наличии):**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ, ИШИТР	Иванова Юлия Александровна	к.т.н.		

**Задание принял к исполнению обучающийся:**

Группа	ФИО	Подпись	Дата
8ВМ13	Куулар Чингис Тимурович		



Министерство науки и высшего образования Российской Федерации  
федеральное государственное автономное  
образовательное учреждение высшего образования  
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники  
Направление подготовки: 09.04.01 Информатика и вычислительная техника  
Уровень образования: Магистратура  
ООП/ОПОП: Искусственный интеллект и машинное обучение  
Отделение школы (НОЦ): Отделение информационных технологий  
Период выполнения весенний семестр 2022/2023 учебного года

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН  
выполнения выпускной квалификационной работы**

Обучающийся:

Группа	ФИО
8ВМ13	Куулар Чингис Тимурович

Тема работы:

<b>Разработка нейросетевого алгоритма восстановления изображений лиц</b>
--

Срок сдачи обучающимся выполненной работы:

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
	Основная часть	70
	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
	Социальная ответственность	10
	Обязательное приложение на иностранном языке	5
	Оформление пояснительной записки	5

**СОСТАВИЛ:**

**Руководитель ВКР**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Иванова Ю.А.	к.т.н.		

**СОГЛАСОВАНО:**

**Руководитель ООП/ОПОП**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын В.Г.	д.т.н.		

**Обучающийся**

Группа	ФИО	Подпись	Дата
8ВМ13	Куулар Чингис Тимурович		

## РЕФЕРАТ

Выпускная квалификационная работа выполнена на 98 страницах, содержит 16 рисунков, 17 таблиц, 47 источников, 2 приложения.

Ключевые слова: восстановление лиц, увеличение разрешения, генеративно-состязательные сети, глубокое обучение, нейронные сети.

Объектом исследования является алгоритм восстановления лиц на изображениях пострадавших от различных деградаций на основе генеративно-состязательной сети.

Цель работы – реализация генеративно-состязательной сети для решения задачи восстановления лиц на изображениях.

В процессе исследования проводились работы по изучению методов восстановления лиц на изображениях и увеличения их разрешения. В ходе работы были рассмотрены существующие подходы к решению задачи восстановления лиц.

В результате исследования была реализована генеративно-состязательная модель, решающая задачу восстановления лиц на изображениях. Также проведены численные эксперименты для оценки качества реализованной модели, и сравнения её с конкурентами.

Область применения: безопасность и криминалистика – улучшение четкости и видимости лиц, с камер наблюдения или мобильных устройств; реставрация старинных фотографий – может восстановить детали и цвета старых или поврежденных фотографий, картин; развлечения и образование – может улучшить визуальный опыт при просмотре фильмов, игр на разных устройствах и экранах.

## СОДЕРЖАНИЕ

Введение.....	9
1. Теоретический обзор решаемой проблемы.....	11
1.1. Проблема суперразрешения.....	11
1.2. Обзор методов.....	14
1.3. Оценка качества моделей.....	15
1.4. Решаемая задача.....	17
2. Обзор реализуемой модели.....	18
2.1. Общий обзор реализуемой модели.....	18
2.2. Модуль удаления деградаций.....	20
2.3. Генеративный модуль.....	22
2.4. Модуль пространственного преобразования.....	24
2.5. Функционал качества модели.....	27
3. Реализация модели.....	28
4. Обучение модели.....	32
Заключение.....	36
5 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение..	38
5.1. Предпроектный анализ.....	38
5.1.1. Потенциальные потребители результатов исследования.....	38
5.1.2. Анализ конкурентных технических решений с позиции ресурсоэффективности и ресурсосбережения.....	40
5.1.3. SWOT-анализ.....	43
5.1.4. Оценка готовности проекта к коммерциализации.....	44
5.1.5. Методы коммерциализации результатов научно-технического исследования.....	46
5.2. Инициация проекта.....	46
5.2.1. Цели и результат проекта.....	46

5.3.	Планирование управления научно-техническим проектом .....	47
5.3.1.	Иерархическая структура работ проекта .....	47
5.3.2.	План проекта .....	48
5.3.3.	Бюджет научного исследования .....	49
5.4.	Определение ресурсной (ресурсосбережение), финансовой эффективности исследования .....	54
5.4.1.	Оценка сравнительной эффективности исследования .....	54
5.5.	Вывод .....	57
ЗАДАНИЕ ДЛЯ РАЗДЕЛА .....		58
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ» .....		58
6.	Социальная ответственность .....	60
6.1.	Введение .....	60
6.2.	Производственная безопасность .....	60
6.2.1.	Вредные производственные факторы .....	60
6.2.2.	Опасные производственные факторы .....	68
6.3.	Экологическая безопасность .....	72
6.4.	Безопасность в чрезвычайных ситуациях .....	73
Заключение .....		75
Список использованной литературы .....		76
Приложение А .....		81
1. METHODS OVERVIEW .....		82
2. OVEVIEW OF ARCHITECTURE USED BY MODEL .....		83
2.1.	Developing model brief overview .....	83
2.2.	Degradation removal module .....	85
Приложение Б .....		94

## Введение

Восстановление лиц является одним из частных случаев проблемы увеличения разрешения изображений. Проблема решает следующую задачу: имеется изображение, содержащее лицо человека, пострадавшее по неизвестным причинам, таким как: уменьшение разрешения изображения вследствие физических ограничений захватывающего устройства, размытие, полученное при получении изображения, артефакты сжатия и так далее; данное изображение далее используется для получения высококачественного изображения лица человека. Задача ещё более усложняется применительно к изображениям, полученным из аналоговых источников таких, как например плёночные фотоаппараты, вследствие более сложной деградации.

В предыдущих работах была использована генеративно состязательная сеть, которая занималась комплексным решением задачи увеличения разрешения изображений. Применительно к изображениям лиц, данная модель также показывала хорошие результаты, однако она была вычислительно сложной, вследствие своей более обобщённой задачи. Было принято решение рассмотреть более частный случай увеличения разрешения изображений, и перейти к задаче восстановления лиц. Это сложная обратная задача, так как существует большое количество возможных отображений деградировавших изображений лиц в высококачественные. Восстановление лиц также, как и увеличение разрешения изображений имеет множество применений.

В настоящее время было предложено большое количество методов, позволяющих получить высококачественное изображение лица из пострадавшего вследствие неизвестных деградаций [21, 22, 23, 24, 25, 26, 27]. Задача восстановления является частным случаем более общей: увеличение разрешения изображений. В предыдущих работах было использовано решение на основе генеративно состязательной сети. Предложенное в рамках

данной работы программное решение будет основано на схожих с предыдущими работами принципах.

В данной работе представлен обзор некоторых методов восстановления изображений лиц, а также описана общая теория и концепции, используемые при решении данной задачи.

Представлена реализация генеративно-сопоставительной сети по реализации восстановления лиц.

Целью данной работы является реализация генеративно-сопоставительной модели для повышения разрешения изображений лиц.

Для достижения поставленной цели, в настоящей работе решаются следующие задачи:

1. изучение проблемы восстановления лиц;
2. изучение метрик, используемых в задаче сверхразрешения;
3. изучение методов восстановления лиц на изображении;
4. изучение проблем производительности методов;
5. реализация генеративно-сопоставительной сети.

## 1. Теоретический обзор решаемой проблемы

### 1.1. Проблема суперразрешения

В большинстве приложений, использующих цифровые изображения, обычно требуются изображения или видео в высоком разрешении для их последующей обработки и анализа. Стремление к высокому разрешению изображений обусловлено двумя причинами: увеличение разрешающей способности изображения способствует улучшению интерпретации изображения человеком и улучшению автоматического машинного восприятия [1]. Разрешение изображения описывает детали, содержащиеся в изображении, чем выше разрешение, тем больше деталей содержится в изображении. Разрешение цифрового изображения можно классифицировать по-разному: разрешение в пикселях, пространственное разрешение, спектральное разрешение, временное разрешение и радиометрическое разрешение. В данном контексте интересным было бы рассмотреть пространственное разрешение.

Цифровое изображение состоит из множества элементов, называемых пикселями. Пространственное разрешение характеризует плотность пикселей в изображении и измеряется в пикселях на единицу площади. На рисунке 1 показана классическая тестовая мишень для определения пространственного разрешения системы формирования изображений [1].

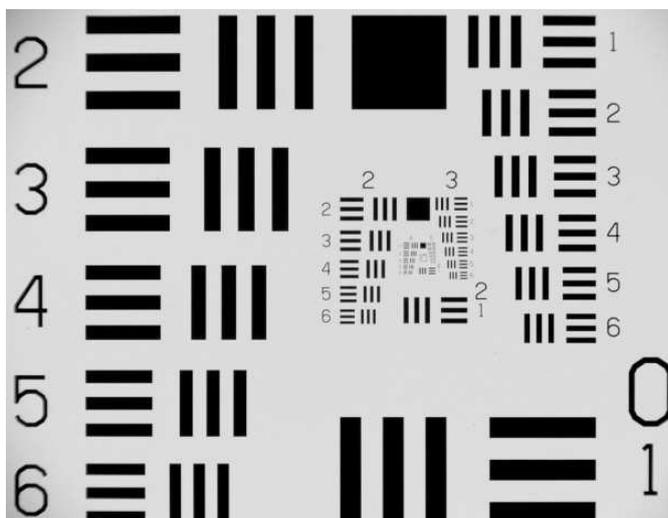


Рисунок 1 – Испытательная мишень

Разрешающая способность пространственного разрешения изображения ограничивается датчиками, фиксирующими изображение, или устройством получения изображения. Современный датчик захвата изображения обычно представляет собой активно-пиксельный датчик с зарядовой связью, так называемая ПЗС матрица, или комплементарную структуру металл-оксид-полупроводник, также называемая КМОП [2]. Эти датчики обычно располагаются в матричном виде для захвата сигналов, соответствующих двумерному изображению. Размер сенсора или количество элементов сенсора на единицу площади определяет пространственное разрешение захватываемого изображения. Чем выше плотность расположения датчиков, тем более высокое пространственное разрешение будет иметь изображение, полученное системой. Система формирования изображений с недостаточным количеством детекторов будет генерировать изображения низкого разрешения с так называемым эффектом «лесенки», по теореме Котельникова, обусловленного низкой частотой пространственной дискретизации. Наиболее простым, с точки зрения обывателя, способом повышения пространственного разрешения системы формирования изображений является простое увеличение плотности датчиков за счет уменьшения их физического размера. Однако при уменьшении физического размера датчика количество света, падающего на каждый датчик, также уменьшается, вызывая, так называемый, дробовой шум. Кроме того, с увеличением плотности датчиков на сенсоре увеличивается стоимость аппаратного обеспечения сенсора. Отсюда можно сделать вывод, что пространственное разрешение получаемого изображения, ограничено размером датчика, устанавливаемого на сенсоре.

Хотя датчики захвата изображения ограничивают пространственное разрешение изображения из-за размытия линз, связанного с функцией точечной коррекции датчика, эффектов абберации линз, дифракции на диафрагме и оптического размытия из-за движения захватывающего устройства во время съёмки, имеется также ограничение на качество и

количество деталей захватываемого изображения, устанавливаемое оптикой [15]. Конструирование микросхем и оптических компонентов для получения изображений с очень высоким разрешением является непомерно дорогим и непрактичным в большинстве реальных приложений, например для широко используемых камер видеонаблюдения и встроенных камер смартфонов. Помимо стоимости, разрешение камеры захвата изображения также ограничено скоростью камеры и аппаратной памятью устройства. Также, например, для камер, используемых в спутниковой съёмке, использование высококачественных фиксирующих устройств ограничено физическими ограничениями самого спутника. Отсюда следует что способ увеличения разрешения изображения посредством увеличения разрешающей способности самого фиксирующего устройства имеет физические ограничения, которые трудно и дорого разрешить. Другим способом разрешения проблемы увеличения пространственного разрешения изображения является, принятие факта ухудшения качества изображения при съёмке, и вместо улучшения камеры, программно обрабатывать полученный сигнал с камеры, для увеличения пространственного изображения устройства. Эти методы называются методами суперразрешения [16].

Суперразрешение (SuperResolution SR) – это метод, который позволяет создавать изображения высокого разрешения (High Resolution) из нескольких наблюдаемых изображений низкого разрешения (Low Resolution), тем самым увеличивая высокочастотные компоненты и устраняя ухудшения, вызванные процессом формирования изображения камерой имеющая низкую разрешающую способность [1]. Задача суперразрешения является одной из наиболее активных областей исследований с момента появления работы Цая и Хуанга в 1984 году [14]. За последние два десятилетия было предложено множество методов, представляющих подходы от частотной до пространственной области и от обработки сигналов до машинного обучения.

В задаче суперразрешения изображения низкого качества могут быть получены из изображений высокого качества при помощи функции деградации по формуле:  $I_x = D(I_y; \sigma)$ , где  $D$  – это функция деградации изображения,  $I_y$  – это изображение с высоким пространственным разрешением,  $I_x$  – это изображение с низким разрешением,  $\sigma$  – это шум [17]. Параметры деградации  $D$  и  $\sigma$  остаются неизвестными, при решении задачи суперразрешения известны только изображение с высоким разрешением (HR) и изображение с низким разрешением (LR). Задачей суперразрешения является нахождение обратной функции  $D^{-1}$  деградации, используя только предоставленные изображения HR и LR.

## 1.2. Обзор методов

Проблеме восстановления лиц на изображениях посвящено большое количество работ. Некоторые работы такие как [21, 22, 23], используют характерные для лиц признаки, такие как: особые якорные точки на лицах [23, 21], применение масок с лицевыми компонентами [22]. Данные работы показывают, что геометрические признаки лица, имеют важное значение для получения очертаний лица и лицевых деталей. Однако данные признаки могут быть найдены только на самих ухудшенных изображениях, поэтому в случае с лицами на старых или нечётких фотографиях, будет очень сложно получить данные характеристики лица. Также хотя они дают хорошее представление о геометрии лица, они не могут дать информацию о текстуре, например цвет, текстуру зрачка.

Также, например модель DFDNet [27], использует схожий подход, основанный на ссылочных признаках. Данные признаки представляют из себя ссылки на изображения, а точнее на части изображений, сходные с тем, которые подаются на вход модели. Модель DFDNet, строит словарь лицевых признаков, такие как глаза, рот, нос на основе изображений из обучающей выборки, при помощи свёрточных слоёв. Далее данный словарь признаков, используется, чтобы помочь процессу восстановления изображения. Однако данный подход имеет недостаток, так хотя модель

хорошо восстанавливает, данные участки на лице, зачастую области, находящиеся вне регионов, из словаря лицевых признаков, претерпевают визуальные ухудшения, например страдают волосы, контур лица, подбородок и так далее. Более того, при ограниченном размере словаря, полученные после генерации изображения, страдают от недостатка разнообразия.

Другой подход использует те же методы что и задача увеличения разрешения изображений. Например, подходы [24, 25, 26] используют генеративную инверсию [24, 25], или вариационные автокодировщики [26]. Подходы, основанные на генеративной инверсии, основаны на том, чтобы входное изображение подать на выходы генератора, для получения векторов в скрытом пространстве генератора, после чего сгенерировать новое изображение. Идея заключается в том, что близкие друг к другу вектора в скрытом пространстве генератора скорее всего получат одинаковое по смыслу изображение, однако сама оно будет более качественным, что обеспечивается состязательной ошибкой и ошибкой восприятия. Подход на основе автокодировщиков использует схожую идею, основанную на кодировании низкокачественного изображения в скрытое пространство и, затем, получении при помощи декодировщика изображения более высокого качества. Идея основана на том, что модель преследуя цель минимизации потери между входным и выходным изображениями, не будет стараться в своих весах запомнить артефакты, характерные для низкокачественных изображений. Однако эти подходы требуют наличия пар низкокачественное – высококачественное изображение. Однако не всегда есть возможность получить высококачественные изображения лиц. Более того данные модели вычислительно сложные из-за того, что они требуют большое количество проходов для запоминания отображения низкокачественного изображения в высококачественное.

### **13. Оценка качества моделей**

Для оценки моделей используются различные метрики. Следует перечислить некоторые из них. Ряд методов оценки качества изображения

используются для оценки разных параметров сгенерированного изображения. Все эти метрики можно в целом классифицировать на две категории – **субъективные** и **объективные**.

Субъективные метрики основаны на оценке восприятия человеком-наблюдателем, тогда как объективные метрики основаны на вычислительных моделях, которые пытаются оценить качество изображения. Субъективные метрики часто являются более «точными для восприятия», однако некоторые из этих метрик неудобны, трудоемки или дороги для вычисления. Другая проблема заключается в том, что эти две категории метрик могут не соответствовать друг другу. Далее некоторые из метрик оценки моделей будут рассмотрены.

Одной из часто используемых метрик является метрика PSNR – пиковое отношение сигнал/шум. Она является объективной, и обратно пропорциональна логарифму средней квадратичной ошибки между настоящим изображением высокого разрешения и сгенерированным [15].

$$MSE = \frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2,$$
$$PSNR = 10 * \log_{10} \left( \frac{L^2}{MSE} \right),$$

где  $L$  – это максимально возможное значение пикселя. Так как PSNR измеряет только разницу между значениями пикселей, то данная метрика может не точно оценивать воспринимаемое качество изображения.

Ещё одной метрикой является метрика структурного сходства (SSIM). Это субъективная метрика, используемая для измерения структурного сходства между изображениями на основе трех относительно независимых сравнений, а именно: яркости, контрастности и структуры [1(25)]. Метрика может быть вычислена как взвешенное произведение яркости, контраста и структуры изображения, вычисленных независимо друг от друга.

$$SSIM(I, \hat{I}) = \frac{(2\mu_I \mu_{\hat{I}} + C_1)(\sigma_{I\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)},$$

где  $\mu_I$  – среднее значение конкретного изображения,  $\sigma_I$  – стандартное отклонение конкретного изображения,  $\sigma_{II}$  – ковариация между изображениями,  $C_1$  и  $C_2$  – константы для стабилизации. Зачастую при оценке пользуются вычислением локальной метрики SSIM, в разных окнах изображения, при последующей взвешенной оценке получившихся результатов, для стабилизации процесса оценки. Поскольку SSIM оценивает качество реконструкции с точки зрения зрительной системы человека, эта метрика лучше соответствует требованиям оценки восприятия. Также для оценки моделей используются такие метрики как: средняя оценка мнений, являющийся ручным способом определения результатов модели, где людей просят оценить изображение, после чего результаты агрегируются, и в качестве показателя используется средний результат оценки, критерий достоверности информации, визуальная достоверность информации.

#### **1.4. Решаемая задача**

В ходе работы была поставлена задача, по восстановлению лиц на изображениях. Задача формулируется следующим образом: необходимо создать модель глубокого обучения, которая на этапе предсказания, получает на вход изображение, на котором присутствуют различные деградации, на выходе, при этом, модель должна, создавать изображение на котором отсутствуют данные деградации, либо же их присутствие сведено к минимуму. Деградации на изображении включают в себя: размытие, артефакты сжатия, наложение неизвестного шума, инверсия цветов различных случайных пикселей, и так далее.

При этом в качестве используемой обучающей выборки, модель имеет исходные изображения, на которых отсутствуют деградации GT (ground truth). Данные изображения используются моделью, для вычисления ошибки, на этапе обучения, при сравнении их с полученными после прямого прохода сгенерированными изображениями. Таким образом представляя из себя целевые значения. Значения, которые подаются на вход модели, генерируются самой моделью, искусственным наложением, различных

деградаций на исходные изображения. Данные изображения являются входными значениями обучающей выборки.

Для решения задачи используется генеративно состязательная модель. Ниже будет представлено описание её архитектуры и деталей реализации.

