



Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный технический университет  
имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)

---

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ

## **РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ**

**НА ТЕМУ:**

**Технология улучшения изображений, основанная на  
глубоком обучении**

Студент ИУ5И-32М

—  
(Группа)  
(И.О.Фамилия)

Джин Шуо

(Подпись, дата)

Руководитель

Гапанюк Ю.Е.

(Подпись, дата)

(И.О.Фамилия)

2024 г.

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный технический университет имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)

УТВЕРЖДАЮ

Заведующий кафедрой ИУ-5  
(Индекс)

В.И.Терехов  
(И.О.Фамилия)

« \_\_\_\_\_ » \_\_\_\_\_ 2024 г.

## ЗАДАНИЕ

### на выполнение научно-исследовательской работы

по теме Технология улучшения изображений, основанная на глубоком обучении

—

Студент группы ИУ5И-32М

Джин Шуо  
(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)  
учебная

Источник тематики (кафедра, предприятие, НИР) учебная тематика

График выполнения НИР: 25% к 12 нед., 50% к 14 нед., 75% к 15 нед., 100% к 16 нед.

**Техническое задание** Обзор основных методов анализа текстовых данных, разработка системы для автоматизированного анализа текстовых данных из социальных медиа с целью выявления трендов и настроений.

#### Оформление научно-исследовательской работы:

Расчетно-пояснительная записка на 11 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

—

Дата выдачи задания « 14 » сентября 2024 г.

Руководитель НИР

Студентка

<u>Галапюк Ю.Е.</u>		
(Подпись,	дата	И.О.Фамилия)
<u>Джин Шуо</u>	<u>. 2024 г.</u>	
(Подпись,	дата	И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

## **Оглавление**

<b>Введение .....</b>	<b>4</b>
<b>1. Обзор методов улучшения изображений и применения глубокого обучения .....</b>	<b>5</b>
1.1 Основные методы улучшения изображений.....	5
1.2 Применение глубокого обучения в улучшении изображений .....	5
1.3 Оценка качества улучшения изображений.....	6
<b>2. Детальное описание выбранных методов .....</b>	<b>7</b>
2.1 Генеративные состязательные сети (GAN).....	7
2.2 Суперразрешение (Super-Resolution) .....	8
2.3 Методы с использованием механизма самовнимания (Self-Attention) .....	9
2.4 Сравнительная таблица методов .....	9
<b>3. Теоретическое сравнение и анализ результатов .....</b>	<b>10</b>
3.1 Основные критерии сравнения .....	10
3.2 Сравнительный анализ.....	10
3.3 Теоретические выводы .....	11
3.4 Предполагаемые результаты при сравнении методов .....	11
3.5 Ограничения методов.....	11
3.6 Перспективы дальнейшего развития.....	12
<b>4. Обсуждение перспектив и выводы .....</b>	<b>13</b>
4.1 Перспективы развития методов улучшения изображений.....	13
4.2 Выводы.....	14
<b>Список использованных источников.....</b>	<b>16</b>

## **Введение**

Глубокое обучение значительно изменило подходы к улучшению изображений. Традиционные методы, такие как фильтрация и интерполяция, имеют ограничения, особенно в сложных задачах, таких как восстановление деталей и повышение качества изображений. С развитием технологий глубокого обучения, таких как сверточные нейронные сети (CNN) и генеративные состязательные сети (GAN), улучшение изображений стало более эффективным, обеспечивая высокое качество и реалистичность.

Глубокое обучение позволяет автоматически извлекать высокоуровневые признаки из изображений, что значительно улучшает их чёткость, детализацию и восстанавливает утраченные элементы. Такие технологии активно применяются в различных областях: от медицины и спутниковой съёмки до искусства и безопасности.

Целью данной работы является теоретическое сравнение трёх популярных методов улучшения изображений: основанных на GAN, методах суперразрешения и технологиях самовнимания. Мы рассмотрим теоретические аспекты каждого подхода, их преимущества и ограничения, а также определим, какой метод лучше подходит для различных типов задач.

# **1. Обзор методов улучшения изображений и применения глубокого обучения**

## **1.1 Основные методы улучшения изображений**

Улучшение изображений — это процесс повышения их визуального качества с целью улучшения видимости важных деталей. Существует множество подходов, от традиционных методов до более современных решений, использующих искусственный интеллект.

