

На правах рукописи



Бережнов Никита Игоревич

**СОВЕРШЕНСТВОВАНИЕ МЕХАНИЗМОВ ВНИМАНИЯ В
ГЛУБОКИХ НЕЙРОННЫХ СЕТЯХ – ТРАНСФОРМЕРАХ В ЗАДАЧАХ
ВОССТАНОВЛЕНИЯ И АУГМЕНТАЦИИ ИЗОБРАЖЕНИЙ**

Специальность 1.2.1. Искусственный интеллект и машинное обучение

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата
технических наук

Воронеж – 2025

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Воронежский государственный университет»

Научный руководитель:

доктор технических наук, профессор
Сирота Александр Анатольевич

Официальные оппоненты:

Приоров Андрей Леонидович

доктор технических наук, профессор, федеральное государственное бюджетное образовательное учреждение высшего образования «Ярославский государственный университет им. П.Г. Демидова», кафедра цифровых технологий и машинного обучения, профессор

Шипко Владимир Вацлавович

доктор технических наук, федеральное государственное казенное военное образовательное учреждение высшего образования «Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина», кафедра робототехнических комплексов и систем воздушного базирования, доцент

Ведущая организация:

Федеральное государственное автономное образовательное учреждение высшего образования «Белгородский государственный национальный исследовательский университет», г. Белгород

Защита состоится 6 февраля 2026 года в 14 часов в конференц-зале на заседании диссертационного совета 24.2.288.11, созданного на базе федерального государственного бюджетного образовательного учреждения высшего образования «Воронежский государственный университет», по адресу: 394018, г. Воронеж, Университетская площадь, 1.

С диссертацией можно ознакомиться в научной библиотеке Воронежского государственного университета и на сайте: <http://www.science.vsu.ru/dissinfo&cand=3550>

Автореферат разослан «__» _____ 2025 г.

Ученый секретарь

диссертационного совета 24.2.288.11
кандидат физико-математических наук,
доцент



Медведева
Ольга Александровна

Актуальность темы диссертации. Современные системы компьютерного зрения предъявляют высокие требования к качеству визуальной информации, получаемой при обработке изображений в автоматическом режиме. В реальных условиях изображения, поступающие на вход таких систем, часто оказываются искаженными из-за наличия шумов, влияния погодных осадков, технических ограничений сенсоров, случайных помех или ошибок передачи данных. Решение задач восстановления и улучшения качества изображений, а также эффективной аугментации обучающих данных становится особенно важным в контексте роста сложности прикладных задач и разнообразия реальных условий съемки.

В работе используются понятия восстановления изображений и улучшение качества изображений. Восстановление направлено на устранение любого рода искажений, возникших в процессе регистрации изображения с целью приблизить изображение к его исходному виду, при этом задается количественный критерий близости (метрика). В задаче улучшения в большей степени исследователи фокусируются на субъективном повышении качества визуального восприятия изображения и его информативности. Хотя эти задачи и близко связаны, но, тем не менее, не всегда эквивалентны. Далее в качестве основной будет рассматриваться задача восстановления изображений.

В последние годы активно развиваются методы глубокого обучения, демонстрирующие высокую устойчивость получаемых при их использовании результатов к сложным типам искажений. Среди них большую роль играют архитектуры глубоких нейронных сетей (ГНС) трансформерного типа, успешно адаптированные для задач компьютерного зрения. Механизм внимания (attention), лежащий в их основе, позволяет эффективно учитывать как локальные, так и глобальные взаимосвязи между элементами изображения, что важно при устранении артефактов, зашумления, затенения или других отклонений от исходного вида изображения.

Однако практическое применение трансформеров в задачах восстановления изображений часто сопряжено с рядом ограничений. Во-первых, внимание может избыточно фокусироваться на отдельных структурах изображений, что снижает эффективность восстановления от шумов и искажений. Во-вторых, большая глубина архитектур делает их склонными к переобучению, особенно при ограниченных объемах обучающих данных. В-третьих, известные трансформерные модели, используемые в задачах обработки изображений, весьма громоздки и требуют больших вычислительных ресурсов (памяти, времени), особенно на этапе обучения. В-четвертых, сбор достаточного объема разнообразных изображений, необходимых для обучения в реальных условиях затруднен, что делает задачу формирования качественного обучающего датасета нетривиальной.