## **2. Обзор реализуемой модели**

### **2.1. Общий обзор реализуемой модели**

Ниже на рисунке 2 представлена общая схема работы используемой модели. Модель состоит из нескольких частей. На вход модели подаётся изображение низкого качества, пострадавшее от различного рода ухудшений, в общем случае неизвестных. На выходе модели получается изображение высокого качества максимально близкое к правде. Модель состоит из следующих частей: сеть U-Net, отвечающая за удаление деградаций на изображении и предобученной сети генератора, в данном случае StyleGAN [27]. StyleGAN был выбран в первую очередь потому, что показал, что он может генерировать правдоподобные изображения лиц, при этом изображения генерируемые сетью высокого разрешения. Сеть генератора StyleGAN используется для того, чтобы получать из её весов достоверные признаки характерные для изображений лиц. Используется тот же принцип что и в генеративной инверсии: скрытое пространство сети U-Net, используется для генерирования изображения лица через модель генератора StyleGAN.

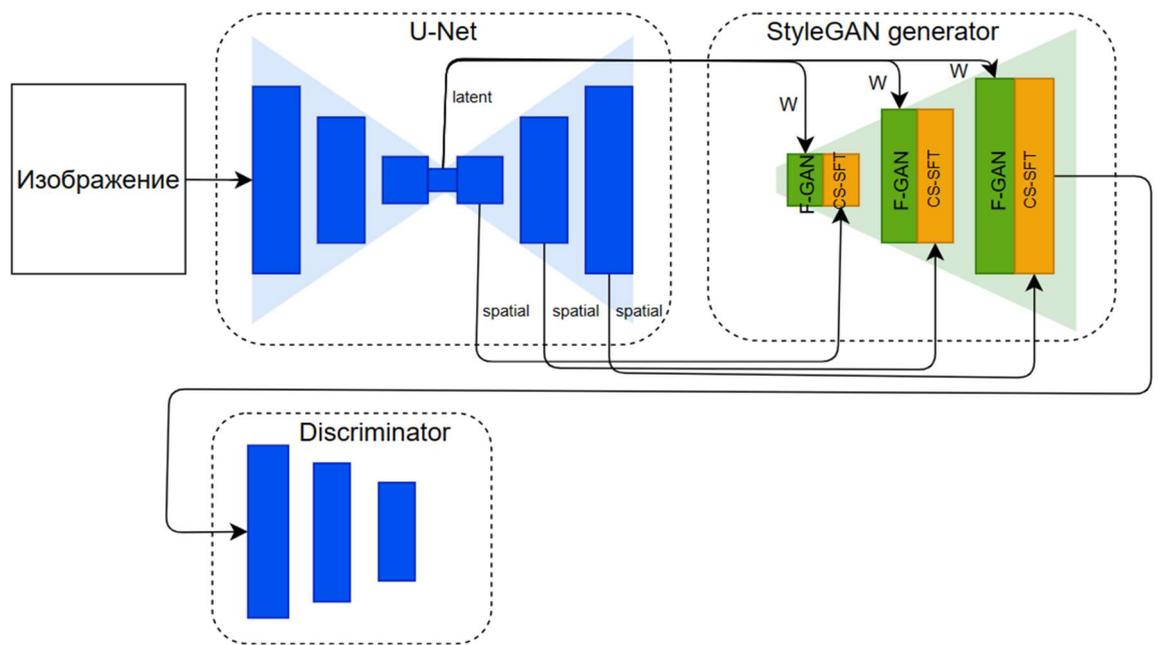


Рисунок 2 – модель сети

Как можно заметить из схемы две сети связаны между собой. В первую очередь вектора из скрытого пространства модели U-Net, проходя через классическую для архитектуры StyleGAN, сеть отображения, подаются на все слои модели StyleGAN. Также вектора из скрытых пространств модели U-Net, далее подаются на блоки преобразования пространственных признаков, которые осуществляют модуляцию сигнала полученного от слоёв генератора Style-GAN. Сеть U-Net в частности используется для получения двух типов векторов в скрытом пространстве: 1) вектор в скрытом пространстве кодирующий скрытые признаки изображения, которые выделила сеть U-Net, и который затем привязывается к наиболее близкому вектору из скрытого пространства генератора Style-GAN; 2) пространственные тензоры получаемые из слоёв модели U-Net, когда она разворачивает вектор из скрытого пространства, которые затем используются для модуляции сигналов получаемых из слоёв генератора модели StyleGAN.

После чего в силу архитектурных особенностей модели StyleGAN, вектор из скрытого пространства модели U-Net привязывается, в несколько векторов  $W$ , несколькими полносвязными слоями. Эта привязка является аффинным преобразованием из пространства скрытого слоя U-Net в скрытый слой модели StyleGAN, так как последняя в своей архитектуре не использует

вектора из скрытого пространства напрямую, вместо этого вектора из скрытого пространства выступают как регуляризаторы слоёв адаптивной нормализации экземпляра, гарантируя линейную делимость признаков из скрытого пространства векторов. Слои сети StyleGAN, при получении вектора близкого к вектору исходного изображения, генерируют промежуточные свёрточные признаки, которые предоставляют разнообразные лицевые детали, так как предобученная сеть, должна была запомнить данные признаки в весах своих слоёв. Пространственные признаки далее используются для модуляции признаков, генерируемых при помощи StyleGAN, при помощи слоёв изменения пространственных признаков. Более того, так как сеть StyleGAN позволяет линейно разделить признаки из скрытого пространства, модуляция будет происходить от грубых признаков к более тонким.

Целью модели обучения будет минимизация общей ошибки которая будет состоять из нескольких компонент, а именно: состязательная ошибка получаемая сетью дискриминатора, которая отвечает за общую достоверность генерируемых изображений, состязательная ошибка получаемая локальными дискриминаторами нацеленными на различение лицевых компонент, то есть глаза и рот, а также используемая ещё с прошлой работы ошибка потери восприятия.

Таким образом во время этапа обучения будут обучаться веса двух моделей, сети U-Net, призванной для извлечения признаков из исходного изображения и отображения его в пространство скрытых векторов, и сети дискриминатора, которая отвечает за общую достоверность модели обучения.

## **2.2. Модуль удаления деградаций**

Ниже на рисунке 3 представлена обобщённая схема модели U-Net [28]. Как можно видеть из рисунка модель вначале понижает размерность пространства, для того чтобы извлечь семантические признаки изображения, то есть отвечает на вопрос, что изображено на картинке, при этом увеличивая

количество каналов в сигнале. После чего идут слои, повышающие размерность сигнала, и понижая количество каналов, восстанавливая изображение используя извлечённые семантические признаки.

Также для того, чтобы модель не теряла пространственные характеристики на изображении в модели, присутствуют пропускающие связи. Так как при понижении размерности изображения, модель не запоминает геометрические признаки изображения, то есть не может ответить на вопрос, где это изображено. Поэтому при помощи остаточных связей, при повышении размерности изображения, модель использует ранее извлечённые признаки, на этапе понижения размерности. Эти признаки позволяют смешать абстрактную информацию высокого уровня, выделенную во время процесса понижения разрешения, с пространственной информацией из предыдущих слоёв сети.

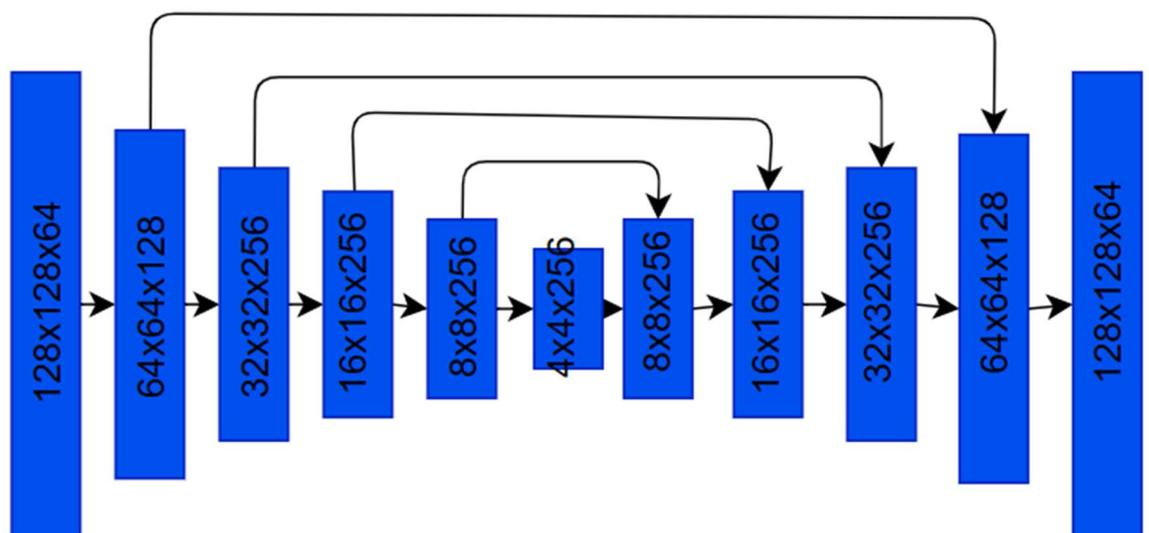


Рисунок 3 – схема сети U-Net

Модуль, основанный на сети U-Net, предназначен для явного удаления деградаций на изображении, как было ранее сказано, данная модель ориентирована на то, чтобы не запоминать информацию о различного рода артефактах деградации: размытие, сжатие и так далее, концентрируясь на семантических и пространственных признаках. Это достигается посредством извлечения чистых от деградации многомерных векторов признаков: скрытых признаков, полученных после последнего слоя свёртки сети, а также

пространственных признаков, получаемых каждый раз после увеличения разрешения изображения.

$$F_{latent}, F_{spatioal} = UNet(x)$$

Вектора скрытых признаков используются для последующей привязки их на пространство векторов модели StyleGAN, а пространственные признаки используются для модуляции признаков, получаемых из модели StyleGAN.

### **2.3. Генеративный модуль**

Ниже на рисунке 4 представлена схема генератора, используемая в модели StyleGAN [27]. Как можно заметить из схемы на рисунке, генератор не использует вектора из скрытого пространства напрямую для генерации новых изображений лиц, вместо этого генератор использует вектора из скрытого пространства в слое адаптивной нормализации экземпляра AdaIN, которое по сути, производит масштабирование и сдвиг входного сигнала относительно вектора из скрытого пространства. Более того формируемый из модели U-Net вектор также не напрямую подаётся на входы генератора, вместо этого они проходят через несколько полносвязных слоёв, что в свою очередь является аффинным преобразованием из одного пространства в другое. В результате которых получаются вектора из скрытого пространства модели генератора.

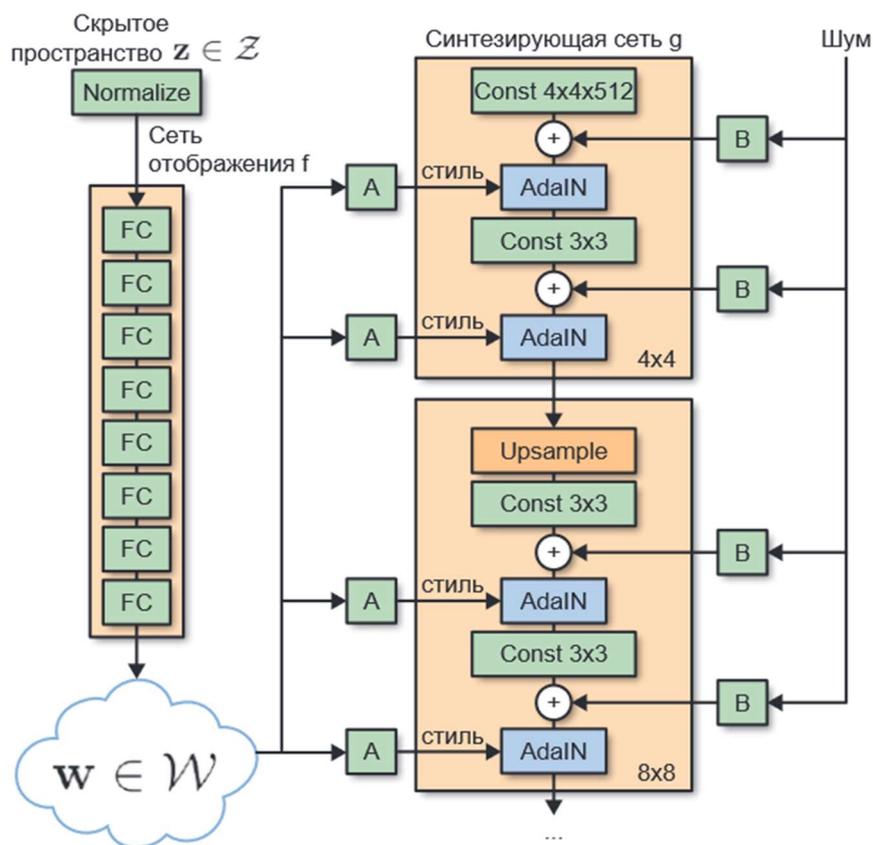


Рисунок 4 – схема генератора StyleGAN.

Была использована предобученная генеративная сеть модели StyleGAN, так как согласно [29, 30], было получено, что генераторы, используемые для получения изображений лиц, захватывают статистики над лицевыми признаками в своих весах в свёрточных слоях. Процесс генерации изображения с улучшенным лицом происходит следующим образом: вектора из скрытого пространства, полученные на предыдущем этапе моделью U-Net, используются генерации признаков из модели StyleGAN, однако данные признаки являются промежуточными. Таким образом генератор не используется для получения высокоточного изображения лица напрямую. Вместо этого, промежуточные признаки модулируются, при помощи пространственных признаков, также получаемых моделью U-Net на предыдущем этапе, для обеспечения хорошей точности с полученным на входе изображением.

Таким образом, вектора полученные на этапе кодирования моделью U-Net  $F_{latent}$ , при помощи предобученных линейных слоёв StyleGAN,

кодируются в скрытое пространство  $W$ , для сохранения семантики изображения, на протяжении этапов генерации в генераторе StyleGAN. После чего вектора из пространства  $W$ , проходят через каждый свёрточный слой генератора, и формируют признаки для каждого разрешения.

$$W = FC(F_{lat})$$

$$F_{StyleGAN} = StyleGAN(W)$$

Следует отметить, что так как в модели используется уже готовый генератор модели StyleGAN, используемая модель имеет представления не только о реалистичных деталях и ярких текстурах на изображении. Так, например, она также может сохранять информацию о цветах, используемых на изображениях из реального мира, которые могут быть встроены в решаемую задачу совместного восстановления изображения и его раскраски. Например, в задачах восстановления старинных фотографий, которые зачастую бывают либо чёрно-белые, либо имеют желтоватый оттенок выцветания, вследствие деградаций накладываемых временем. В случае использования уже готового предобученного генератора, можно надеяться, что на этапе восстановления такого рода старинного изображения, модель, сможет на основе имеющихся у неё представлениях о цветах, попутно с основной задачей, раскрасить изображение в правдоподобные цвета. Также есть надежда, что используемая модель имеет представления о геометрии лица, её трёхмерном представлении, для обеспечения большего правдоподобия с образцами из реального мира.

#### **2.4. Модуль пространственного преобразования**

Для лучшего сохранения точности модели на этапе генерации изображения лица, используется модуль пространственного преобразования, который использует извлечённые U-Net, пространственные признаки изображения. Модуль модулирует сигнал, полученный в ходе генерации генеративных признаков  $F_{StyleGAN}$  из модели StyleGAN. Модуль необходим в первую очередь для того, чтобы на основе пространственной информации на поданном на вход изображении, можно было бы сохранить локальные

характеристики лица, для повышения точности процесса увеличения разрешения изображения, так как разные зоны на лице требуют различного подхода к увеличению разрешения. Это достигается при помощи генерации параметров для аффинного преобразования, для пространственной модуляции признаков, извлечённых генератором, данный подход показал свою эффективность на примере таких работ как [31, 32] для восстановления изображений и [33] для генерации изображений.

Таким образом, для признаков, получаемых на каждом слое генератора StyleGAN, формируется пара параметров для аффинного преобразования  $(\alpha, \beta)$ , из входных пространственных признаков  $F_{spatial}$ , несколькими свёрточными слоями. После чего модуляция производится при помощи масштабирования и сдвига признаков, получаемых сетью генератора:

$$\alpha, \beta = Conv(F_{spatial})$$
$$F_{output} = SFT(F_{GAN} | \alpha, \beta) = \alpha * F_{GAN} + \beta$$

Для получения лучшего баланса точности и реалистичности, преобразование пространственных признаков, производится не для всего сигнала, получаемого генератором, а для части каналов извлекаемых признаков, оставшиеся каналы в сигнале продолжают идти на вход следующего слоя генератора без изменений. Схема представлена на рисунке 5.

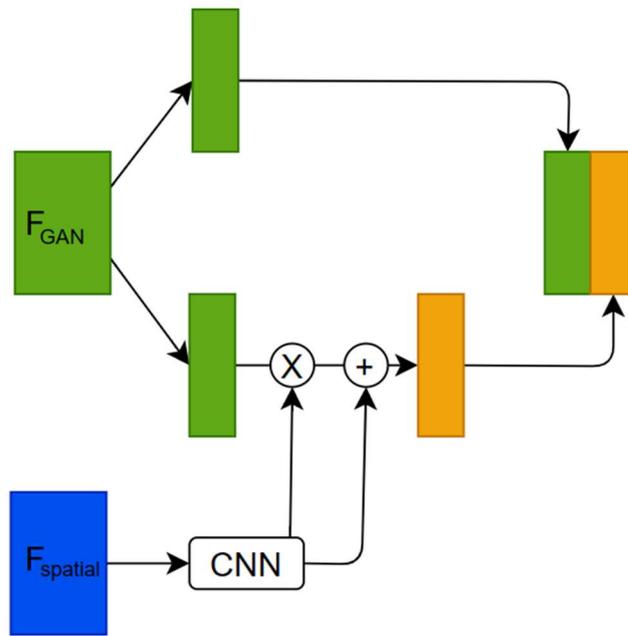


Рисунок 5 – Схема пространственного преобразования признаков

$$F_{output} = CSSFT(F_{GAN} | \alpha, \beta) = Concat(F_{GAN}^{split}, SFT(F_{GAN}^{split} | \alpha, \beta))$$

Где  $F_{GAN}^{split}$ , это часть сигнала  $F_{GAN}$ , поделённого в канальной размерности, а функция  $Concat$ , простая конкатенация двух тензоров признаков.

В результате данный модуль позволяет как использовать извлечённые из генератора StyleGAN, признаки характерные для лиц, так и добиться точности путём модуляции этого сигнала, таким образом достигая баланса между правдоподобностью сгенерированного изображения, так и точностью для процесса увеличения разрешения. Также данный модуль позволяет сэкономить вычислительные ресурсы, так как для его работы требуется меньшее количество каналов, также, как и в модели GhostNet [23].

Модуляция при помощи модуля пространственного преобразования признаков, проводится для каждого выхода в переработанной модели StyleGAN, ровно, как и вычисление параметров для сдвига и масштабирования. Получая таким образом готовое восстановленное изображение  $\hat{y}$ .

## 2.5. Функционал качества модели

Как уже было сказано функционал качества модели основан из нескольких компонент: 1) состязательная ошибка, вычисляемая дискриминатором модели, для получения реалистичных изображений, 2) локальная состязательная ошибка, вычисляемая для лицевых компонент, а также 3) ошибка восприятия.

Состязательная ошибка используется для того, чтобы, модель создавала изображения правдоподобные к изображениям из реального мира, так как дискриминатор, также, как и любой другой модели обучения, основанной на состязании генератора и дискриминатора, учится отличать реальные изображения от сгенерированных моделью обучения. Также, как и для любого другого дискриминатора ошибка определяется как:

$$L_{adv} = -\lambda_{adv} E_{\hat{y}} \text{softmax}(D(\hat{y}))$$

Где  $D$  – это ответ дискриминатора на вопрос сгенерировано ли изображение моделью, а  $\lambda_{adv}$  обозначает вес данной ошибки в общей ошибке модели.