### **Традиционные методы улучшения изображений**

Традиционные алгоритмы, такие как фильтрация, коррекция яркости и контраста, а также алгоритмы интерполяции, позволяют улучшить изображение в случае потери чёткости или при наличии шума. Например, гистограммная обработка используется для улучшения контраста изображения, а фильтрация Гаусса помогает сглаживать изображение и уменьшать шум. Однако такие методы часто ограничены, когда речь идет о сложных или сильно повреждённых изображениях, где важны высокоуровневые детали, как текстуры и структуры.

### **Современные методы на основе глубокого обучения**

Совсем недавно с развитием глубокого обучения появились новые методы, значительно улучшившие результаты по сравнению с традиционными. Глубокие нейронные сети позволяют извлекать скрытые особенности изображений, улучшать качество деталей и делать их более естественными. К наиболее популярным методам относятся:

Сверточные нейронные сети (CNN) — один из самых широко используемых типов нейронных сетей для обработки изображений. CNN эффективно применяются в задачах восстановления деталей и повышения разрешения изображений.

Генеративные состязательные сети (GAN) — эта архитектура включает две нейронные сети: генератор, который создаёт изображение, и дискриминатор, который оценивает его правдоподобие. GAN широко применяются в задачах переноса стиля и генерации новых изображений, которые выглядят максимально реалистично.

Суперразрешение (Super-Resolution) — методы суперразрешения направлены на улучшение качества изображений с низким разрешением путём их масштабирования до более высокого разрешения с сохранением или улучшением деталей. Наиболее распространённым подходом является использование сверточных нейронных сетей для восстановления утраченной информации.

## **1.2 Применение глубокого обучения в улучшении изображений**

Глубокое обучение революционизировало подходы к улучшению изображений благодаря своей способности извлекать сложные зависимости и контексты из данных.

Рассмотрим несколько ключевых направлений применения глубокого обучения в улучшении изображений.

#### 1. Удаление шума и восстановление изображений

Одной из ключевых задач в улучшении изображений является удаление шума, который может быть добавлен при захвате изображения. Глубокие нейронные сети, особенно сверточные сети, продемонстрировали высокую эффективность в фильтрации шума, восстанавливая при этом важные детали изображения. Например, сети, обученные на изображениях с шумом, могут восстанавливать оригинальные изображения, что невозможно с помощью традиционных методов.

#### 2. Суперразрешение изображений

С помощью методов суперразрешения можно увеличивать разрешение изображений без значительных потерь в качестве. Эта технология активно используется в таких областях, как медицинская диагностика (для улучшения изображений рентгеновских снимков или МРТ), спутниковая съёмка и криминалистика. Современные модели, такие как ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), значительно улучшили результаты, предлагая методы восстановления деталей, которых не хватает в исходном изображении.

#### 3. Перенос стиля и генерация изображений

Перенос стиля и генерация изображений с использованием GAN является одними из самых впечатляющих достижений в области улучшения изображений. GAN способны не только восстанавливать изображения, но и создавать новые изображения, стилизованные под определённые художественные направления. Например, CycleGAN может преобразовывать фотографии в картины в стиле различных художников, что находит широкое применение в искусстве и дизайне.

#### 4. Обработка и улучшение медицинских изображений

В области медицины глубокое обучение активно используется для улучшения качества медицинских изображений. Например, модели глубокого обучения могут улучшать рентгеновские снимки или МРТ-изображения, помогая врачам более точно диагностировать заболевания. Методы суперразрешения позволяют врачам получать более чёткие снимки с меньшими дозами облучения.

### 1.3 Оценка качества улучшения изображений

Для оценки качества улучшенных изображений в практике часто используются несколько метрик. Одной из самых распространённых является PSNR (Peak Signal-to-Noise Ratio), которая измеряет уровень шума на изображении и помогает определить, насколько улучшилось качество по сравнению с исходным. SSIM (Structural Similarity Index) — это метрика, которая оценивает структурные и текстурные изменения между двумя изображениями, предлагая более точную оценку визуального восприятия. Оба этих показателя важны для анализа эффективности моделей улучшения изображений.

PSNR рассчитывается как отношение максимальной возможной мощности сигнала к мощности шума. Высокий PSNR обычно указывает на хорошее качество восстановленного изображения, но он не всегда соответствует реальному восприятию человеком.

SSIM учитывает восприятие человеческим зрением и используется для оценки визуального качества изображений, сопоставляя структурные, текстурные и яркостные элементы.

## **2. Детальное описание выбранных методов**

В этой главе будет представлено подробное описание трёх методов улучшения изображений на основе глубокого обучения: Генеративные состязательные сети (GAN), суперразрешение (Super-Resolution) и методы с использованием механизма самовнимания (Self-Attention).