Решение проблемы нехватки данных требует применения как классических приемов аугментации, так и генеративных нейросетевых моделей, способных формировать новые условно-реалистичные изображения. Особенно актуален этот подход в задачах, связанных с нечеткими, нелинейными или аппликативными видами шумов, где традиционные методы фильтрации оказываются неэффективными. В этом плане применение генеративных ГНС позволяет формировать синтетические изображения, имитирующие сложные искажения, атмосферные осадки, цифровые дефекты, шумы различных типов. Это открывает возможности как для создания обширных обучающих наборов данных, так и для реализации архитектур, в которых восстановление и аугментация изображений выступают как взаимосвязанные стадии единого процесса обработки информации. В частности, генерация зашумленных или стилизованных изображений может быть непосредственно добавлена в этапы обучения нейросетевых моделей, обеспечивая лучшее понимание вариативности входных данных.

Однако подобные подходы требуют более тщательной регуляризации и контроля процесса обучения. В противном случае возникают риски искажения исходного распределения данных, доменного смещения данных (bias) и ухудшения обобщающей способности модели. Кроме того, существующие генеративные модели имеют сложности в процессе обучения и склонность к неконтролируемому синтезу изображений. Это требует разработки гибридных архитектур и дальнейшего совершенствования механизмов внимания с учетом специфики решаемой задачи.

Таким образом, актуальными являются исследования комбинированных подходов для решения задачи восстановления изображений, сочетающих относительно «легковесные» архитектуры трансформеров с модифицированными механизмами канального и пространственного внимания и современные методы аугментации данных, обеспечивающие увеличение устойчивости, точности и обобщающей способности обучаемых нейросетевых моделей. Такие решения позволяют не только обеспечить качественное восстановление изображений, но и повысить устойчивость систем компьютерного зрения в реальных и, зачастую, нестабильных условиях.

Степень разработанности темы диссертации. Алгоритмы восстановления и аугментации изображений достаточно давно развиваются в области компьютерного зрения и машинного обучения. Классические подходы, основанные на фильтрации, частотных преобразованиях и методах регуляризации, подробно рассматривались в трудах Н.Н. Бондиной, В.С. Сизикова, В.П. Кузнецова, В.К. Клочко, Р.С. Гонсалеса, С.О. Емельянова, В.И. Тихонова. Однако подобные методы часто оказываются неэффективными при обработке изображений, содержащих сложные и нелинейные искажения, такие как шумы нестандартной природы, атмосферные эффекты или структурные дефекты.

Существенный прогресс в решении этих задач достигнут за счет использования методов глубокого обучения, в частности, сверточных ГНС и архитектур автокодировщиков. Однако основные успехи здесь связаны с появлением трансформеров, реализующих принципы механизмов внимания (самовнимания), адаптированных к задачам обработки изображений. Архитектуры Vision Transformer (ViT), Swin Transformer, Restormer и их производные, предложенные в работах А. Dosovitskiy, Z. Liu, Н. Zhao, S.W. Zamir, J. Valanarasu, P. Isola и других исследователей, продемонстрировали высокую эффективность в задачах восстановления изображений, улучшения визуального качества и удаления шумов.

Развитие механизмов внимания и их модификаций на основе структурной регуляризации и стохастического сглаживания подробно рассматривались в исследованиях А. Vaswani, В. Li, W. Zhou, Н. Lu, X. Chen, К. He и др. Тем не менее, несмотря на достигнутый прогресс, актуальной остается проблема переобучения и повышения эффективности моделей трансформерного типа, особенно в условиях недостаточности обучающих выборок.

Для решения данной проблемы используются также методы аугментации изображений. Наряду с эвристическими подходами все более широкое распространение получают генеративные методы, включая модели класса GAN, VAE, диффузионные модели и трансформеры с перекрестным вниманием. Они позволяют создавать синтетические изображения с реалистичными атмосферными искажениями, шумами и стилем, что существенно расширяет возможности моделирования входных данных.