Ошибка лицевых компонент, используется для того, чтобы усилить важные для лица признаки, а именно глаза и рот. Ошибка основана на локальных дискриминаторах, вычисляющих свой ответ для лицевых компонент на лице: левый, правый глаз и рот. Для вычисления ошибки, вначале изображения при помощи готового фреймворка [34] выделяет лицевые компоненты, после чего, обрезает изображение в три разных изображения, представляющих разные лицевые компоненты. Далее для каждого региона обучаются три различных дискриминатора, для того чтобы различать, реальны ли предоставленные изображения лицевых компонент или получены при помощи модели.

$$L_{local} = -\lambda_{local} \sum_{Region} E_{\hat{y}_{Region}} (1 - \text{softmax}(D(\hat{y}_{Region})))$$

Где  $\hat{y}_{Region}$  – это участок изображения: глаза, рот.  $D$ , обозначает ответ дискриминатора применительно к данному региону, а  $\lambda_{local}$  обозначает вес данной компоненты ошибки.

Ошибка восприятия, используется ещё с прошлых работ по увеличению разрешения изображений, и также используется для оценки точности модели применительно к задаче увеличения разрешения. Для вычисления данной компоненты ошибки также используется предобученная сеть VGG-19, которая также имеет знания о статистиках большого количества изображений, для того чтобы различать одну картинку от другой. Ошибка вычисляется следующим образом: на вход модели VGG-19 подаются, сгенерированное изображение и реальное изображение высокого разрешения, после чего, изображения проходятся через модель VGG-19, у которой были убраны последние полносвязных слои, отвечающие за классификацию изображения, далее полученные после прохода изображений вектора, используются для вычисления евклидова расстояния между векторами. В конечном счёте модель, в целях минимизации ошибки, должна генерировать изображения близкие к идентичному, так как ошибка заставляет модель генерировать изображения, скрытое пространство которых в модели VGG, близко к скрытому пространству истинного изображения.

$$L_{id} = \lambda_{id} \|\eta(\hat{y}) - \eta(y)\|,$$

где  $\eta$  – это модель VGG-19,  $\lambda_{id}$  – это вес данной компоненты потерь в общей формуле потерь модели.

В итоге формула функционала качества имеет следующую форму:

$$L_{total} = L_{adv} + L_{local} + L_{id}.$$

### 3. Реализация модели

Модель была реализована на языке Python, с использованием библиотек PyTorch и BasicSR. PyTorch используется для создания модели, она предоставляет удобные интерфейсы для создания моделей глубокого обучения. Модуль BasicSR – это модуль, который упрощает создание моделей, направленных на увеличение разрешения изображений,

автоматизируя этапы создания конвейера обучения, загрузки данных в модель, сохранения модели, её валидации. Вместо этого, можно просто определить архитектуру нейронных сетей, и модель обучения.

Обучение проводилось на основе датасета FFHQ, которая содержит 70 тысяч высококачественных снимков. Все изображения во время этапа обучения были приведены к разрешению 512 на 512 пикселей. Как можно видеть, при формировании обучающей выборки сходное изображение, вначале аугментируется, то есть, изображения поворачиваются на случайный угол, дополняя таким образом исходную выборку увеличивая количество элементов в нём. Далее на изображение накладывается размытие, при помощи фильтра Гаусса, значения в матрице фильтра, формируются случайно стандартное отклонение значений лежит от минус  $-\pi$  до  $+\pi$ . Далее изображение при линейной интерполяции сжимается. После чего на сжатое изображение накладывается гауссовый шум. Далее на полученное изображение накладываются артефакты сжатия “jpeg”, после чего изображение при помощи линейной интерполяции возвращается к исходному размеру. Далее изображение переводится в тензор, после чего полученный на предыдущем этапе тензор нормализуется.

Прямой проход модели генератора, реализован в виде метода. На вход метод получает деградировавшее изображение, на выходе нейронной сети, получается изображение с восстановленным на нём лицом.

Прямой проход генератора выглядит следующим образом. Как и было описано в разделе общей архитектуры модели, в нейронной сети генератора используются три компоненты – это модель U-Net и предобученная сеть StyleGAN, выходные тензоры которой модулируются, масштабированием и сдвигом, полученными из компоненты пространственного преобразования признаков. Вначале модель, получая на входе изображение, проводит его через кодировщик сети U-Net, получая по итогу вектор скрытого пространства сети генератора, именно из данного вектора `style_code`, впоследствии генерируется итоговое изображение. При этом каждый раз

вычисляются значения векторов в связях-пропусках. Далее во время прохода вектора из скрытого пространства через декодировщик модели U-Net, модель берёт значения тензоров, получающихся в результате развёртки, данные тензоры затем подаются на вход модуля пространственного преобразования признаков, которые на выходе дают два тензора – значения масштаба и сдвига, используемые при модуляции генератора StyleGAN. Далее полученные значения подаются на вход модуля модифицированного StyleGAN, модификация заключается в том, что на этапе прямого прохода, помимо значений вектора скрытого пространства, модель также получает параметры для модуляции генерируемых StyleGAN выходных тензоров.

Метод повторяет те же самые шаги, которые используются сетью генератора StyleGAN. То есть она также переводит при помощи модуля `mlp` переводит скрытый вектор, сформированный сетью U-Net, в вектора скрытого пространства модели StyleGAN, после чего проводит над ними адаптивную нормализацию. Далее полученные тензоры отправляются на вход слоёв обратной свёртки, к результату которых добавляется нормальный шум. После чего идёт модуляция полученного результата при помощи параметров пространственного преобразования признаков. Полученный результат затем подаётся на следующие слои. Результат впоследствии переводится в модель RGB.

Для генеративно-сопоставительных сетей, используется следующая модель обучения:

1. Замораживаются веса дискриминатора, происходит прямой проход модели, генератором создаются новые изображения.
2. Дискриминатор при прямом проходе вычисляет свой вердикт, который используется в ошибке генератора, происходит обратный проход модели.
3. Веса дискриминатора размораживаются, веса генератора, наоборот, замораживаются.

4. Происходит прямой проход модели, на этом этапе дискриминатор получает изображения от генератора, обратный проход которого происходил на предыдущем шаге.

5. Вычисляется ошибка дискриминатора, при вычислении ошибки реальные изображения имеют метку 1, а сгенерированные метку 0, происходит обратный проход модели.

В реализуемой модели, обучение происходит таким же образом. Вначале вычисляются ошибки всех компонент входящих в ошибку модели. Вначале при помощи фреймворка `goi`, выделяются лицевые компоненты на сгенерированном и исходном изображениях, далее вычисляются ошибки восприятия между регионами из исходного и сгенерированного изображений, далее вычисляется компонента ошибки восприятия модели, при помощи как уже упоминалось предобученной сети `vgg19`. Далее вычисляется ошибка критика. Все полученные значения ошибок суммируются после чего на основе этой ошибки происходит обратный проход модели. Далее веса дискриминатора модели размораживаются, и происходит уже вычисление ошибки для него. Отдельно вычисляются ошибки для дискриминатора лицевых компонент, и ошибки критика. В сети для увеличения точности и стабильности, используется модель критика из сети `WGAN`. После чего происходит обратный проход для критика.

Тестирование не представляет из себя сложной задачи. Загружается модель, полученная на этапе обучения, далее на основе загруженной модели, создаётся объект `restorer`, который будет заниматься восстановлением лиц. Далее каждая фотография подаётся на вход метода `enhance`, который возвращает результат модели для данной фотографии.

#### 4. Обучение модели.

Обучение проводилось на основе датасета FFHQ, которая содержит 70 тысяч высококачественных снимков. Все изображения во время этапа обучения были приведены к разрешению 512 на 512 пикселей.

Для тестирования использовался тот же датасет, однако он был заранее поделён на тренировочную, и тестовую.

Также обучение изначально планировалось провести в Google Colab, но из-за временных ограничений по использованию графического процессора обучение производилось на ноутбуке:

- Процессор: Intel core i7;
- Оперативная память 16,0 ГБ;
- Графический процессор NVIDIA GeForce RTX 3060 Laptop 6 ГБ.

Следовательно, обучение проводилось с последующим сохранением состояний сетей дискриминатора и генератора.

Ниже представлены примеры изображений до прохода через загрузчика данных и после прохода через загрузчика:



Рисунок 6 – результаты деградации фотографии загрузчиком



Рисунок 7 – результаты деградации фотографии

Ниже представлено изображение процесса обучения модели.

```

2023-05-19 03:12:29,374 INFO: Model [GFPGANModel] is created.
2023-05-19 03:12:35,307 INFO: Start training from epoch: 0, iter: 0
[W reducer.cpp:1298] Warning: find_unused_parameters=True was specified in DDP constructor, but did not find any unused parameters in the forward pass. This flag results in an extra traversal of the autograd graph every iteration, which can adversely affect performance. If your model indeed never has any unused parameters in the forward pass, consider turning this flag off. Note that this warning may be a false positive if your model has flow control causing later iterations to have unused parameters. (function operator())
2023-05-19 03:14:12,049 INFO: [train.][epoch: 0, iter: 100, lr:(2.000e-03,)] [eta: 7 days, 11:09:17, time (data): 0.967 (0.067)] l_g_pix: 3.1978e-02 l_p_8: 3.0470e-01 l_p_16: 2.7244e-01 l_p_32: 2.6720e-01 l_p_64: 2.3555e-01 l_p_128: 2.9211e-01 l_p_256: 2.2756e-01 l_p_512: 4.8750e-01 l_g_percep: 1.2134e+01 l_g_style: 3.9146e+00 l_g_gan: 9.8581e-02 l_d: 9.7610e-01 real_score: 8.7023e-01 fake_score: -5.0321e-01
2023-05-19 03:15:31,086 INFO: [train.][epoch: 0, iter: 200, lr:(2.000e-03,)] [eta: 7 days, 9:22:18, time (data): 0.879 (0.034)] l_g_pix: 2.3207e-02 l_p_8: 2.5357e-01 l_p_16: 2.4622e-01 l_p_32: 2.1342e-01 l_p_64: 1.9646e-01 l_p_128: 2.0468e-01 l_p_256: 2.0440e-01 l_p_512: 3.4053e-01 l_g_percep: 1.1318e+01 l_g_style: 3.1614e+00 l_g_gan: 1.0178e-01 l_d: 1.3107e+00 real_score: -2.8747e-01 fake_score: -5.6035e-01
2023-05-19 03:16:50,472 INFO: [train.][epoch: 0, iter: 300, lr:(2.000e-03,)] [eta: 7 days, 9:00:59, time (data): 0.794 (0.002)] l_g_pix: 2.6969e-02 l_p_8: 2.9765e-01 l_p_16: 2.8853e-01 l_p_32: 2.4341e-01 l_p_64: 2.4223e-01 l_p_128: 2.3775e-01 l_p_256: 2.4337e-01 l_p_512: 2.6808e-01 l_g_percep: 1.2008e+01 l_g_style: 3.2118e+00 l_g_gan: 1.6170e-01 l_d: 1.2252e+00 real_score: -3.3611e-01 fake_score: -1.3799e+00
2023-05-19 03:18:10,018 INFO: [train.][epoch: 0, iter: 400, lr:(2.000e-03,)] [eta: 7 days, 8:54:57, time (data): 0.795 (0.002)] l_g_pix: 3.1153e-02 l_p_8: 2.8662e-01 l_p_16: 2.4548e-01 l_p_32: 2.4657e-01 l_p_64: 2.4198e-01 l_p_128: 2.4576e-01 l_p_256: 2.5757e-01 l_p_512: 2.6710e-01 l_g_percep: 1.1073e+01 l_g_style: 2.4307e+00 l_g_gan: 1.0429e-01 l_d: 7.2980e-01 real_score: 2.5124e+00 fake_score: -5.3978e-01 l_dr1: 3.3891e-01

```

Рисунок 8 – процесс обучения модели

300 эпоха [total\_loss=5.2e-5]



Рисунок 9 – Результат работы модели

2400 эпоха [total\_loss=2.89e-6]

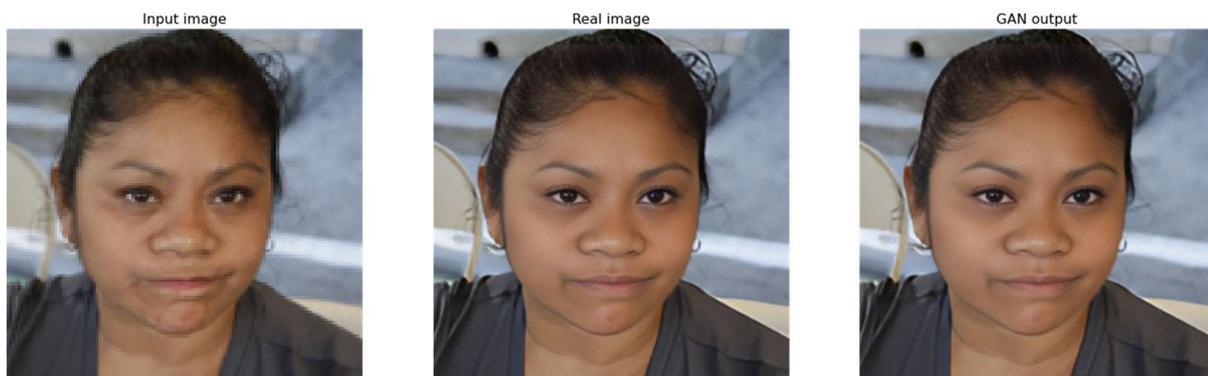


Рисунок 10 – Результат работы модели

2461 эпоха [total\_loss =4.33e-6]

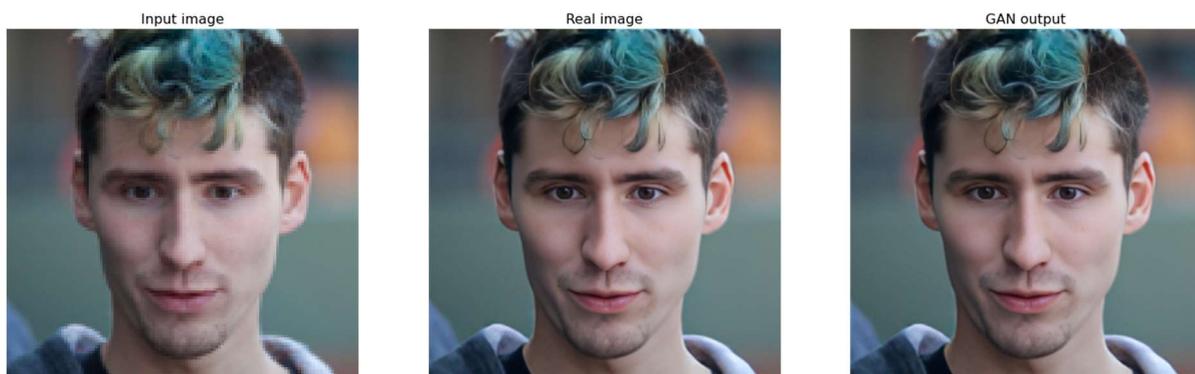


Рисунок 11 – Результат работы модели

2499 эпоха [total\_loss =4.05e-6]

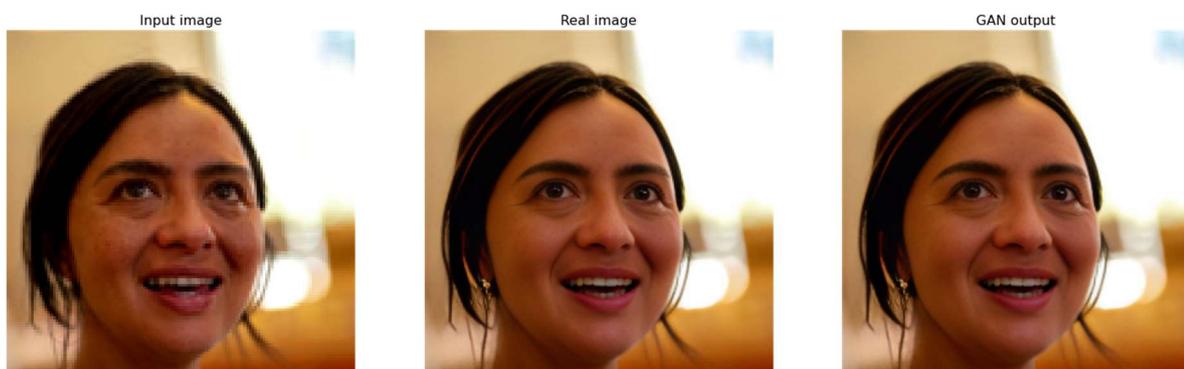


Рисунок 12 – Результат работы модели

По результатам сети можно сделать вывод, что генератор сумел научиться создавать реалистичные и точные изображения лиц.

На основе просмотра работ также решающих данную задачу по восстановлению изображений лиц [40, 44, 67], была составлена сравнительная характеристика модели:

Таблица 1 – сравнительная характеристика модели.

Метод	PSNR	SSIM
DeblurGANv2 [40]	25.91	0.6952
HiFaceGAN [67]	24.92	0.6195
DFDNet [44]	23.68	0.6622
GFP-GAN (реализованный)	25.08	0.6777

Как видно из сравнительной характеристики, предложенное решение имеет значение метрик сопоставимое со своими конкурентами.

## Заключение

Приведённый обзор методов получения суперразрешения изображений демонстрирует широкий спектр существующих задач и подходов. Каждый из методов обладает своими уникальными особенностями. Некоторые, как например [1] не требуют, по сути, никаких данных для обучения кроме самого деградировавшего изображения. Все рассмотренные модели обладают высокой производительностью, и хорошо себя показывают в рамках решения задачи увеличения разрешения, показывая отличные результаты при вычислении различных метрик оценки производительности моделей. Нам наиболее интересным показался подход показанный в [5], вследствие необычного подхода, связанного с остаточными блоками обработки изображения во множестве пространственных разрешений, способных обмениваться информацией друг с другом.

В работе была реализована собственная модель для восстановления лица на изображении при помощи генеративно-согласительной сети. Однако результаты показали, что, генеративно-согласительных сетей есть один существенный недостаток – их очень сложно обучать. Несмотря на это, полученное в результате решение обладает хорошей степенью надёжности и точности.

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА  
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И  
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

<b>Группа</b>	<b>ФИО</b>
8BM13	Куулар Чингис Тимурович

<b>Школа</b>	<b>ИШИТР</b>	<b>Отделение школы (НОЦ)</b>	<b>ОИТ</b>
Уровень образования	Магистр	Направление/специальность	Информатика и вычислительная техника

**Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:**

1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	Среднерыночные цены РФ для определения стоимости материальных ресурсов. Нормативные документы НИ ТПУ, ФЗ «О минимальном размере оплаты труда» для определения оплата труда исполнителей проекта.
2. <i>Нормы и нормативы расходования ресурсов</i>	Тариф электроэнергии 3,16 руб. кВт/ч., Премияльный коэффициент 0,3 Коэффициент доплат и надбавок 0,2 Районный коэффициент 1,3 Коэффициент дополнительной заработной платы 0,12 Накладные расходы 16%
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	Коэффициент отчислений на уплату во внебюджетные фонды 1,3

**Перечень вопросов, подлежащих исследованию, проектированию и разработке:**

1. <i>Оценка коммерческого и инновационного потенциала НТИ</i>	Провести предпроектный анализ
2. <i>Разработка устава научно-технического проекта</i>	Представить Устав научного проекта магистерской работы
3. <i>Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок</i>	Разработать план управления НТИ
4. <i>Определение ресурсной, финансовой, экономической эффективности</i>	Рассчитать сравнительную эффективность исследования

**Перечень графического материала (с точным указанием обязательных чертежей):**

1. Карта сегментирования рынка услуг по разработке системы восстановления лиц. 2. Оценочная карта для сравнения конкурентных технических решений. 3. Матрица SWOT-анализа 4. Оценка готовности проекта к коммерциализации 5. Цели и результат проекта 6. Иерархическая структура работ проекта 7. Перечень этапов, работ и распределение исполнителей 8. Временные показатели проведения НИ 9. Группировка затрат по статьям 10. Оценка характеристик вариантов исполнения 11. Сравнительная эффективность разработки.