### **2.1 Генеративные состязательные сети (GAN)**

Генеративные состязательные сети, предложенные Иэном Гудфеллоу в 2014 году, представляют собой одну из наиболее мощных технологий глубокого обучения. GAN состоят из двух нейронных сетей — генератора и дискриминатора, которые соревнуются друг с другом.

Генератор создаёт изображения, пытаясь воспроизвести исходное распределение данных.

Дискриминатор оценивает, является ли изображение настоящим (из исходных данных) или сгенерированным генератором.

Соревновательный процесс приводит к улучшению качества создаваемых изображений, так как генератор учится обманывать дискриминатор, создавая всё более реалистичные изображения.

Применение GAN в улучшении изображений

GAN широко используются для решения задач улучшения изображений:

Восстановление повреждённых изображений — GAN могут восстанавливать утраченные области изображения (inpainting). Например, алгоритмы, такие как DeepFill, применяются для заполнения пропусков на изображениях.

Увеличение разрешения — модели, такие как SRGAN (Super-Resolution GAN), используются для преобразования изображений низкого разрешения в высококачественные изображения.

Перенос стиля — CycleGAN и Pix2Pix используются для преобразования изображений из одного домена в другой (например, преобразование фотографии в рисунок).

Преимущества и недостатки GAN

Преимущества: способность генерировать реалистичные изображения, универсальность применения.

Недостатки: трудности в обучении (нестабильность), требовательность к вычислительным ресурсам.

## 2.2 Суперразрешение (Super-Resolution)

Суперразрешение — это процесс преобразования изображения низкого разрешения в изображение высокого разрешения. Основной целью является восстановление утраченных деталей и улучшение чёткости изображения.

Архитектура моделей суперразрешения

Наиболее популярной архитектурой для суперразрешения являются сверточные нейронные сети (CNN).

SRCNN (Super-Resolution Convolutional Neural Network) — одна из первых CNN для суперразрешения. Она использует три основных слоя для интерполяции, нелинейного отображения и восстановления.

ESRGAN (Enhanced Super-Resolution GAN) — улучшенная версия SRGAN, которая обеспечивает более высокое качество изображений благодаря использованию perceptual loss, учитывающей восприятие человеком.

Применение суперразрешения

Медицинская визуализация: улучшение рентгеновских снимков или изображений МРТ для диагностики.

Спутниковые изображения: повышение детализации снимков для анализа поверхности земли.

Улучшение фотографий: масштабирование изображений для профессиональной печати.

Преимущества и недостатки суперразрешения

Преимущества: высокая эффективность в восстановлении утраченных деталей, простота архитектур.

Недостатки: ограниченность при обработке сильно повреждённых изображений.



## 2.3 Методы с использованием механизма самовнимания (Self-Attention)

Механизм самовнимания стал ключевым компонентом современных моделей глубокого обучения, таких как трансформеры. Самовнимание позволяет модели фокусироваться на важных частях изображения, что особенно полезно для задач, где необходимо учитывать глобальные и локальные зависимости.

### Архитектура моделей с самовниманием

Image Transformer — адаптация трансформеров для обработки изображений, где пиксели взаимодействуют друг с другом через механизм внимания.

Vision Transformer (ViT) — использует механизм самовнимания для разделения изображения на небольшие фрагменты (patches) и обработки их как последовательности данных.

### Применение самовнимания в улучшении изображений

Восстановление деталей — модели с самовниманием могут анализировать глобальные зависимости, что делает их особенно эффективными в задачах восстановления сильно повреждённых изображений.

Перенос стиля — самовнимание помогает моделям учитывать контекст изображения, улучшая качество стилизации.

### Преимущества и недостатки самовнимания

Преимущества: высокая точность за счёт анализа глобального контекста, эффективность в задачах с высоким разрешением.

Недостатки: высокая вычислительная сложность, требовательность к памяти.

## 2.4 Сравнительная таблица методов

Метод	Применение	Преимущества	Недостатки
GAN	Генерация и улучшение изображений	Реалистичные результаты	Сложность обучения, ресурсоёмкость
Суперразрешение	Увеличение разрешения	Восстановление утраченных деталей	Ограниченность в сложных задачах
Самовнимание	Восстановление, перенос стиля	Учет глобального контекста	Высокая вычислительная сложность

### 3. Теоретическое сравнение и анализ результатов

В этой главе мы проведём теоретическое сравнение трёх выбранных методов улучшения изображений: GAN, суперразрешение и методы с самовниманием, проанализируем их преимущества, недостатки и подходящие области применения.