<b>Дата выдачи задания для раздела по линейному графику</b>	
---	--

**Задание выдал консультант:**

<b>Должность</b>	<b>ФИО</b>	<b>Ученая степень, звание</b>	<b>Подпись</b>	<b>Дата</b>
доцент ОСГН, ШБИП	Былкова Татьяна Васильевна	канд.экон.наук		

**Задание принял к исполнению студент:**

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
8BM13	Куулар Чингис Тимурович		

## **5      ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ**

### **5.1.    Предпроектный анализ**

#### **5.1.1. Потенциальные потребители результатов исследования**

Целевым рынком разработки являются организации, работающие в сфере информационной безопасности. Данный выбор обусловлен тем, что разрабатываемый продукт позволяет повышать чёткость лиц на различных изображениях. Данный аспект программного продукта, позволяет организациям уменьшить временные и вычислительные затраты на обработку изображений, повысить точность детектирования, что в конечном итоге сократит себестоимость и повысит точность биометрических считывателей лиц.

Для анализа рынка результата исследования проведено сегментирование рынка. При этом было выбрано два наиболее значимых критерия (таблица 2): типы организаций и тип предоставляемого преимущества.

Можно выделить следующие типы предоставляемых компаниями аутсорсерами преимуществ организациям: увеличение точности детектирования, уменьшение времени считывания лица, уменьшение потребляемых сканером вычислительных ресурсов. По типу организаций, работающих в сфере безопасности, можно выделить следующие типы: разработчики мобильных систем, разработчики встраиваемых систем, разработчики серверных решений.

Таблица 2 – Карта сегментирования рынка услуг по разработке системы восстановления лиц.

		Тип предоставляемого преимущества		
		Увеличение точности	Уменьшение времени работы	Уменьшение ресурсов
Типы организаций в сфере безопасности	Мобильные разработчики			
	Разработчики встраиваемых систем			
	Разработчики серверных решений			

Haar cascades
  TrueDepth
  DFDNet
  HiFaceGAN

Haar cascades, TrueDepth, DFDNet, HiFaceGAN – это программные решения, которые схожи по функционалу, предоставляемому программному обеспечению считывателей биометрических данных, которые удовлетворяют потребностям указанных организаций, работающих в сфере безопасности.

В приведенной карте сегментирования видно, какие ниши на рынке по предоставлению услуг организациям в сфере информационной безопасности не заняты конкурентами или его уровень конкуренции низок. Следовательно, разрабатываемый продукт в виде библиотеки по восстановлению лиц для разработчиков серверных приложений наиболее актуален. Таким образом,

данная разработка будет привлекать потенциальных потребителей (независимых разработчиков).

### **5.1.2. Анализ конкурентных технических решений с позиции ресурсоэффективности и ресурсосбережения**

Для анализа конкурентных решений были взяты программные библиотеки со сходными характеристиками целевого рынка:

1. TrueDepth;
2. DFDNet;
3. HiFaceGAN,

Данные системы относятся к тем сегментам рынка, где разрабатываемое должно бороться за конкуренцию.

В качестве критериев для анализа конкурентных решений были взяты следующие критерии:

1. Гибкость настройки системы – наличие и степень возможности преднастройки параметров системы перед непосредственным восстановлением изображения
2. Сложность настройки системы – степень сложности и количество шагов, которые необходимо проделать пользователю, для настройки системы, перед возможностью использования системы
3. Возможность обучения на своей выборке – наличие и степень возможность системы обучаться на примерах, поданных пользователем системы, для её персонализации
4. Точность восстановления – качество изображения, полученного после прогона системой
5. Степень восстановления – оценка степени деградаций, накладываемых на изначальное изображение, которые система в принципе способна минимизировать
6. Количество потребляемых ресурсов – количество необходимых вычислительных ресурсов (такты частота процессора,

количество оперативной памяти), необходимых для работы системы

7. Время восстановления изображения – время, затрачиваемое системой, для восстановления лица на одном стандартном изображении.
8. Документация – качество и доступность документации системы, для комфортного использования потребителем
9. Размер готовой библиотеки – количество места на постоянном накопителе системы пользователя, которое будет занимать система восстановления лиц
10. Уровень проникновения на рынок – оценка возможных и/или существующих отзывов продукта после выхода на рынок
11. Сопровождение – возможность сопровождения системы после выхода на рынок

В ходе анализа были протестированы приложения и выставлены оценки, по критериям, представленным в таблице 3.

Таблица 3 – Оценочная карта для сравнения конкурентных технических решений (разработок)

Критерии оценок	Вес критерия	Баллы				Конкурентоспособность			
		Б <sub>ф</sub>	Б <sub>к1</sub>	Б <sub>к2</sub>	Б <sub>к3</sub>	К <sub>ф</sub>	К <sub>к1</sub>	К <sub>к2</sub>	К <sub>к3</sub>
1	2	3	4	5	6	7	8	9	10
<b>Технические критерии оценки ресурсоэффективности</b>									
Гибкость настройки системы	0,05	5	2	5	4	0.25	0.04	0.25	0.2
Сложность настройки системы	0,19	2	5	2	3	0.38	0.95	0.38	0.57
Возможность	0,03	5	1	5	5	0,15	0,03	0,15	0,15

обучения на своей выборке									
Точность восстановления	0,16	5	4	4	4	0,8	0,64	0,64	0,64
Степень восстановления	0,15	5	3	5	4	0,75	0,45	0,75	0,6
Количество потребляемых ресурсов	0,15	2	4	2	3	0,3	0,6	0,3	0,45
Время восстановления изображения	0,1	4	5	3	4	0,4	0,5	0,3	0,4
Документация	0,05	4	2	5	3	0,2	0,1	0,25	0,15
Размер готовой библиотеки	0,02	2	4	2	3	0,04	0,08	0,04	0,06
<b>Экономические критерии оценки эффективности</b>									
Уровень проникновения на рынок	0,05	4	5	4	4	0,2	0,25	0,2	0,2
Сопровождение	0,05	3	5	4	4	0,15	0,25	0,2	0,2
<b>Итого</b>	<b>1</b>					<b>3,62</b>	3,89	3,46	3,62

Из таблицы 3 можно сделать вывод, что разрабатываемая система может составить хорошую конкуренцию уже существующим решениям. Поскольку система, имеет большую точность восстановления лица на изображении, и может восстанавливать изображения, имеющие большее количество деградаций.

### 5.1.3. SWOT-анализ

Таблица 4 – SWOT-анализ проекта

Внешние факторы	Внутренние факторы		
		<p>Сильные стороны проекта:</p> <ol style="list-style-type: none"> <li>1. Высокая конкурентоспособность</li> <li>2. Низкая стоимость</li> <li>3. Гибкость настройки</li> <li>4. Высокая точность восстановления</li> </ol>	<p>Слабые стороны проекта:</p> <ol style="list-style-type: none"> <li>1. Высокие интеллектуальные затраты</li> <li>2. Меньшее количество начальной выборки</li> <li>3. Сложность предварительной настройки</li> </ol>
	<p>Возможности:</p> <ol style="list-style-type: none"> <li>1. Не наполненность ниши технологий восстановления лиц</li> <li>2. Большая часть технологий закрыта</li> </ol>	<p>Возможность создания приложения, обладающего лучшей по сравнению с конкурентами характеристикой, цена качество. Использование разработанного проекта так как другие технологии закрыты.</p>	<p>В связи с небольшим опытом разработки существует вероятность что не другие конкуренты могут оказаться на том же уровне что и разрабатываемый проект</p>
<p>Угрозы:</p> <ol style="list-style-type: none"> <li>1. Конкуренты совершенствуют свои уже отточенные</li> </ol>	<p>Конкуренты могут улучшать свои технологии, пока проект разрабатывается, однако в конечном итоге</p>	<p>Поскольку опыт разработки данного вида решений небольшой, то есть вероятность</p>	

	<p>технологии</p> <p>2. Нехватка времени на реализацию</p>	<p>открытость системы, а также низкая стоимость использования могут это компенсировать</p>	<p>невыполнения разработки в срок. Поскольку база данных небольшая возможна в конечном итоге меньшая репрезентативность разработки</p> <p>Так как на рынке уже существуют решения, сложный механизм настройки может оттолкнуть пользователей.</p>
--	--	--	---

Анализируя результаты SWOT-анализа, можно сделать вывод, что, несмотря на угрозы и слабые стороны проекта, проект можно считать перспективным и успешным.

#### 5.1.4. Оценка готовности проекта к коммерциализации

Была произведена оценка готовности проекта к коммерциализации, заполнения форма, содержащая показатели о степени проработанности научно-исследовательской работы с позиции коммерциализации и компетентности разработчика, приведены в таблице 5.

Таблица 5 – Оценка готовности проекта к коммерциализации

№ п/п	Наименование	Степень проработанности научно-исследовательской работы	Уровень имеющихся знаний у разработчика
1	Определен имеющийся научно-технический задел	3	4

2	Определены перспективные направления коммерциализации научно-технического задела	2	3
3	Определены отрасли и технологии (товары, услуги) для предложения на рынке	3	3
4	Определена товарная форма научно-технического задела для представления на рынке	2	2
5	Определены авторы и осуществлена охрана их прав	2	2
6	Проведена оценка стоимости интеллектуальной собственности	1	1
7	Проведены маркетинговые исследования рынков сбыта	1	2
8	Разработан бизнес-план коммерциализации научной разработки	1	1
9	Определены пути продвижения научной разработки на рынок	2	3
10	Разработана стратегия (форма) реализации научной разработки	1	2
11	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	1	1
12	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	1	1
13	Проработаны вопросы финансирования коммерциализации научной разработки	2	2
14	Имеется команда для коммерциализации научной	1	1

	разработки		
15	Проработан механизм реализации проекта	научного	2
	<b>ИТОГО БАЛЛОВ</b>		<b>25</b>
			<b>30</b>

В результате проведения оценки выявлено, что разработка имеет перспективность коммерциализации ниже среднего, и разработчик обладает средними знаниями для ее коммерциализации. Так как, отсутствует заказ на научно-исследовательскую работу, и разработка демоверсии представит возможность поиска заказчиков.

Для успешной коммерциализации проекта необходимо привлекать дополнительных разработчиков и специалистов в сфере права и финансов, а также заниматься поиском инвесторов проекта.

#### **5.1.5. Методы коммерциализации результатов научно-технического исследования**

В качестве метода коммерциализации была выбрана передача ноу=хау. Так как проект, предоставляет уже готовую функцию по восстановлению лица человека на изображении, при этом, имея преимущества в производительности, превосходя в этом своих конкурентов. Разработка, таким образом может быть использована любой организацией, без необходимости подстраиваться по неё. В таком случае разработчик системы чувствует, в процессе встраивания, системы, в бизнес-процессы организации.

### **5.2. Инициация проекта**

#### **5.2.1. Цели и результат проекта**

В рамках процессов инициации в таблице 6 представлены изначальные цели проекта и критерии их достижения. Научно-исследовательская работа иницирована студентом и научным руководителем.

Таблица 6 – Цели и результат проекта

Цели проекта:	Убрать или минимизировать деградации на изображениях, для облегчения процедуры их последующей обработки, или улучшения восприятия человеком.
Ожидаемые результаты проекта:	Программная библиотека, включающая функции и интерфейсы, которая может быть использована другими разработчиками.

### 5.3. Планирование управления научно-техническим проектом

#### 5.3.1. Иерархическая структура работ проекта

Для отображения содержания всего проекта была построена иерархическая структура работ, представленной на рисунке 13.

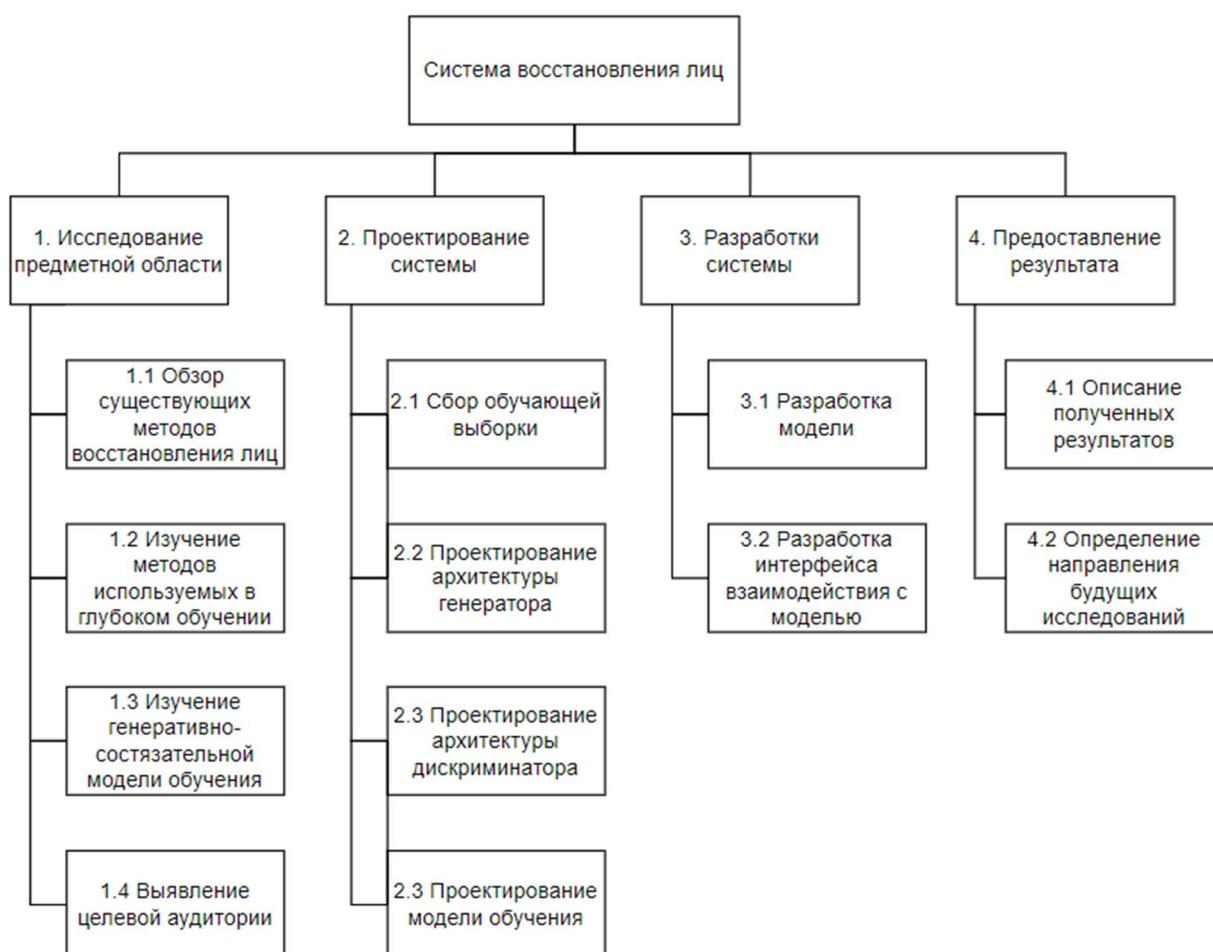


Рисунок 13 – Иерархическая структура работ проекта

По которой видно, что подготовительным этапом к созданию алгоритма является исследование предметной области. Далее формируются начальные требования к системе и осуществляется её проектирование в

зависимости от этих требований. После стадии проектирования следует стадия разработки предложенного решения. После стадии разработки наступает стадия описания полученного решения.

### 5.3.2. План проекта

В рамках планирования научно-исследовательского проекта был составлен календарный план проекта, таблица 7. Основные исполнители работ: руководитель (Р) и инженер (И).

Таблица 7 – Временные показатели проведённого исследования

Код работы	Название	Длительность, дни	Дата начала работ	Дата окончания работ	Состав участников
1.1	Обзор существующих методов восстановления лиц	7	31.01.2022	06.02.2023	Р, И
1.2	Изучение методов, используемых в глубоком обучении	6	07.02.2023	12.02.2023	И
1.3	Изучение генеративно состязательной модели обучения	3	13.02.2023	15.02.2023	И
1.4	Выявление целевой аудитории	2	16.02.2023	17.02.2023	И
2.1	Сбор обучающей выборки	3	18.02.2023	20.02.2023	Р, И
2.2	Проектирование архитектуры генератора	4	21.02.2023	24.02.2023	И
2.3	Проектирование архитектуры дискриминатора	4	25.02.2023	28.02.2023	И
2.4	Проектирование модели обучения	4	01.03.2023	04.03.2023	И
3.1	Разработка модели	50	05.03.2023	23.04.2023	И
3.2	Разработка интерфейса взаимодействия с моделью	20	24.04.2023	13.05.2023	И
4.1	Описание полученных результатов	14	14.05.2023	27.05.2023	Р, И
4.2	Определение направления	1	28.05.2023	28.05.2023	И

	будущих исследований				
<b>Итого</b>		<b>118</b>			

На основе таблицы 7 была построена диаграмма Ганта (рисунок 14), где отображены выполняемые работы в соответствии с датой их начала и окончания.

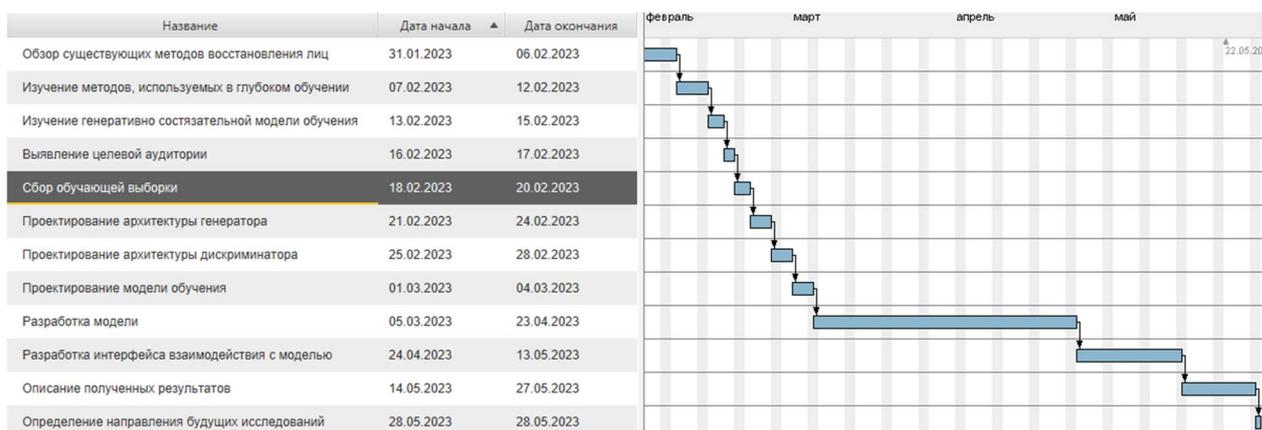


Рисунок 14 – Диаграмма Ганта

### 5.3.3. Бюджет научного исследования

Для определения бюджета исследования необходимо представить возможные альтернативы проведения научных исследований Морфологическая матрица представлена таблицей 8.

Таблица 8 – Морфологическая матрица

	1	2	3
А. Способ создания датасета для этапа обучения	Предобработка готового датасета отдельным исполнителем	Предобработка искусственно сформированного датасета	Использование готовых пар изображений
Б. Вариант запуска модели	В рамках одной вычислительной машины	Распределённо на кластере	
В. Вариант	При помощи	При помощи	

обмена сообщениями между исполнителями	интерфейса передачи сообщений	очереди сообщений	
Г. Способ передачи начальной конфигурации	Посредством файла конфигурации	Через СУБД	

Из данной морфологической матрицы можно выделить следующие варианты реализации проекта:

Вариант 1: А2Б2В1Г2

Вариант 2: А3Б1В1Г1

Вариант 3: А1Б2В2Г2

***Калькуляция плановой себестоимости НИ работы***

На основании рассчитанных затрат была составлена калькуляция плановой себестоимости НИ работы по форме, приведенной в таблице 8.

Таблица 8 – Группировка затрат по статьям для вариантов исполнения НИ

<b>Статья затрат Исполнения 1</b>	<b>Сумма, руб.</b>
Спецоборудование для научных работ	91000
Материальные затраты	3560
Основная заработная плата	77314
Отчислений на социальные нужды	20643
Накладные расходы	30802,7
<b>Итого:</b>	223319,7
<b>Статья затрат Исполнения 2</b>	
Спецоборудование для научных работ	91000
Материальные затраты	3560
Основная заработная плата	80488
Отчислений на социальные нужды	21596
Накладные расходы	31463

<b>Итого:</b>	228107
<b>Статья затрат Исполнения 3</b>	
Спецоборудование для научных работ	96400
Материальные затраты	3560
Основная заработная плата	82869
Отчислений на социальные нужды	22310
Накладные расходы	32822,2
<b>Итого:</b>	<b>237961,2</b>

***Специальные оборудования для научных (экспериментальных) работ***

В таблице 9 представлены все материальные затраты, необходимые для проведения исследования. Определение стоимости специальных оборудований берется по действующим прейскурантам.

Таблица 9 – Расчёт затрат по статье «Спецоборудование для научных работ»

№ п/п	Наименование оборудования	Кол-во единиц оборуд.	Цена единицы оборуд., тыс.руб.	Общая стоимость оборуд. руб.
1.	Ноутбук Acer Aspire 7	1	90	90000
2.	Периферийные устройства (мышь)	1	1	1000
<b>Итого:</b>				<b>91000</b>

***Материальные затраты***

Таблица 10 – Расчёт затрат по статье «Материальные затраты»

№ п/п	Наименование	Цена за единицы	Общая стоимость, руб.
1	Электроэнергия, руб./кВт*час	5,51	3560

### **Основная заработная плата**

Данная статья расходов включает основную заработную плату с учетом доплат для исполнителей проекта: инженера (дипломника), научного руководителя от ТПУ непосредственно участвующих в выполнении работ по данной теме.

Инженер работает по 5-дневной рабочей недели, руководитель работает по 6-дневной рабочей неделе. В таб. 9 приведен баланс рабочего времени для руководителя и инженера за 2023 год.

Таблица 11 – Баланс основной заработной платы

Показатели рабочего времени	Руководитель	Инженер
Календарное число дней	365	365
Количество нерабочих дней		
– Выходные дни	53	114
– Праздничные дни	14	14
Потери рабочего времени		
– Отпуск	56	24
– Невыходы по болезни	0	0
Действительный годовой фонд рабочего времени ( $F_d$ )	242	213

Базовый оклад ( $Z_6$ ) для руководителя (доцент со степенью кандидата наук) – составляет 37700 руб., инженера – 23800 руб.

Месячный должностной оклад ( $Z_m$ ) рассчитывается по формуле:

$$Z_m = Z_6 * K_p$$

где  $K_p$  – районный коэффициент (для Томской области 1,3)

Среднедневная заработная плата руководителя:

$$Z_{дн} = \frac{Z_m * M}{F_d}$$

где  $M$  – количество месяцев работы без отпуска в течение года:

при отпуске в 24 раб. дня  $M = 11,2$  месяца, 5-дневная неделя;

при отпуске в 48 раб. дней  $M = 10,4$  месяца, 6-дневная неделя;

$F_d$  – действительный годовой фонд рабочего времени, раб.дн.

Основная заработная плата ( $Z_{осн}$ ) рассчитывается по следующей формуле:

$$Z_{\text{осн}} = Z_{\text{дн}} * T_p$$

где  $T_p$  – продолжительность работ, выполняемых научно-техническим работником, раб.дн.

Результаты расчетов основной заработной платы исполнителей приведены в таблице 11. Продолжительность работы инженера составляет 78 рабочих дней от 118, с учетом 40 не рабочих дней (выходных/праздничные дни).

Таблица 12 – Расчет затрат по статье «Основная заработная плата»

Исполнители	$Z_b$ , руб.	$K_p$	$Z_m$ , руб.	$Z_{дн}$ , руб.	$T_p$ , раб.дн.	$Z_{осн}$ , руб.
Руководитель	39300	1,3	51090	2125	4	8501
Инженер	16242	1,3	21115	794	76	60311
<b>Итого «Исполнение 1»:</b>						68813
Руководитель	39300	1,3	51090	2125	4	8501
Инженер	16242	1,3	21115	794	80	63486
<b>Итого «Исполнение 2»:</b>						71987
Руководитель	39300	1,3	51090	2125	4	8501
Инженер	16242	1,3	21115	794	83	65866
<b>Итого «Исполнение 3»:</b>						74368

### ***Отчисления на социальные нужды***

Расчет величины отчислений во внебюджетные фонды приведена в таблице 12, и определяется по формуле:

$$C_{\text{внеб}} = K_{\text{внеб}} * (Z_{\text{осн}} + Z_{\text{доп}})$$

где  $K_{\text{внеб}}$  – коэффициент отчислений на уплату во внебюджетные фонды, равный 30% (налоговая нагрузка) по тарифу на 2023 год.

Таблица 12 – Расчёт затрат по статье «Отчислений на социальные нужды»

	Руководитель	Инженер (Вариант 1)	Инженер (Вариант 2)	Инженер (Вариант 3)
<b>Заработная плата исполнителя</b>	8501	68813	71987	74368

Коэффициент отчислений во внебюджетные фонды ( $K_{внеб}$ )	0,3	0,3	0,3	0,3
Отчислений во внебюджетные фонды ( $C_{внеб}$ )	2550	18093	19046	19760
<b>Итого:</b>	97957		102084	105179

### ***Накладные расходы***

Накладные расходы учитывают прочие затраты, не попавшие в предыдущие статьи расходов: оплата услуг связи, и т.д. Их величина определяется по следующей формуле:

$$Z_{накл} = K_{пр} * (\text{сумма всех статей})$$

где  $K_{пр}$  – коэффициент накладных расходов, можно взять в размере 16%.

Результат расчета приставлен в таблице 12

## **5.4. Определение ресурсной (ресурсосбережение), финансовой эффективности исследования**

### **5.4.1. Оценка сравнительной эффективности исследования**

Определение эффективности происходит на основе расчета интегрального показателя эффективности научного исследования. Его нахождение связано с определением двух средневзвешенных величин: финансовой эффективности и ресурсоэффективности. В качестве вариантов исполнения проекта рассматриваются три варианта рассмотренные выше, варианты отличаются методом исполнения разработки, варианты влияют на время реализации разработки инженером.

Интегральный показатель ресурсоэффективности ( $I_{pi}$ ) вариантов исполнения объекта исследования рассчитаны и приведены в таблице 13.

Таблица 13 – Сравнительная оценка характеристик вариантов исполнения проекта

Критерии \ ПО	Вес критерия	Исполнение 1	Исполнение 2	Исполнение 3
Гибкость настройки системы	0,05	5	5	4
Сложность настройки системы	0,19	4	3	3
Возможность обучения на своей выборке	0,03	5	5	5
Точность восстановления	0,16	5	5	5
Степень восстановления	0,15	5	5	5
Количество потребляемых ресурсов	0,15	3	4	3
Время восстановления изображения	0,1	4	5	4
Документация	0,05	5	5	5
Размер готовой библиотеки	0,02	5	4	4
Уровень проникновения на рынок	0,05	4	4	4
Сопровождение	0,05	5	5	5
<b>Итого</b>	<b>1</b>	<b>4,36</b>	<b>4,4</b>	<b>4,1</b>

Так как интегральный показатель ресурсоэффективности ( $I_{pi}$ ) вариант исполнения 2 является наибольшим, то в качестве максимального финансового значения ( $\Phi_{max}$ ) соотносится его значение.

Интегральный финансовый показатель разработки для Исполнения 1, 2 и 3:

$$I_{\text{фин.исп.1}} = \frac{223319,7}{237961,2} = 0,94$$

$$I_{\text{фин.исп.1}} = \frac{228107}{237961,2} = 0,96$$

$$I_{\text{фин.исп.1}} = \frac{237961,2}{237961,2} = 1$$

Полученная величина интегрального финансового показателя разработки отражает соответствующее численное уменьшения бюджета затрат разработки.

Интегральный показатель эффективности разработки ( $I_{финр}$ ) определяется на основании соотношения интегрального показателя ресурсоэффективности ( $I_m$ ) к интегральному финансовому показателю ( $I_ф$ ).

Сравнение интегрального показателя эффективности ( $Э_{ср}$ ) Исполнения 1, Исполнения 2 и Исполнения 3 позволит определить сравнительную эффективность проекта. Для этого нужно произвести соотношение  $I_{фин.р}/I_{фин.а}$ .

Таблица 14 – Сравнительная эффективность разработки

№	Показатели	Исполнение 1	Исполнение 2	Исполнение 3
1	Интегральный показатель ресурсоэффективности разработки	4,36	4,4	4,1
2	Интегральный финансовый показатель разработки	0,94	0,96	1
3	Интегральный показатель эффективности	4,64	4,58	4,1
4	Сравнительная эффективность вариантов исполнения	1	1,01	1,13

Оценка сравнительной эффективности показала, что разработка в первом варианте исполнения эффективнее второго и третьего вариантов исполнения на 1 % и 13%. Поэтому текущая разработка в первом исполнении является более эффективным вариантом решения поставленной в магистерской диссертации технической задачи с позиции финансовой и ресурсной эффективности.

## 5.5. Вывод

Для определения эффективности были выполнены следующие задачи:

- определен целевой рынок проекта;
- построена карта сегментирования;
- выявлены ниши на рынке услуг, незанятые конкурентами или где уровень конкуренции низок;
- произведен SWOT-анализа, в ходе которого были рассмотрены слабые и сильные стороны разработки инструментария;
- выявлены возможные перспективы его создания и рассмотрены варианты минимизации влияния угроз, которые могут этому помешать.

Для повышения эффективности работы необходимо минимизировать ошибки, контролировать процесс создания и проверять работу информационной технологии в соответствии с поставленными требованиями.

Также была определена трудоемкость проекта, которая составила 147 дней для студента. На основании трудоемкости была построена диаграмма Ганта.

Была рассчитана сравнительная эффективность проекта с вариантом разработки номер 3, которая составила 1,13. Отсюда следует что вариант разработки номер 3 является предпочтительным в плане эффективности, но разрабатываемая технология не сильно уступает в эффективности

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА  
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

<b>Группа</b>	<b>ФИО</b>
8ВМ13	Куулар Чингис Тимурович

**Тема магистерской диссертации: «Нейросетевой алгоритм восстановления изображений лиц»**

<b>Школа</b>	<b>ИШИТР</b>	<b>Отделение</b>	<b>ОИТ</b>
<b>Уровень образования</b>	Магистратура	<b>Направление/специальность</b>	09.04.01 «Информатика и вычислительная техника»

**Исходные данные к разделу «Социальная ответственность»**

1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Работа посвящена разработке алгоритмов восстановления изображений лиц из низкокачественных фотографий, с низким разрешением, искаженных шумом, размытием, артефактами сжатия и тд.
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<p><b>1. Производственная безопасность:</b></p> <p>1.1. Анализ выявленных вредных факторов</p> <ul style="list-style-type: none"> <li>• Природа воздействия</li> <li>• Действие на организм человека</li> <li>• Нормы воздействия и нормативные документы (для вредных факторов)</li> <li>• СИЗ коллективные и индивидуальные</li> </ul> <p>1.2. Анализ выявленных опасных факторов</p> <ul style="list-style-type: none"> <li>• Термические источники опасности</li> <li>• Электробезопасность</li> <li>• Пожаробезопасности</li> </ul>	<p>1. Вредные факторы:</p> <p>1.1. Нарушения микроклимата, оптимальные и допустимые параметры;</p> <p>1.2. Шум, ПДУ, СКЗ, СИЗ;</p> <p>1.3. Повышенный уровень электромагнитного излучения, ПДУ, СКЗ, СИЗ;</p> <p>1.4. Недостаточная освещенность; Проведен расчет освещения рабочего места; представлен рисунок размещения светильников на потолке с размерами в системе СИ;</p> <p>2. Опасные факторы:</p> <p>2.1. Электроопасность; класс электроопасности помещения, безопасные номиналы I, U, R<sub>заземления</sub>, СКЗ, СИЗ;</p> <p>2.2. Пожароопасность, категория пожароопасности помещения, марки огнетушителей, их назначение и ограничение применения; Приведена схема эвакуации.</p>
<p><b>2. Экологическая безопасность:</b></p> <ul style="list-style-type: none"> <li>• Выбросы в окружающую среду</li> <li>• Решения по обеспечению экологической безопасности</li> </ul>	Наличие промышленных отходов (бумага-черновики, перегоревшие люминесцентные лампы, оргтехника) и способы их утилизации;
<p><b>3. Безопасность в чрезвычайных ситуациях:</b></p> <p>1. перечень возможных ЧС при разработке и эксплуатации проектируемого решения;</p> <p>2. разработка превентивных мер по предупреждению ЧС;</p> <p>3. разработка действий в результате возникшей ЧС и мер по ликвидации её последствий.</p>	<p>Рассмотрены 2 ситуации ЧС:</p> <p>1) природная – сильные морозы зимой, (аварии на электро-, тепло-коммуникациях, водоканале, транспорте);</p> <p>2) техногенная – несанкционированное проникновение посторонних на рабочее место (возможны проявления вандализма, диверсии, промышленного шпионажа), представлены мероприятия по обеспечению устойчивой работы производства в том и другом случае.</p>

<b>4.Перечень нормативно-технической документации.</b>	– ГОСТы, СанПиНы, СНиПы
--	-------------------------

Дата выдачи задания для раздела по линейному графику	
--	--

**Задание выдал консультант:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ТПУ	Федорчук Ю.М.	д.т.н.		28.05.23 г.

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ВМ13	Куулар Чингис Тимурович		28.05.23 г.

## **6. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ**

### **6.1. Введение**

Проект, разработанный в рамках исследовательской работы, направлен на увеличение пространственного разрешения изображения с использованием методов глубокого обучения, а именно увеличение разрешения изображения путем генерации изображения высокого разрешения на основе изображения низкого разрешения. Сферы применения разработки: от нуждобычных пользователей до специализированных задач. Взаимодействие пользователя с разработанной программой производится с помощью программных и аппаратных средств ПЭВМ, а также с помощью периферийных устройств, подключенных к ПЭВМ.

Данный раздел посвящен анализу вредных и опасных факторов производственной среды для операторов ПЭВМ, а также разработке программ по минимизации воздействия вредоносного и опасного влияния, снижению вредных воздействий на окружающую среду и защите в чрезвычайных ситуациях.

Выпускная квалификационная работа студента выполнялась в домашних условиях, по месту проживания.

### **6.2. Производственная безопасность**

#### **6.2.1. Вредные производственные факторы**

##### **6.2.1.1. Отклонение показателей микроклимата в помещении**

Проанализируем микроклимат в помещении, где находится рабочее место. Микроклимат производственных помещений определяют следующие параметры: температура, относительная влажность, скорость движения воздуха. Эти факторы влияют на организм человека, определяя его самочувствие.

Оптимальные параметры микроклимата на рабочих местах должны соответствовать величинам, приведенным в Таблица. и Таблица. .

Таблица. 15. Оптимальные нормы микроклимата

Период года	Температура воздуха, С°	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	19-23	40-60	0.1
Теплый	23-25		0.2

Таблица. 16. Допустимые нормы микроклимата

Период года	Температура воздуха, С°		Относительная влажность воздуха, %	Скорость движения воздуха, м/с
	Нижняя допустимая граница	Верхняя допустимая граница		
Холодный	15	24	20-80	<0.5
Теплый	22	28	20-80	<0.5

Общая площадь рабочего помещения составляет 21м<sup>2</sup>, объем составляет 74,5м<sup>3</sup>. По СанПиН 2.2.2/2.4.1340-03 санитарные нормы составляют 6,5 м<sup>2</sup> и 20 м<sup>3</sup> объема на одного человека. Исходя из приведенных выше данных, можно сказать, что количество рабочих мест соответствует размерам помещения по санитарным нормам.

После анализа габаритных размеров рассмотрим микроклимат в этой комнате. В качестве параметров микроклимата рассмотрим температуру, влажность воздуха, скорость ветра.

В помещении осуществляется естественная вентиляция посредством наличия легко открываемого оконного проема (форточки), а также дверного проема. По зоне действия такая вентиляция является общеобменной. Основной недостаток - приточный воздух поступает в помещение без предварительной очистки и нагревания. Согласно нормам, СанПиН 2.2.2/2.4.1340-03 объем воздуха необходимый на одного человека в помещении без дополнительной вентиляции должен быть более 40м<sup>3</sup>. В нашем случае объем воздуха на одного человека составляет 37,25м<sup>3</sup>, из этого следует, что дополнительная вентиляция требуется. Параметры

микроклимата поддерживаются в холодное время года за счет систем водяного отопления с нагревом воды до 100°С, а в теплое время года – за счет кондиционирования, с параметрами согласно. Нормируемые параметры микроклимата, ионного состава воздуха, содержания вредных веществ должны соответствовать требованиям.

#### **6.2.1.2. Превышение уровней шума**

Одним из наиболее распространенных в производстве вредных факторов является шум. Он создается вентиляционным и рабочим оборудованием, преобразователями напряжения, рабочими лампами дневного света, а также проникает снаружи. Шум вызывает головную боль, усталость, бессонницу или сонливость, ослабляет внимание, память ухудшается, реакция уменьшается.

Основным источником шума в комнате являются компьютерные охлаждающие вентиляторы. Уровень шума варьируется от 35 до 42 дБА. Согласно СанПиН 2.2.2 / 2.4.1340-03, при выполнении основных работ на ПЭВМ уровень шума на рабочем месте не должен превышать 80 дБА.

При значениях выше допустимого уровня необходимо предусмотреть средства индивидуальной защиты (СИЗ) и средства коллективной защиты (СКЗ) от шума.

Средства коллективной защиты:

1. устранение причин шума или существенное его ослабление в источнике образования;
2. изоляция источников шума от окружающей среды (применение глушителей, экранов, звукопоглощающих строительных материалов, например любой пористый материал – шамотный кирпич, микропористая резина, поролон и др.);
3. применение средств, снижающих шум и вибрацию на пути их распространения;

Средства индивидуальной защиты;

1. применение спецодежды и защитных средств органов слуха: наушники, беруши, антифоны.

### **6.2.1.3. Повышенный уровень электромагнитных излучений**

Источником электромагнитных излучений в нашем случае являются дисплеи ПЭВМ. Монитор компьютера включает в себя излучения рентгеновской, ультрафиолетовой и инфракрасной области, а также широкий диапазон электромагнитных волн других частот.

Согласно СанПиН 2.2.2/2.4.1340-03 напряженность электромагнитного поля по электрической составляющей на расстоянии 50 см вокруг ВДТ не должна превышать 25В/м в диапазоне от 5Гц до 2кГц, 2,5В/м в диапазоне от 2 до 400кГц [1]. Плотность магнитного потока не должна превышать в диапазоне от 5 Гц до 2 кГц 250нТл, и 25нТл в диапазоне от 2 до 400кГц. Поверхностный электростатический потенциал не должен превышать 500В [1].

В ходе работы использовалась ПЭВМ типа ASUSGL703VD со следующими характеристиками: напряженность электромагнитного поля 3В/м; поверхностный потенциал составляет 560 В (основы противопожарной защиты предприятий ГОСТ 12.1.004 и ГОСТ 12.1.010 – 76.)[42].