#### 3.1 Основные критерии сравнения

Для объективного анализа методов улучшения изображений выделим несколько ключевых критериев:

1. **Качество результатов:** степень визуального улучшения изображения, включая восстановление деталей и снижение артефактов.
2. **Сложность обучения:** вычислительные ресурсы и время, необходимые для обучения модели.
3. **Гибкость и универсальность:** способность модели адаптироваться к различным задачам и типам изображений.
4. **Стабильность и надёжность:** предсказуемость результата при работе с различными данными.

#### 3.2 Сравнительный анализ

Критерий	GAN	Суперразрешение	Методы с самовниманием
Качество результатов	Высокое: реалистичное улучшение деталей, особенно в задачах генерации.	Высокое: точное восстановление деталей, особенно при увеличении разрешения.	Высокое: глубокий анализ глобального контекста, полезен для сложных изображений.
Сложность обучения	Очень высокая: обучение требует настройки баланса между генератором и дискриминатором.	Средняя: стабильные архитектуры, требующие меньшего количества параметров.	Высокая: требует значительных вычислительных ресурсов из-за сложности архитектуры.
Гибкость и универсальность	Универсальный метод, применимый к генерации, восстановлению и переносу стиля.	Узконаправленный метод, подходит для задач масштабирования.	Универсальный, особенно эффективен в задачах обработки сложных изображений.

<b>Стабильность</b>	Может быть нестабильным из-за проблемы модового коллапса.	Высокая: результаты предсказуемы.	Высокая, но может страдать от переобучения на больших данных.
---------------------	---	-----------------------------------	---

### 3.3 Теоретические выводы

GAN демонстрируют превосходное качество в задачах, где требуется создание или восстановление изображений, однако сложность их обучения и возможная нестабильность ограничивают их использование в реальном времени. Они лучше всего подходят для художественных задач (например, перенос стиля) и восстановления сильно повреждённых изображений.

Суперразрешение является более стабильным и лёгким в обучении методом. Оно эффективно для задач масштабирования изображений, особенно в областях, где требуется точное восстановление утраченных деталей, таких как медицинская визуализация или спутниковая съёмка. Однако этот метод ограничен узким спектром задач.

Методы с самовниманием предлагают революционные подходы к обработке изображений, учитывая глобальный контекст и мелкие детали одновременно. Они особенно полезны в задачах, где требуется анализ сложных зависимостей, например, при восстановлении сильно повреждённых или высокодетализированных изображений. Однако их высокая вычислительная сложность ограничивает использование в системах с ограниченными ресурсами.

### 3.4 Предполагаемые результаты при сравнении методов

Теоретический анализ показывает, что для различных типов задач каждый метод имеет свои преимущества:

- Для задач восстановления повреждённых изображений лучше всего подходят GAN, так как они способны генерировать реалистичные фрагменты.
- Для задач масштабирования и повышения разрешения более оптимальны методы суперразрешения.
- Для задач, требующих глубокого анализа структуры изображения, например, переноса стиля или обработки высокодетализированных данных, превосходство демонстрируют методы с самовниманием.

### 3.5 Ограничения методов

Несмотря на их преимущества, каждый метод имеет свои ограничения:

1. **GAN:** высокая вероятность модового коллапса, сложность настройки.
2. **Суперразрешение:** ограниченность в задачах, требующих генерации новых данных.
3. **Самовнимание:** высокая потребность в памяти и вычислительных ресурсах.

### 3.6 Перспективы дальнейшего развития

1. Интеграция методов: использование GAN и механизмов самовнимания в единой архитектуре для повышения качества и стабильности.
2. Оптимизация ресурсов: разработка облегчённых моделей, способных эффективно работать на устройствах с ограниченными вычислительными мощностями.
3. Адаптивное обучение: создание моделей, которые могут автоматически подстраиваться под специфику задачи, минимизируя необходимость ручной настройки.

## 4. Обсуждение перспектив и выводы

В этой главе мы подведём итоги анализа методов улучшения изображений на основе глубокого обучения, обсудим перспективы их развития и возможные направления для будущих исследований и практического применения.

### 4.1 Перспективы развития методов улучшения изображений

Методы улучшения изображений на основе глубокого обучения стремительно развиваются, и существует несколько ключевых направлений, которые могут значительно повлиять на будущее этой области.