При длительном постоянном воздействии электромагнитного поля (ЭМП) радиочастотного диапазона при работе на ПЭВМ у человеческого организма сердечно-сосудистые, респираторные и нервные расстройства, головные боли, усталость, ухудшение состояния здоровья, гипотония, изменения сердечной мышцы проводимости. Тепловой эффект ЭМП характеризуется увеличением температуры тела, локальным селективным нагревом тканей, органов, клеток за счет перехода ЭМП на теплую энергию.

Предельно допустимые уровни (ПДУ) облучения (по ОСТ 54 30013-83):

- а) до 10 мкВт./см<sup>2</sup>, время работы (8 часов);
- б) от 10 до 100 мкВт/см<sup>2</sup>, время работы не более 2 часов;
- в) от 100 до 1000 мкВт/см<sup>2</sup>, время работы не более 20 мин. при условии

пользования защитными очками;

г) для населения в целом ППМ не должен превышать 1 мкВт/см<sup>2</sup>.

Защита человека от опасного воздействия электромагнитного излучения осуществляется следующими способами:

Средства коллективной защиты (СКЗ):

1. защита временем;
2. защита расстоянием;
3. снижение интенсивности излучения непосредственно в самом источнике излучения;
4. заземление экрана вокруг источника;
5. защита рабочего места от излучения.

Средства индивидуальной защиты (СИЗ):

1. Очки и специальная одежда, выполненная из металлизированной ткани (кольчуга). При этом следует отметить, что использование СИЗ возможно при кратковременных работах и является мерой аварийного характера. Ежедневная защита обслуживающего персонала должна обеспечиваться другими средствами.

2. Вместо обычных стекол используют стекла, покрытые тонким слоем золота или диоксида олова (SnO<sub>2</sub>).

#### **6.2.1.4. Недостаточная освещенность рабочей зоны**

Для обеспечения требуемой освещенности необходимо использовать совмещенное освещение, создаваемое сочетанием естественного и искусственного освещения. При данном этапе развития осветительной техники целесообразно использовать люминесцентные лампы, которые по сравнению с лампами накаливания имеют большую светоотдачу на ватт потребляемой мощности и более естественный спектр.

Минимальный уровень средней освещенности на рабочих местах с постоянным пребыванием людей должен быть не менее 200 лк.

В расчётном задании должны быть решены следующие вопросы:

- выбор системы освещения;

- выбор источников света;
- выбор светильников и их размещение;
- выбор нормируемой освещённости;
- расчёт освещения методом светового потока.

В данном расчётном задании для всех помещений рассчитывается общее равномерное освещение.

Таблица 17. Параметры помещения.

Параметр	Обозначение	Значение, м
Длина	A	7
Ширина	B	3
Высота помещения	H	3,5
Свес	h <sub>с</sub>	0,4
Высота Р.П.	h <sub>рп</sub>	0,8
Высота от светильника до Р.П.	h	H- h <sub>рп</sub> - h <sub>с</sub>
Коэффициент отражения стен	ρ <sub>ст</sub>	70 %
Коэффициент отражения потолка	ρ <sub>п</sub>	70
Коэффициент запаса	K <sub>з</sub>	1.5
Коэффициент неравномерности	Z	1.1

Расчёт общего равномерного искусственного освещения горизонтальной рабочей поверхности выполняется методом коэффициента светового потока, учитывающим световой поток, отражённый от потолка и стен.

Световой поток лампы определяется по формуле:

$$\Phi_{\text{рас}} = \frac{E_{\text{н}} * S * K_{\text{з}} * Z}{N * \eta}$$

Где  $E_{\text{н}}$  – нормируемая минимальная освещённость по СНиП 23-05-95, лк;  $S$  – площадь освещаемого помещения, м<sup>2</sup>;  $K_{\text{з}}$  – коэффициент запаса, учитывающий загрязнение светильника (источника света, светотехнической арматуры, стени пр., т.е. отражающих поверхностей), наличие ватмо

сфере потолка, пыли;  $Z$  – коэффициент неравномерности освещения, отношение  $E_{ср}/E_{min}$ . Для люминесцентных ламп при расчётах берётся равным 1,1;  $N$  – число ламп в помещении;  $\eta$  – коэффициент использования светового потока.

Коэффициент использования светового потока показывает, какая часть светового потока лампы попадает на рабочую поверхность. Он зависит от индекса помещения  $i$ , типа светильника, высоты светильников над рабочей поверхностью  $h$  и коэффициентов отражения стен  $\rho_c$  и потолка  $\rho_n$ .

Индекс помещения определяется по формуле:

$$i = S/h(A + B)$$

Проведем расчет индекса помещения:

Площадь помещения:

$$S = A * B = 7 * 3 = 21 \text{ м}^2$$

Индекс:

$$i = \frac{S}{h * (A + B)} = \frac{21}{(3,5 - 0,8 - 0,4) * (7 + 3)} = 0,91$$

Согласно этим данным, коэффициент использования светового потока будет равен 43 % или в долях = 0,43.

Коэффициенты отражения оцениваются субъективно (таблица 16) [47].

Согласно указанной методике, выбираем тип источника света.

Наиболее подходящим вариантом является 40 ваттная лампа ЛБ, у которой  $\Phi = 2800$  лм. Для выбранного типа лампы подходит светильник ОД-2-40 с размерами: длина = 1230 мм, ширина = 266 мм.

Количество ламп для помещения:

$$N = \frac{E_H * S * K_3 * Z}{\Phi * \eta} = \frac{200 * 21 * 1,5 * 1,1}{2800 * 0,43} = 5,2$$

Принимаем  $N = 6$  лампы или 3 светильника.

Размещаем светильники в 1 ряд в ряду с соблюдением условий:  $L$  – расстояние между соседними светильниками или рядами (если по длине ( $A$ ) и ширине ( $B$ ) помещения расстояния различны,

то они обозначаются  $L_A$  и  $L_B$ ),

$l$ –

расстояние от крайних светильников до рядов стен.

Оптимальное расстояние от крайнего ряда светильников до стены рекомендуется принимать равным  $L/3$ .

Сначала определим световой поток расчетный.

$$\Phi = \frac{E_H * S * K_3 * Z}{N * \eta} = \frac{200 * 21 * 1,5 * 1,1}{6 * 0,43} = 2686 \text{ лм}$$

$$\Phi = \frac{E_H * S * K_3 * Z}{N * \eta} = \frac{200 * 42 * 1,5 * 1,1}{12 * 0,48} = 2406 \text{ лм};$$

Теперь определим расстояния между светильниками по длине помещения.

$$7000 = 2 * L_A + 3 * 1230 + \frac{2}{3} * L_A;$$

$$L_A = (7000 - 3690) * \frac{3}{8} = 1241 \text{ мм};$$

$$\frac{L_A}{3} = 414 \text{ мм};$$

Так как у нас всего один ряд светильников, то установим их по середине ширины.

$$3000 = 2l + 266;$$

$$l = (3000 - 266) * \frac{1}{2} = 1367 \text{ мм};$$

Рисуем схему размещения светильников на потолке для обеспечения общего равномерного освещения.

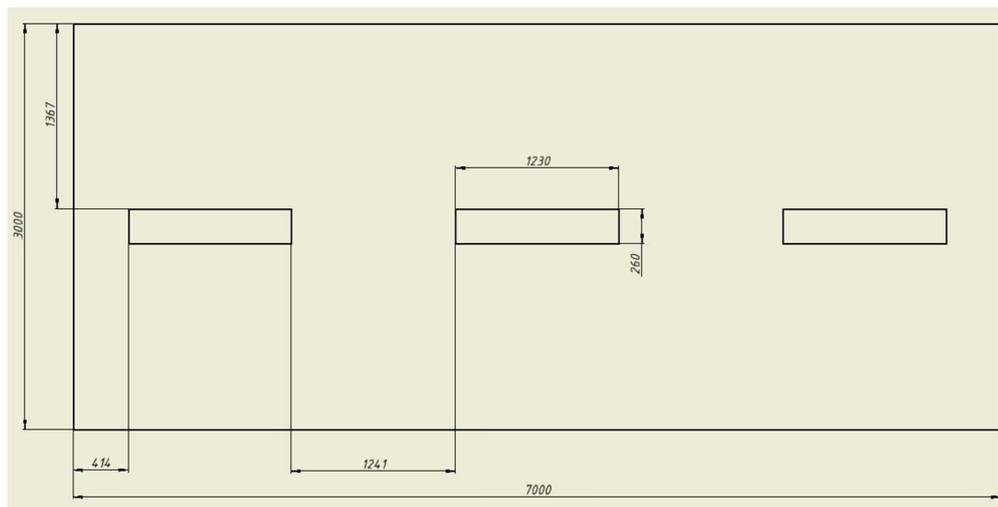


Рисунок 15 – План размещения светильников на потолке.

Проведем проверку выполнения условия соответствия:

$$-10\% \leq \frac{\Phi_{\text{л.станд}} - \Phi_{\text{л.расч}}}{\Phi_{\text{л.станд}}} * 100\% \leq 20\%$$

$$-10\% \leq \frac{2800 - 2686}{2800} * 100\% \leq 20\%$$

$$-10\% \leq 14\% \leq 20\%$$

Результат расчета укладывается в поле допуска.

Определим мощность осветительной установки:

$$P = N * P_i = 6 * 40 = 240\text{ВТ}$$

## 6.2.2. Опасные производственные факторы

### 6.2.2.1. Поражение электрическим током

К опасным факторам можно отнести наличие в помещении большого количества аппаратуры, использующей однофазный электрический ток напряжением 220 В и частотой 50Гц. По опасности электропоражения комната относится к помещениям без повышенной опасности, так как отсутствует повышенная влажность, высокая температура, токопроводящая пыль и возможность одновременного сприкосновения токоведущих элементов с заземленными металлическими корпусами оборудования [43].

Лаборатория относится к помещению без повышенной опасности поражения электрическим током. Безопасными номиналами являются:  $I < 0,1$  А;  $U < (6-42)$  В;  $R_{\text{зазем}} < 4$  Ом.

Для защиты от поражения электрическим током используют СКЗ и СИЗ.

Средства коллективной защиты:

1. — защитное заземление, зануление;
2. — малое напряжение;
3. — электрическое разделение сетей;
4. — защитное отключение;

5. — изоляция токоведущих частей;

6. — оградительные устройства.

Использование щитов, барьеров, клеток, ширм, а также заземляющих и шунтирующих штанг, специальных знаков и плакатов.

Средства индивидуальной защиты:

1. — диэлектрические перчатки, изолирующие клещи и штанги;

2. — слесарные инструменты с изолированными рукоятками;

3. — указатели величины напряжения, калоши, боты, подставки и коврики.

#### **6.2.2.2. Пожароопасность**

По взрывопожарной и пожарной опасности помещения подразделяются на категории А, Б, В1-В4, Г и Д.

Согласно НПБ 105-03 учебная комната относится к категории В – горючие и трудно горючие жидкости, твердые горючие и трудно горючие вещества и материалы, вещества и материалы, способные при взаимодействии с водой, кислородом воздуха или друг с другом только гореть, при условии, что помещения, в которых находится, не относятся к категории наиболее опасных А или Б.

По степени огнестойкости данное помещение относится к 1-й степени огнестойкости по СНиП 2.01.02-85 (выполнено из кирпича, которое относится к трудносгораемым материалам).

Возникновение пожара при работе с электронной аппаратурой может быть по причинам как электрического, так и неэлектрического характера.

Причины возникновения пожара неэлектрического характера:

а) халатное неосторожное обращение с огнем (курение, оставленные без присмотра нагревательные приборы, использование открытого огня);

Причины возникновения пожара электрического характера: короткое замыкание, перегрузки по току, искрение и электрические дуги, статическое электричество и т. п.

Для локализации или ликвидации загорания на начальной стадии используются первичные средства пожаротушения. Первичные средства пожаротушения обычно применяют до прибытия пожарной команды.

Огнетушители водо-пенные (ОХВП-10) используют для тушения очагов пожара без наличия электроэнергии. Углекислотные (ОУ-2) и порошковые огнетушители предназначены для тушения электроустановок, находящихся под напряжением до 1000В. Для тушения токоведущих частей и электроустановок применяется переносной порошковый огнетушитель, например ОП-5.

В общественных зданиях и сооружениях на каждом этаже должно размещаться не менее двух переносных огнетушителей. Огнетушители следует располагать на видных местах вблизи от выходов из помещений на высоте не более 1,35 м. Размещение первичных средств пожаротушения в коридорах, переходах не должно препятствовать безопасной эвакуации людей.

Для предупреждения пожара и взрыва необходимо предусмотреть:

1. специальные изолированные помещения для хранения и разлива легковоспламеняющихся жидкостей (ЛВЖ), оборудованные приточно-вытяжной вентиляцией во взрывобезопасном исполнении - соответствии с ГОСТ 12.4.021-75 и СНиП 2.04.05-86;

2. специальные помещения (для хранения в таре пылеобразной канифоли), изолированные от нагревательных приборов и нагретых частей оборудования;

3. первичные средства пожаротушения на производственных участках (передвижные углекислые огнетушители ГОСТ 9230-77, пенные огнетушители ТУ 22-4720-80, ящики с песком, войлок, кошма или асбестовое полотно);

4. автоматические сигнализаторы (типа СВК-3 М 1) для сигнализации о присутствии в воздухе помещений предвзрывных концентраций горючих паров растворителей и их смесей.

Лаборатория полностью соответствует требованиям пожарной безопасности, а именно, наличие охранно-пожарной сигнализации, плана эвакуации, изображенного на рисунке 16, порошковых огнетушителей с поверенным клеймом, табличек с указанием направления к запасному (эвакуационному) выходу.



Рисунок 16 – План эвакуации. сплошные линии —основной выход, штриховые линии —запасный выход.

### 6.3. Экологическая безопасность

Во время выполнения выпускной квалификационной работы вынуждены использовать черновики (предварительная запись информации) на бумажном носителе. Записи несут в себе конфиденциальную, а иногда даже секретную информацию. Чтобы повторно использовать бумагу для записей необходимо бумагу с записями shredировать с помощью шредера, спрессовать для уменьшения объема, упаковать в герметичную упаковку и хранить на складе до накопления объема для 1 транспортной единицы, после чего отправить на утилизацию макулатуры в ближайший ее пункт приема.

В компьютерах огромное количество компонентов, которые содержат токсичные вещества и представляют угрозу, как для человека, так и для окружающей среды.

К таким веществам относятся:

- свинец (накапливается в организме, поражая почки, нервную систему);
- ртуть (поражает мозг и нервную систему);
- никель и цинк (могут вызывать дерматит);
- щелочи (прожигают слизистые оболочки и кожу);

Поэтому компьютер требует специальных комплексных методов утилизации.

Таким образом утилизацию компьютера можно провести следующим образом:

- отделить металлические детали от неметаллов;
- разделить углеродистые металлы от цветмета;
- пластмассовые изделия (крупногабаритные) измельчить для уменьшения объема;
- кофир-порошок упаковать в отдельную упаковку, точно также, как и все проклассифицированные и измельченные компоненты оргтехники, и после накопления на складе транспортных количеств отправить

предприятиям и фирмам, специализирующимся по переработке отдельных видов материалов.

Люминесцентные лампы утилизируют следующим образом. Не работающие лампы немедленно после удаления из светильника должны быть упакованы в картонную коробку, бумагу или тонкий мягкий картон, предохраняющий лампы от взаимного соприкосновения и случайного механического повреждения. После накопления ламп объемом в 1 транспортную единицу их сдают на переработку на соответствующее предприятие. Недопустимо выбрасывать отработанные энергосберегающие лампы вместе с обычным мусором, превращая его в ртутьсодержащие отходы, которые загрязняют ртутными парами

#### **6.4. Безопасность в чрезвычайных ситуациях**

Природная чрезвычайная ситуация – обстановка на определенной территории или акватории, сложившейся в результате возникновения источника природной чрезвычайной ситуации, который может повлечь или повлек за собой человеческие жертвы, ущерб здоровью людей и (или) окружающей природной среде, значительные материальные потери и нарушение условий жизнедеятельности людей.

Производство находится в городе Томске с континентально-циклоническим климатом. Природные явления (землетрясения, наводнения, засухи, ураганы и т. д.), в данном городе отсутствуют.

Возможными ЧС на объекте в данном случае, могут быть сильные морозы и диверсия (вандализм, хулиганство, шпионаж).

Для Сибири в зимнее время года характерны морозы. Достижение критически низких температур приводит к авариям систем тепло- и водоснабжения, сантехнических коммуникаций и электроснабжения, приостановке работы. В этом случае при подготовке к зиме следует предусмотреть

- а) газобаллонные калориферы (запасные обогреватели),
- б) дизель или бензоэлектродгенераторы;

в) запасы питьевой и технической воды на складе (не менее 30 л на 1 человека);

г) теплый транспорт для доставки работников на работу и с работы домой в случае отказа муниципального транспорта. Их количества и мощности должно хватать для того, чтобы работа на производстве не прекратилась.

В лаборатории кибернетического центра наиболее вероятно возникновение чрезвычайных ситуаций (ЧС) техногенного характера.

Для предупреждения вероятности осуществления вышесказанных диверсии предприятие необходимо оборудовать системой видеонаблюдения, круглосуточной охраной, пропускной системой, надежной системой связи, а также исключения распространения информации о системе охраны объекта, расположении помещений и оборудования в помещениях, системах охраны, сигнализаторах, их местах установки и количестве. Должностные лица раз в полгода проводят тренировки по отработке действий на случай экстренной эвакуации.

## ЗАКЛЮЧЕНИЕ

В работе был приведён обзор методов позволяющих увеличить разрешение лиц на изображениях, а также восстанавливать деградировавшие изображения. Все рассмотренные в работе методы используют модели глубокого обучения. Различаются подходы, которые они используют в модели обучения, так некоторые из них используют те детали, которые можно извлечь непосредственно из деградировавших изображений, другие же используют наработки других исследователей в области компьютерного зрения, в виде априорных знаний о структуре лиц.

В работе была реализована комплексная модель глубокого обучения, которая в своей основе представляет из себя генеративно состязательную сеть. В работе была подробно рассмотрена архитектура предложенной модели, а также детали её реализации. Кроме того, были рассмотрены технические детали структурных элементов реализованной модели. Также было дано теоретическое обоснование, того почему модель должна справляться со своей задачей. Результаты показали, что генеративно-состязательные сети имеют существенный недостаток – их очень сложно обучать. Поэтому для получения хороших результатов необходимо провести достаточно большее число эпох обучения. Изображения, полученные на выходе модели, получились, на субъективный взгляд, относительно неплохими. Для оценки качества изображений были рассчитаны метрики PSNR и SSIM, которые показали, что обученная модель конкурентоспособна со своими аналогами DeblurGANv2, HiFaceGAN, DFDNet.

Был проведен анализ финансовой и ресурсной эффективности рассматриваемого проекта, который показал, что проект конкурентоспособен и эффективен в коммерциализации.

Выявлены возможные негативные факторы, которые могут возникнуть во время производства, и пути их решения.

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Jianchao Yang, Thomas Huang. Image super-resolution: Historical overview and future challenges. siteseerx. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.309.1145&rep=rep1&type=pdf> (дата обращения 21.12.2021)
2. Жан М. Рабаи, Ананта Чандракасан, Боривож Николич. Цифровые интегральные схемы. Методология проектирования = Digital Integrated Circuits. — 2-е изд. — М.: «Вильямс», 2007. — С. 912. — ISBN 0-13-090996-3
3. Интерполяция и дискретизация, зачем они нужны при проективном преобразовании изображения? Habr. URL: <https://habr.com/ru/company/smartengines/blog/500576/> (дата обращения 9.01.2022)
4. Liang-Chieh Chen, George Papandreou. Rethinking Atrous Convolution for Semantic Image Segmentation. [Text] // google inc. 2017. – 14с.
5. Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, Ling Shao. Learning Enriched Features for Real Image Restoration and Enhancement. [Text] // Google Research. 2020. – 20с.
6. Zhihao Wang, Jian Chen, Steven C.H. Hoi. Deep Learning for Image Super-resolution: A Survey. [Text] // Cornell University archive. 2019. – 24с.
7. Younghyun Jo, Sejong Yang, Seon Joo Kim, Yonsei University. Investigating Loss Functions for Extreme Super-Resolution. [Text] // Computer Vision Foundation. 2020. – 8с.
8. Jaejun Yoo, Namhyuk Ahn, Kyung-Ah Sohn. Rethinking Data Augmentation for Image Super-resolution: A Comprehensive

- Analysis and a New Strategy. [Text] // Google Research survey. 2020. – 18c.
9. Justin Johnson, Alexandre Alahi, Li Fei-Fei. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. [Text] // Cornell University archive. 2016. – 18c.
  10. Leon A. Gatys, Alexander S. Ecker, Matthias Bethge. Image Style Transfer Using Convolutional Neural Networks. [Text] // Computer Vision Foundation archive. 2014. – 10c.
  11. Dmitry Ulyanov, Andrea Vedaldi, Victor Lempitsky. Deep Image Prior. [Text] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 9446-9454
  12. X. Jia F. Li and D. Fraser. Universal HMT based super resolution for remote sensing images. In IEEE International Conference on Image Processing, pages 333–336, 2008.
  13. J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
  14. R. Y. Tsai and T. S. Huang. Multipleframe image restoration and registration. In *Advances in Computer Vision and Image Processing*, pages 317–339. Greenwich, CT: JAI Press Inc., 1984.
  15. M. S. Alam, J. G. Bognar, R. C. Hardie, and B. J. Yasuda. Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames. *IEEE Transactions on Instrumentation and Measurement*, 49(5):915–923, 2000.
  16. S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1167–1183, 2002.

17. Sean Borman and Robert L. Stevenson. Super-resolution from image sequences A review. In Proceedings of the 1998 Midwest Symposium on Circuits and Systems, pages 374–378, 1998.
18. Burger, H.C., Schuler, C.J., Harmeling, S.: Image denoising: Can plain neural networks compete with bm3d? In: CVPR, pp. 2392–2399 (2012).
19. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: ECCV. pp. 694–711. Springer (2016).
20. VGG16 and VGG19. URL: <https://keras.io/api/applications/vgg/> (дата обращения 22.06.22)
21. Yu Chen, Ying Tai, Xiaoming Liu, Chunhua Shen, and Jian Yang. Fsrnet: End-to-end learning face super-resolution with facial priors. In CVPR, 2018.
22. Xin Yu, Basura Fernando, Bernard Ghanem, Fatih Porikli, and Richard Hartley. Face super-resolution guided by facial component heatmaps. In ECCV, pages 217–233, 2018.
23. Chaofeng Chen, Xiaoming Li, Lingbo Yang, Xianhui Lin, Lei Zhang, and Kwan-Yee K. Wong. Progressive semantic-aware style transformation for blind face restoration. arXiv:2009.08709, 2020.
24. Xiaoming Li, Ming Liu, Yuting Ye, Wangmeng Zuo, Liang Lin, and Ruigang Yang. Learning warped guidance for blind face restoration. In ECCV, 2018.
25. Xiaoming Li, Wenyu Li, Dongwei Ren, Hongzhi Zhang, Meng Wang, and Wangmeng Zuo. Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion. In CVPR, 2020.

26. Berk Dogan, Shuhang Gu, and Radu Timofte. Exemplar guided face image super-resolution without facial landmarks. In CVPRW, 2019
27. Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In CVPR, 2020
28. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015
29. Jinjin Gu, Yujun Shen, and Bolei Zhou. Image processing using multi-code gan prior. In CVPR, 2020
30. Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. Exploiting deep generative prior for versatile image restoration and manipulation. In ECCV, 2020
31. Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In CVPR, 2018
32. Xiaoming Li, Chaofeng Chen, Shangchen Zhou, Xianhui Lin, Wangmeng Zuo, and Lei Zhang. Blind face restoration via deep multi-scale component dictionaries. In ECCV, 2020
33. Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. Semantic image synthesis with spatially-adaptive normalization. In CVPR, 2019
34. Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In ECCV, 2016.
35. Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In ICCV, 2017

36. Xiaoming Li, Chaofeng Chen, Shangchen Zhou, Xianhui Lin, Wangmeng Zuo, and Lei Zhang. Blind face restoration via deep multi-scale component dictionaries. In ECCV, 2020.
37. СанПиН 2.2.4.548-96. Гигиенические требования к микроклимату производственных помещений.
38. ГОСТ Р 12.1.019-2009. Электробезопасность. Общие требования и номенклатура видов защиты.
39. ГОСТ 12.4.123-83. Средства коллективной защиты от инфракрасных излучений. Общие технические требования.
40. [8] ГОСТ Р 12.1.019-2009. Электробезопасность. Общие требования и номенклатура видов защиты.
41. ГОСТ 12.1.030-81. Электробезопасность. Защитное заземление. Зануление.
42. ГОСТ 12.1.004-91. Пожарная безопасность. Общие требования.
43. ОСТ 12.2.037-78. Техника пожарная. Требования безопасности.
44. СанПиН 2.1.6.1032-01. Гигиенические требования к качеству атмосферного воздуха.
45. ГОСТ 30775-2001. Ресурсосбережение. Обращение с отходами. Классификация, идентификация и кодирование отходов.
46. СНиП 21-01-97. Противопожарные нормы.
47. ГОСТ 12.4.154. Система стандартов безопасности труда. Устройства, экранирующие для защиты от электрических полей промышленной частоты. Общие технические требования, основные параметры и размеры.

## Приложение А

(справочное)

### Section 1

#### Related methods overview

### Section 3

#### Model using architecture

Студент

Группа	ФИО	Подпись	Дата
8ВМ13	Куулар Чингис Тимурович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент, ОИТ	Иванова Юлия Александровна	к.т.н.		

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент, ШБИП	Сидоренко Татьяна Валерьевна	к.п.н.		

## 1. METHODS OVERVIEW

A large number of works are devoted to the problem of restoration of faces on images. Some works, such as [4, 5, 6], use characteristic features of faces, such as: special anchor points on faces [6, 4], the use of masks with facial components [5]. These works show that geometric facial features, are important for obtaining facial outlines and facial details. However, these features can be found only on the degraded images themselves, so in the case of faces on old or fuzzy photos, it will be very difficult to obtain these facial characteristics. Also, while they give a good idea of facial geometry, they cannot give information about texture, such as color, pupil texture.

Also, for example, the DFDNet model [10] uses a similar approach based on reference features. These features are references to images, or, more precisely, to parts of images similar to those fed to the input of the model. The DFDNet model, builds a dictionary of facial features, such as eyes, mouth, nose based on images from the training sample, using convolution layers. This feature dictionary is then used to help the image reconstruction process. However, this approach has a drawback, because although the model reconstructs well, these areas on the face, often areas outside the regions from the dictionary of facial features undergo visual degradation, such as hair, facial contour, chin, and so on suffer. Moreover, with the limited size of the dictionary, the images obtained after generation suffer from a lack of diversity.

Another approach uses the same methods as the problem of increasing the resolution of images. For example, the approaches [7, 8, 9] use generative inversion [7, 8], or variational autoencoders [9]. Generative inversion approaches are based on feeding the input image to the outputs of the generator, to obtain vectors in the hidden space of the generator, and then generate a new image. The idea is that vectors close to each other in the hidden space of the generator will most likely get the same image, but it will be of higher quality, which is provided by the adversarial error and perceptual error. The autoencoder approach uses a similar idea, based on encoding a low-quality image into a hidden space and then

obtaining a higher-quality image with the decoder. The idea is based on the fact that the model, pursuing the goal of minimizing the loss between the input and output images, will not try in its scales to remember artifacts characteristic of low-quality images. However, these approaches require low-quality to high-quality image pairs. However, it is not always possible to obtain high-quality images of faces. Moreover, these models are computationally complex due to the fact that they require a large number of passes to memorize the mapping of a low-quality image into a high-quality image.

## **2. OVEWIEW OF ARCHITECTURE USED BY MODEL**

### **2.1. Developing model brief overview**

The overall framework of a model is depicted below in figure 2. The model consists of several parts. Given an input facial image suffering from unknown, various degradation, which is generally unknown. The aim of training model is obtaining a high-quality image, which should be similar as close as possible to the ground-truth, in terms of realness and fidelity. The model consists of the following parts: U-Net network responsible for removing degradation from input and pre-trained generator network, in this case StyleGAN [18]. StyleGAN was chosen primarily because it showed that it could generate high-resolution face images, mostly similar natural manifold. The StyleGAN generator network is used to derive from its weights reliable features specific to face images. The same principle is used in generative inversion approaches: the latent space of the U-Net network is used to generate a face image through the StyleGAN generator model.

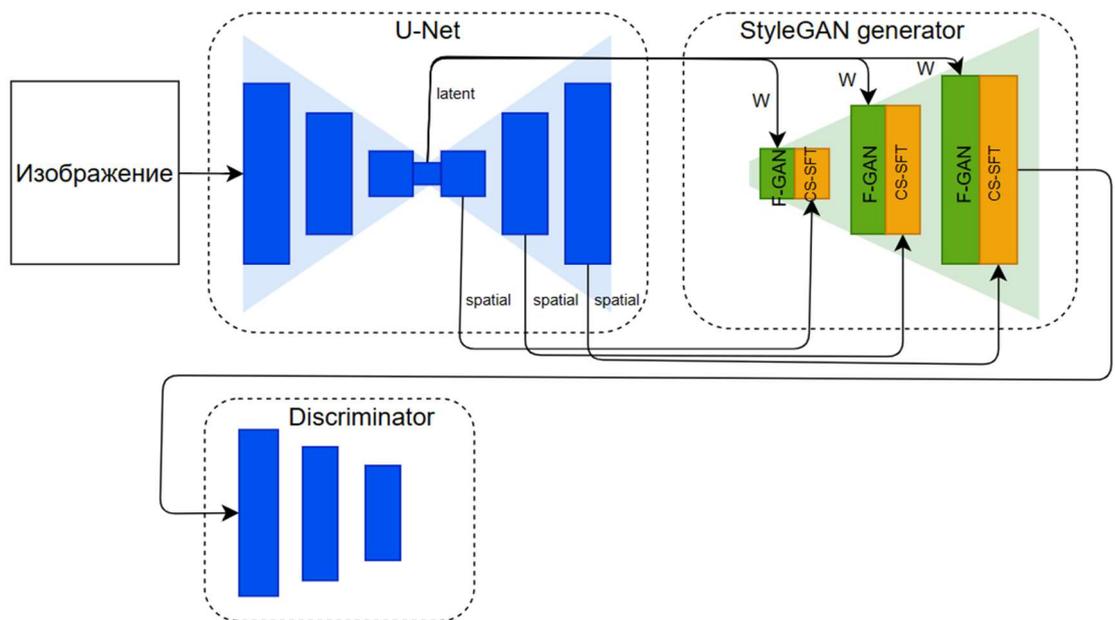


Figure 2 – brief model architecture overview

As you can see from the figure above, the two networks are interconnected. First of all, the vectors from the latent space of the U-Net model, like in classical StyleGAN approach, pass through the mapping feed-forward nets, after that they fed to all layers of StyleGAN model. Also, the vectors from the hidden spaces of the U-Net model from upsampling layers, are then fed to the spatial feature transformation blocks, which modulate the signal received from the layers of the generator of StyleGAN model. The U-Net is specifically used to derive two types of latent-space vectors: 1) a latent-space vector encoding the latent features of the image that the U-Net has extracted, and which is then mapped to the closest vector from the latent space of the Style-GAN generator; 2) spatial tensors obtained from the layers of the U-Net model when it unwraps a vector from the latent space, which are then used to modulate the signals received from the layers of the StyleGAN model generator.

Then due to StyleGAN model architecture, vector from the latent-space of the U-Net model is mapped to several  $W$  vectors, with several fully connected layers. These mapping procedures is affine transformation from the latent-space of U-Net model to latent-space of StyleGAN model, since latter in its architecture does not use vectors from the hidden space directly, instead, vectors from the hidden space proceed as regularizes of the adaptive instance normalization layers,

guaranteeing linear separability of features from the latent vector space. The layers of the StyleGAN network, upon receiving a vector close to the vector of the original image, generate intermediate convolutional features that provide a variety of facial details, as the pre-trained network, had to memorize these features in the weights of its layers. The spatial features are further used to modulate the features generated with StyleGAN using the spatial feature transform layers. Moreover, since the StyleGAN network allows linear separation of features from the latent-space, modulation will occur from coarse traits to finer ones.

The aim of the training model will be minimizing the overall error, which will consist of several components, namely: the adversarial error obtained by the discriminator network, which is responsible for the overall fidelity of the generated images, adversarial error obtained by local discriminators aimed at distinguishing the face components, that is the eyes and mouth, as well as used from the past work the error of perception loss.

Thus, during the learning phase will be trained weights of two models, the network U-Net, which is designed to extract features from the original image and map it into the space of hidden vectors, and the discriminator network, which is responsible for the overall fidelity of the training model.

## **2.2. Degradation removal module**

Figure 3 below shows a generalized scheme of the U-Net model [10]. As can be seen from the figure, the model first lowers the dimensionality of space, in order to extract the semantic features of the image, that is, answers the question of what is shown in the picture, while increasing the number of channels in the signal. After that come the layers that increase the dimensionality of the signal, and lowering the number of channels, reconstructing the image using the extracted semantic features.

Also, in order for the model not to lose spatial features in the image, there are skipping connections in the model. Since when decreasing the dimensionality of the image, the model does not remember the spatial features of the image, that is, it cannot answer the question of where it is depicted. Therefore, with the help of

residual connections, when increasing the dimensionality of the image, the model uses the previously extracted features, in the step of decreasing the dimensionality. These features allow us to mix the high-level abstract information extracted during the downsizing process with the spatial information from the previous layers of the network.

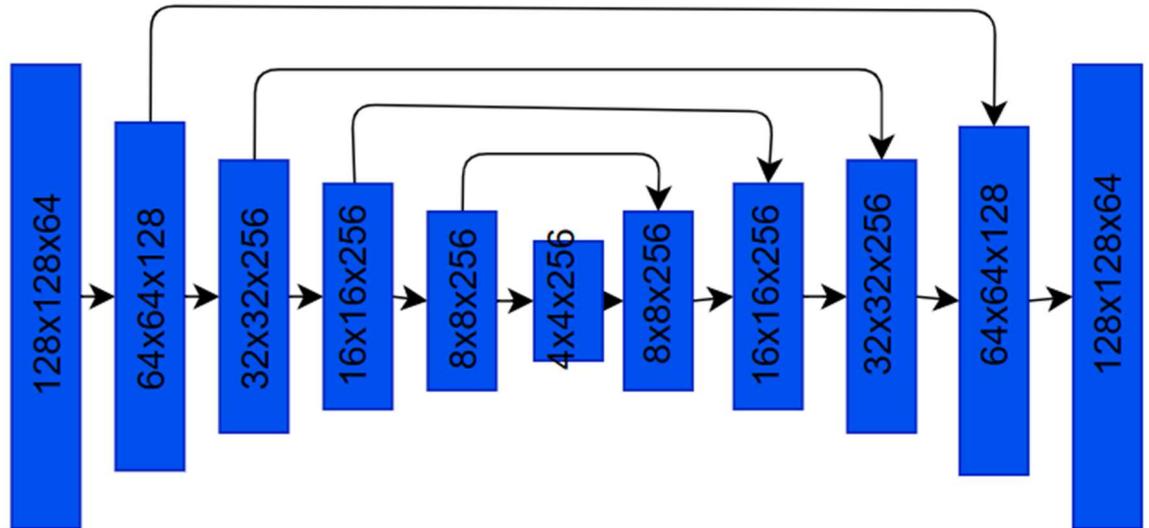


Figure 3 – overall scheme of U-Net model

The module based on the U-Net network is designed to explicitly remove degradation in the image, as previously mentioned, this model is focused on not remembering information about various kinds of degradation artifacts: blurring, compression and so on, concentrating on semantic and spatial features. This is achieved by extracting degradation-clean multidimensional feature vectors: hidden features obtained after the last layer of network convolution, as well as spatial features obtained each time the image resolution is increased.

$$F_{latent}, F_{spatioal} = UNet(x)$$

Hidden feature vectors are used to subsequently map them to the StyleGAN model vector space, and spatial features are used to modulate features obtained from the StyleGAN model.

### 1.3 Generative module

Figure 4 below shows a schematic view of the generator used in the StyleGAN model [9]. As can be seen from the diagram in the figure, the generator does not use vectors from the latent space directly to generate new face images,

instead, the generator uses vectors from the latent space in the adaptive normalization layer of the AdaIN instance, which essentially produces a scaling and shift of the input signal relative to the vector from the hidden space. Moreover, the vectors generated from the U-Net model are also not directly fed to the generator inputs, instead, they pass through feedforward network, which in turn is an affine transformation from one space to another. As a result, the vectors from the hidden space of the generator model are obtained.

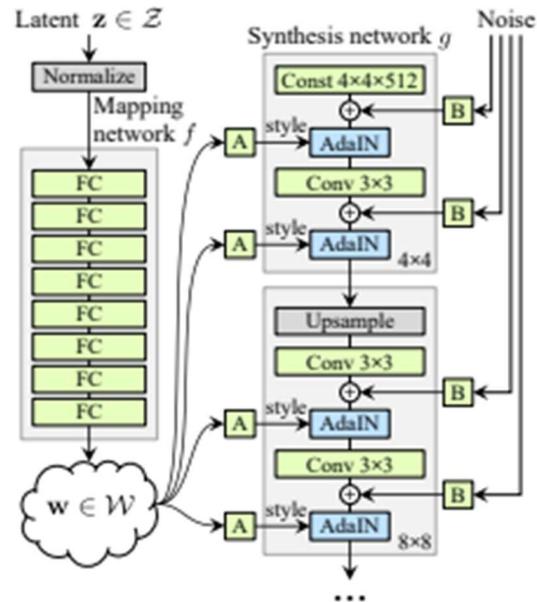


Figure 4 – overall scheme of StyleGAN model

The pre-trained generative network of the StyleGAN model was used, since according to [11, 12], it has been obtained that the generators used to obtain face images capture statistics over face features in their weights in convolutional layers. The process of generating an improved face image is as follows: vectors from the hidden space, obtained in the previous step by the U-Net model, are used to generate features from the StyleGAN model, but these features are intermediate. Thus, the generator is not used to obtain a high-precision image of the face directly. Instead, the intermediate features are modulated, using the spatial features also obtained by the U-Net model in the previous step, to ensure good accuracy with the input image.

Thus, the vectors obtained at the encoding stage by the U-Net model  $F_{latent}$ , using the pre-trained linear StyleGAN layers, are encoded into the hidden

space  $\mathbf{W}$ , to preserve the semantics of the image, during the generation stages in the StyleGAN generator. After that, vectors from space  $\mathbf{W}$ , pass through each convolutional layer of the generator, and form features for each resolution.

$$\mathbf{W} = FC(F_{latent})$$

$$F_{StyleGAN} = StyleGAN(\mathbf{W})$$

It should be noted that since the model uses the pretrained StyleGAN model generator, the model used has rich priors not only of the realistic details and vivid textures in the image. But also, for example, it can also store information about the colors used in real-world images, which can be built into the problem to be solved in the joint image restoration and coloring task. For example, in the task of restoring vintage photographs, which are often either black and white or have a yellowish tinge of fading due to degradation imposed by time. In the case of using an already pre-trained generator, it is hoped that at the stage of restoration of this kind of vintage image, the model, based on its notions of colors, will be able to paint the image in believable colors, along with the main task. It is also hoped that the model used has ideas about the geometry of the face, its three-dimensional representation, to provide greater verisimilitude with real-world samples.