#### Интеграция различных методов

Совмещение различных подходов, таких как генеративные состязательные сети (GAN), суперразрешение и методы с самовниманием, открывает новые возможности для улучшения качества изображений. Например, можно интегрировать возможности GAN для генерации деталей с преимуществами методов суперразрешения для увеличения разрешения, а также использовать механизмы самовнимания для анализа глобального контекста. Такое объединение позволит добиться ещё более высококачественных и стабильных результатов, которые могут применяться в различных сферах, от медицины до развлечений.

#### Уменьшение вычислительных затрат

Одной из главных проблем современных моделей глубокого обучения является высокая вычислительная нагрузка, которая ограничивает их применение в реальных условиях, особенно на устройствах с ограниченными ресурсами (например, мобильных телефонах или встроенных системах). Разработка более лёгких и оптимизированных моделей, таких как MobileNet или EfficientNet, позволит значительно сократить время обработки и требования к вычислительным ресурсам. Это, в свою очередь, обеспечит доступность технологий улучшения изображений для более широкого круга пользователей.

#### Адаптивное обучение

В будущем возможно создание моделей, которые смогут автоматически адаптироваться под специфику задач и данных. Например, алгоритмы могут обучаться не только на большом объёме данных, но и использовать онлайн-обучение, что позволит адаптироваться к новым типам изображений без необходимости повторного обучения модели с нуля. Такие подходы особенно полезны для динамичных сфер, таких как видеонаблюдение, где данные постоянно изменяются.

## Использование мультимодальных данных

С увеличением доступности данных из различных источников (например, с камер с высокой частотой кадров, спутниковых снимков, медицинских изображений и т.д.), есть перспектива использования мультимодальных данных для улучшения качества изображений. Комбинированное использование данных с разных сенсоров или источников может помочь не только улучшить изображения, но и добавить дополнительные слои информации, такие как контекст или временные изменения. Это может открыть новые горизонты в таких областях, как мониторинг окружающей среды или анализ медицинских снимков.

## 4.2 Выводы

На основе проведённого анализа и теоретического сравнения различных методов улучшения изображений можно сделать несколько важных выводов:

GAN остаются одним из самых мощных инструментов для создания и восстановления изображений, обеспечивая реалистичные результаты. Однако их высокая сложность и нестабильность ограничивают их применение в реальных приложениях.

Суперразрешение — это более стабильный и менее ресурсоёмкий метод, который показывает отличные результаты в задачах увеличения разрешения изображений. Однако он ограничен в применении к задачам генерации новых данных и может не справляться с сильно повреждёнными изображениями.

Методы с самовниманием предлагают революционные подходы для анализа изображений, учитывая как локальные, так и глобальные зависимости. Эти методы особенно полезны для сложных изображений, требующих глубокого анализа. Однако они требуют значительных вычислительных ресурсов и могут быть ограничены в применении на устройствах с ограниченными мощностями.

Перспективы комбинирования методов открывают новые возможности для достижения лучших результатов в различных областях. Интеграция разных подходов позволит использовать сильные стороны каждого из них, создавая более гибкие и эффективные системы.

Будущее технологий улучшения изображений связано с уменьшением вычислительных затрат, развитием адаптивных моделей и применением мультимодальных данных. Это обеспечит широкое распространение технологий и их интеграцию в различные сферы жизни, включая медицину, искусство, образование и безопасность.

В целом, технологии улучшения изображений на основе глубокого обучения продолжают развиваться, и их потенциал для различных приложений огромен. Ожидается, что в будущем эти методы станут более доступными, точными и быстрыми, что откроет новые возможности для использования в реальном времени.

## **Список использованных источников**

- [1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems (NeurIPS)*, 27, 2672-2680.
- [2] Ledig, C., Theis, L., Huszár, F., Caballero, J., Aitken, A. P., Tejani, A., ... & Wang, Z. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 105-114.
- [3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. A., ... & Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 5998-6008.
- [4] Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 38(2), 295-307.
- [5] Zhang, X., & Wu, Y. (2018). Image Restoration Using Deep Learning: A Survey. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 1937-1945.
- [6] Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. *Proceedings of the 2nd International Conference on Learning Representations (ICLR)*.
- [7] Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *Proceedings of the International Conference on Machine Learning (ICML)*, 2716-2724.
- [8] Zhao, Z., & Cheng, H. (2020). Attention-based Deep Learning Models for Image Enhancement. *Journal of Imaging Science and Technology*, 64(3), 30301-1-30301-10.
- [9] Chen, L., & Zhang, J. (2020). Deep Learning for Image Super-Resolution: A Survey. *Journal of Visual Communication and Image Representation*, 70, 102763.
- [10] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.