#### 1.4 Spatial feature transform module

To better preserve the accuracy of the model during the facial image generation phase, a spatial feature transformation module, which uses the spatial features of the image extracted by U-Net, is used. The module modulates the signal obtained during generation of generative features  $F_{StyleGAN}$  from the StyleGAN module. The module is necessary, first of all, in order to preserve the local characteristics of the face, based on the spatial information on the input image, in order to increase the accuracy of the process of increasing the resolution of the image, as different zones on the face require different approach to increase the resolution. The module is necessary, first of all, in order to preserve the local characteristics of the face, based on the spatial information on the input image, in order to increase the accuracy of the process of increasing the resolution of the image, as different zones on the face require different approach to increase the

resolution. This is achieved by generating parameters for the affine transform, to spatially modulate the features extracted by the generator, this approach has shown to be effective in works such as [13, 14] for image reconstruction and [13] for image generation.

Thus, for the features obtained on each layer of the StyleGAN generator, a pair of parameters for affine transformation  $(\alpha, \beta)$ , from the input spatial features  $F_{spatial}$ , is generated by several convolution layers. Then the modulation is done by scaling and shifting the features obtained by the generator network:

$$\alpha, \beta = Conv(F_{spatial})$$

$$F_{output} = SFT(F_{GAN} | \alpha, \beta) = \alpha * F_{GAN} + \beta$$

To get a better balance of fidelity and realness, the transformation of spatial features is not performed for the entire signal received by the generator, but for part of the channels of extracted features, the remaining channels in the signal continue to go to the input of the next layer of the generator without change. The scheme is shown in Figure 5.

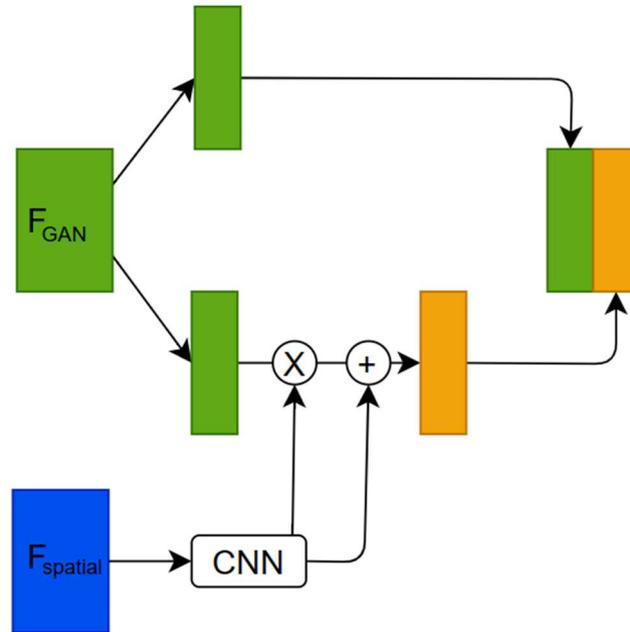


Figure 5 – overview of SFT module

$$F_{output} = CSSFT(F_{GAN} | \alpha, \beta) = Concat(F_{GAN}^{split}, SFT(F_{GAN}^{split} | \alpha, \beta))$$

Where  $F_{GAN}^{split}$ , is part of the  $F_{GAN}$  signal divided in the channel dimension, and the *Concat* function, a simple concatenation of the two feature tensors.

As a result, this module allows both the use of features extracted from the StyleGAN generator, characteristic for faces, and to achieve accuracy by modulating this signal, thus achieving a balance between the realness of the generated image, and the fidelity for the process of increasing the resolution. This module also saves computational resources, since it requires fewer channels, just as in the GhostNet model [1].

Modulation with the help of the spatial transformation of features module, is carried out for each output in the processed StyleGAN model, exactly as the calculation of parameters for shift and scaling. Thus, obtaining a final restored image  $\hat{y}$ .

### 1.5 Model aim

As already mentioned, the model aim is based on several components: 1) the adversarial error computed by the model discriminator to produce realistic images, 2) the local adversarial error computed for the face components, and 3) the perceptual error.

The adversarial error is used to ensure that the model generates images that have a close distance to the real face distribution and natural image manifold, respectively, since the discriminator, just like any other learning model based on an adversarial generator and discriminator, learns to distinguish real images from those generated by the learning model. Also, as for any other discriminator, the error is defined as:

$$L_{adv} = -\lambda_{adv} E_{\hat{y}} \text{softmax}(D(\hat{y}))$$

Where  $D$  is the discriminator's answer to the question whether the image was generated by the model, and  $\lambda_{adv}$  denotes the weight of this error in the total model error.

Facial component error, is used to reinforce features important to the face, namely eyes and mouth. The error is based on local discriminators that calculate their answer for the facial components on the face: left eye, right eye, and mouth. To calculate the error, first the image using an off-the-shelf framework [34] extracts the facial components, then, crops the image into three different images

representing different facial components. Then, three different discriminators are trained for each region in order to distinguish whether the provided images of the facial components are real or derived from the model.

$$L_{local} = -\lambda_{local} \sum_{Region} E_{\hat{y}_{Region}} \left( 1 - softmax \left( D(\hat{y}_{Region}) \right) \right)$$

Where  $\hat{y}_{Region}$  is the region of the image: eyes, mouth,  $D$ , denotes the discriminator response with respect to a given region, and  $\lambda_{local}$  denotes the weight of a given error component.

The perceptual error, which has been used since past works on image resolution enhancement, is also used to estimate the accuracy of the model as applied to the problem of resolution enhancement. To calculate this error component, the pre-trained VGG-19 network is also used, which also has knowledge of the statistics of a large number of images, in order to distinguish one image from another. It is calculated as follows: the generated image and the real high-resolution image are fed to the input of the VGG-19 model, then the images are passed through the VGG-19 model in which the last dense layers responsible for image classification have been removed, then the vectors obtained after passing the images are used to calculate the Euclidean distance between vectors. Eventually, the model, in order to minimize the error, should generate images close to the identical one, since the error forces the model to generate images whose latent space in the VGG model, is close to the latent space of the true image.

$$L_{id} = \lambda_{id} \|\eta(\hat{y}) - \eta(y)\|$$

Where  $\eta$  is the VGG-19 model,  $\lambda_{id}$  is the weight of this loss component in the total model loss formula.

As a result, the formula for the quality functional has the following form:

$$L_{total} = L_{adv} + L_{local} + L_{id}$$

## References

1. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: ECCV. pp. 694–711. Springer (2016).
2. VGG16 and VGG19. URL: <https://keras.io/api/applications/vgg/> (дата обращения 22.06.22)
3. Yu Chen, Ying Tai, Xiaoming Liu, Chunhua Shen, and Jian Yang. Fsrnet: End-to-end learning face super-resolution with facial priors. In CVPR, 2018.
4. Xin Yu, Basura Fernando, Bernard Ghanem, Fatih Porikli, and Richard Hartley. Face super-resolution guided by facial component heatmaps. In ECCV, pages 217–233, 2018.
5. Chaofeng Chen, Xiaoming Li, Lingbo Yang, Xianhui Lin, Lei Zhang, and Kwan-Yee K. Wong. Progressive semantic-aware style transformation for blind face restoration. arXiv:2009.08709, 2020.
6. Xiaoming Li, Ming Liu, Yuting Ye, Wangmeng Zuo, Liang Lin, and Ruigang Yang. Learning warped guidance for blind face restoration. In ECCV, 2018.
7. Xiaoming Li, Wenyu Li, Dongwei Ren, Hongzhi Zhang, Meng Wang, and Wangmeng Zuo. Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion. In CVPR, 2020.
8. Berk Dogan, Shuhang Gu, and Radu Timofte. Exemplar guided face image super-resolution without facial landmarks. In CVPRW, 2019
9. Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In CVPR, 2020
10. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015
11. Jinjin Gu, Yujun Shen, and Bolei Zhou. Image processing using multi-code gan prior. In CVPR, 2020

12. Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. Exploiting deep generative prior for versatile image restoration and manipulation. In ECCV, 2020
13. Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In CVPR, 2018
14. Xiaoming Li, Chaofeng Chen, Shangchen Zhou, Xianhui Lin, Wangmeng Zuo, and Lei Zhang. Blind face restoration via deep multi-scale component dictionaries. In ECCV, 2020
15. Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. Semantic image synthesis with spatially-adaptive normalization. In CVPR, 2019
16. Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In ECCV, 2016.
17. Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In ICCV, 2017
18. Xiaoming Li, Chaofeng Chen, Shangchen Zhou, Xianhui Lin, Wangmeng Zuo, and Lei Zhang. Blind face restoration via deep multi-scale component dictionaries. In ECCV, 2020.

## Приложение Б

Метода класса DataSet, отвечающий за загрузку изображений в модель на этапе обучения.

```
def __getitem__(self, index):

    if self.file_client is None:
        self.file_client = FileClient(self.io_backend_opt.pop('type'),
**self.io_backend_opt)

    gt_path = self.paths[index]
    img_bytes = self.file_client.get(gt_path)
    img_gt = imfrombytes(img_bytes, float32=True)

    # аугментация данных
    img_gt, status = augment(img_gt, hflip=self.opt['use_hflip'], rotation=False,
return_status=True)
    h, w, _ = img_gt.shape

    # Наложение размытия на изображение
    kernel = degradations.random_mixed_kernels(
        self.kernel_list,
        self.kernel_prob,
        self.blur_kernel_size,
        self.blur_sigma,
        self.blur_sigma, [-math.pi, math.pi],
        noise_range=None)
    img_lq = cv2.filter2D(img_gt, -1, kernel)
    # Сжатие изображения
    scale = np.random.uniform(self.downsample_range[0], self.downsample_range[1])
    img_lq = cv2.resize(img_lq, (int(w // scale), int(h // scale)),
interpolation=cv2.INTER_LINEAR)
    # Добвление шума
    if self.noise_range is not None:
        img_lq = degradations.random_add_gaussian_noise(img_lq, self.noise_range)
    # Наложение jpeg артефактов
    if self.jpeg_range is not None:
        img_lq = degradations.random_add_jpg_compression(img_lq, self.jpeg_range)

    # Возвращение к исходному разрешению
    img_lq = cv2.resize(img_lq, (w, h), interpolation=cv2.INTER_LINEAR)

    # Формируем тензоры из изображений
    img_gt, img_lq = img2tensor([img_gt, img_lq], bgr2rgb=True, float32=True)

    # Нормализация
    normalize(img_gt, self.mean, self.std, inplace=True)
    normalize(img_lq, self.mean, self.std, inplace=True)

    return {'lq': img_lq, 'gt': img_gt, 'gt_path': gt_path}
```

Метод прямого прохода модели генератора.

```
def forward(self, x, return_latents=False, return_rgb=True, randomize_noise=True,
**kwargs):

    conditions = []
    unet_skips = []
    out_rgbs = []

    # Энкодер U-Net
```

```

feat = self.conv_body_first(x)
for i in range(self.log_size - 2):
    feat = self.conv_body_down[i](feat)
    unet_skips.insert(0, feat)

feat = self.final_conv(feat)

# Вектор скрытого пространства
style_code = self.final_linear(feat.view(feat.size(0), -1))
if self.different_w:
    style_code = style_code.view(style_code.size(0), -1, self.num_style_feat)

# Декодер U-Net
for i in range(self.log_size - 2):
    # Добавление выходов из отсеченных связей
    feat = feat + unet_skips[i]
    # Апсэмпл U-Net
    feat = self.conv_body_up[i](feat)
    # Генерация тензоров масштаба и сдвига. модуля пространственного преобразования
    признаков
    scale = self.condition_scale[i](feat)
    conditions.append(scale.clone())
    shift = self.condition_shift[i](feat)
    conditions.append(shift.clone())

# Декодер StyleGAN учитывающий пространственное преобразование
image, _ = self.stylegan_decoder([style_code],
                                conditions)

return image

```

### Листинг прямого прохода модифицированного генератора StyleGAN

```

def forward(self,
            styles,
            conditions,
            input_is_latent=False,
            truncation=1,
            truncation_latent=None,
            inject_index=None,
            return_latents=False):
    # Генерация векторов скрытого пространства модели StyleGAN, при помощи полносвязных
    сетей mlp
    if not input_is_latent:
        styles = [self.style_mlp(s) for s in styles]
    # StyleGAN, после каждого блока обратной свёртки добавляет к результату нормальный
    шум, для добавления случайности в изображение
    noise = [None] * self.num_layers # Шум добавляется после каждой обратной свёртки,
    так как чем глубже слой, тем более структурными получаются изменения в изображениях

    # Точно также как и в модели BigGAN, модель StyleGAN использует усечённое нормальное
    распределение для формирования более достоверных изображений так как уменьшается
    изменчивость вектора из скрытого слоя
    style_truncation = []
    for style in styles:
        style_truncation.append(truncation_latent + truncation * (style -
truncation_latent))
    styles = style_truncation

    # Адаптивная нормализация вектора из скрытого пространства модели
    if inject_index is None:
        inject_index = random.randint(1, self.num_latent - 1)

```

```

latent1 = styles[0].unsqueeze(1).repeat(1, inject_index, 1)
latent2 = styles[1].unsqueeze(1).repeat(1, self.num_latent - inject_index, 1)
latent = torch.cat([latent1, latent2], 1)

# Основной блок генерации
out = self.constant_input(latent.shape[0])
out = self.style_conv1(out, latent[:, 0], noise=noise[0])
skip = self.to_rgb1(out, latent[:, 1])

i = 1
for conv1, conv2, noise1, noise2, to_rgb in zip(self.style_convs[::2],
self.style_convs[1::2], noise[2::2], self.to_rgbs):
    out = conv1(out, latent[:, i], noise=noise1) # обратная свёртка модели, на вход
подаются вектора после адаптивной нормализации

    # Модуляция полученного результата
    if i < len(conditions):
        # Модуляция подразумевает масштабирование результата свёртки и его сдвиг
        out = out * conditions[i - 1] + conditions[i]

    out = conv2(out, latent[:, i + 1], noise=noise2)
    skip = to_rgb(out, latent[:, i + 2], skip) # признаки конвертируются в
пространство RGB
    i += 2

image = skip

if return_latents:
    return image, latent
else:
    return image, None

```

Листинг этапа оптимизации обучаемых параметров на этапе обратного распространения ошибки.

```

def optimize_parameters(self, current_iter):
    # Вначале происходит перевычисление весов генератора, поэтому замораживаем веса
дискриминатора
    for p in self.net_d.parameters():
        p.requires_grad = False
    self.optimizer_g.zero_grad()

    # также замораживаем веса дискриминатора лицевых компонент
    for p in self.net_d_left_eye.parameters():
        p.requires_grad = False
    for p in self.net_d_right_eye.parameters():
        p.requires_grad = False
    for p in self.net_d_mouth.parameters():
        p.requires_grad = False

    # Получение регионов глаз и рта моделью
    self.get_roi_regions(eye_out_size=80, mouth_out_size=120)

    l_g_total = 0
    loss_dict = OrderedDict()
    if (current_iter % self.net_d_iters == 0 and current_iter > self.net_d_init_iters):
        # Вычисление ошибки восприятия
        if self.cri_perceptual:
            l_g_percep, l_g_style = self.cri_perceptual(self.output, self.gt)

```

```

        if l_g_percep is not None:
            l_g_total += l_g_percep
            loss_dict['l_g_percep'] = l_g_percep
        if l_g_style is not None:
            l_g_total += l_g_style
            loss_dict['l_g_style'] = l_g_style

    # Генеративная ошибка
    fake_g_pred = self.net_d(self.output)
    l_g_gan = self.cri_gan(fake_g_pred, True, is_disc=False)
    l_g_total += l_g_gan
    loss_dict['l_g_gan'] = l_g_gan

    # Ошибка лицевых компонент
    if self.use_facial_disc:
        # Левый глаз
        fake_left_eye, fake_left_eye_feats = self.net_d_left_eye(self.left_eyes,
return_feats=True)
        l_g_gan = self.cri_component(fake_left_eye, True, is_disc=False)
        l_g_total += l_g_gan
        loss_dict['l_g_gan_left_eye'] = l_g_gan
        # Правый глаз
        fake_right_eye, fake_right_eye_feats = self.net_d_right_eye(self.right_eyes,
return_feats=True)
        l_g_gan = self.cri_component(fake_right_eye, True, is_disc=False)
        l_g_total += l_g_gan
        loss_dict['l_g_gan_right_eye'] = l_g_gan
        # Рот
        fake_mouth, fake_mouth_feats = self.net_d_mouth(self.mouths,
return_feats=True)
        l_g_gan = self.cri_component(fake_mouth, True, is_disc=False)
        l_g_total += l_g_gan
        loss_dict['l_g_gan_mouth'] = l_g_gan

    # Обратный проход
    l_g_total.backward()
    self.optimizer_g.step()

    # EMA
    self.model_ema(decay=0.5**((32 / (10 * 1000))))

    # Вычисление ошибки дискриминатора
    for p in self.net_d.parameters():
        p.requires_grad = True
    self.optimizer_d.zero_grad()
    # Вычисление ошибки для дискриминатора лицевых компонент
    for p in self.net_d_left_eye.parameters():
        p.requires_grad = True
    for p in self.net_d_right_eye.parameters():
        p.requires_grad = True
    for p in self.net_d_mouth.parameters():
        p.requires_grad = True
    self.optimizer_d_left_eye.zero_grad()
    self.optimizer_d_right_eye.zero_grad()
    self.optimizer_d_mouth.zero_grad()

    fake_d_pred = self.net_d(self.output.detach())
    real_d_pred = self.net_d(self.gt)
    l_d = self.cri_gan(real_d_pred, True, is_disc=True) + self.cri_gan(fake_d_pred,
False, is_disc=True)
    loss_dict['l_d'] = l_d
    # Для критика GAN модели Вассерштайна, потеря лежит в районе -1 до 1

```

```

loss_dict['real_score'] = real_d_pred.detach().mean()
loss_dict['fake_score'] = fake_d_pred.detach().mean()
l_d.backward()

# Потеря L1 регуляризатора критика
if current_iter % self.net_d_reg_every == 0:
    self.gt.requires_grad = True
    real_pred = self.net_d(self.gt)
    l_d_r1 = r1_penalty(real_pred, self.gt)
    l_d_r1 = (self.r1_reg_weight / 2 * l_d_r1 * self.net_d_reg_every + 0 *
real_pred[0])
    loss_dict['l_d_r1'] = l_d_r1.detach().mean()
    l_d_r1.backward()

# Обратный проход для критика
self.optimizer_d.step()

```

Листинг тестирования получившейся после обучения модели.

```

arch = 'original'
channel_multiplier = 1
model_name = 'GFPGAN'

model_path = os.path.join('gfpgan/weights', model_name + '.pth')

restorer = GFPGANer(
    model_path=model_path,
    arch=arch,
    channel_multiplier=channel_multiplier)

# Восстановление лиц
for img_path in img_list:
    # Загрузка изображения gt
    img_name = os.path.basename(img_path)
    print(f'Processing {img_name} ...')
    basename, ext = os.path.splitext(img_name)
    input_img = cv2.imread(img_path, cv2.IMREAD_COLOR)

    # Восстановление фотографии
    restored_img = restorer.enhance(
        input_img,
        has_aligned=args.aligned,
        only_center_face=args.only_center_face,
        weight=args.weight)

    # Сохранение фотографий
    if restored_img is not None:
        if args.ext == 'auto':
            extension = ext[1:]
        else:
            extension = args.ext

        if args.suffix is not None:
            save_restore_path = os.path.join(args.output, 'restored_imgs',
f'{basename}_{args.suffix}.{extension}')
        else:
            save_restore_path = os.path.join(args.output, 'restored_imgs',
f'{basename}.{extension}')
        imwrite(restored_img, save_restore_path)

```