Тем не менее, несмотря на большое количество исследований, остаются нерешенными многие задачи интеграции модифицированных механизмов внимания и алгоритмов генерации синтетических искажений в единую схему обработки, способную одновременно восстанавливать изображения и эффективно увеличивать набор обучающих данных. Актуальной задачей также остается разработка малозатратных архитектур для обработки изображений, обеспечивающих качество, сопоставимое с лучшими известными большими моделями, и устойчивых к проблемам, возникающим при несбалансированной генерации и некорректной стилизации изображений.

Таким образом, настоящая работа продолжает и развивает научные исследования, направленные на повышение эффективности нейросетевых архитектур восстановления изображений на основе сетей трансформерного типа с модифицированными механизмами внимания и аугментацией обучающих данных для создания реалистических изображений в условиях погодных искажений и воздействия различных шумов. Работа находится в русле актуальных направлений исследований в области искусственного интеллекта и машинного обучения.

Цель и задачи исследования. Целью диссертационной работы является совершенствование алгоритмов восстановления изображений в условиях различных искажений

(включая шумы, атмосферные осадки, аппликативные помехи) на основе архитектур ГНС трансформерного типа с модифицированными механизмами внимания и применение средств аугментации данных, направленных на повышение обобщающей способности нейронных сетей, достигаемой в процессе обучения.

Для достижения указанной цели в работе решаются следующие задачи:

1. Анализ современных подходов к восстановлению и аугментации изображений, выявление ограничений существующих алгоритмов и обоснование необходимости внедрения новых архитектурных решений в моделях глубокого обучения.

2. Теоретическое обоснование и разработка методов модификации и регуляризации механизма внимания в трансформерах, направленных на повышение их устойчивости к переобучению и улучшение качества восстанавливаемых изображений.

3. Синтез архитектур трансформеров с усовершенствованными канальным и пространственным механизмами внимания для восстановления изображений в условиях сложных помех и искажений.

4. Разработка алгоритмов генерации и стилизации изображений на основе ГНС в целях аугментации обучающих данных, включая синтез реалистичных изображений объектов в условиях атмосферных осадков.

5. Разработка программного комплекса и методики его применения, реализующих комплексное применение предложенных алгоритмов восстановления и аугментации изображений, проведение экспериментального анализа для выявления их эффективности в типовых задачах компьютерного зрения.

Объект исследования. Объектом исследования являются системы компьютерного зрения для восстановления и аугментации изображений.

Предмет исследования. Предметом исследования являются модели и алгоритмы глубокого обучения на основе трансформеров с модифицированными механизмами внимания, а также алгоритмы генерации и стилизации условно-реальных изображений, используемые для повышения качества изображений и увеличения обучающих выборок в задачах компьютерного зрения.

Методы исследования. В ходе выполнения диссертационной работы использовались методы математического анализа, линейной алгебры и оптимизации, методы теории вероятностей, методы цифровой обработки изображений, модели и методы глубокого машинного обучения, технологии разработки многослойных ГНС (в том числе, сверточных и трансформерных архитектур), модели и методы генерации и стилизации изображений, методы и средства имитационного моделирования, технологии программирования ГНС с использованием современных инструментальных сред.

Научная новизна диссертации заключается в следующем.

1. Предложен и теоретически обоснован способ структурной регуляризации механизма внимания в нейронных сетях трансформерного типа, отличающийся использованием мультипликативной и аддитивной стохастической составляющих, вносимых при вычислении матриц весов внимания, что обеспечивает сглаживание распределения весов для предотвращения их неконтролируемого роста в процессе обучения.

2. Предложен и теоретически обоснован способ структурной регуляризации процесса обучения трансформеров, отличающийся использованием обучаемой матрицы масштабных коэффициентов, что позволяет оказывать позитивное влияние в ситуациях насыщения активационной функции механизма внимания.

3. Разработаны и исследованы модификации канального механизмов внимания в трансформерах, отличающиеся использованием сжатия канальных признаков. Предложены способы структурной регуляризации пространственного внимания. Это позволило повысить качество восстановления изображений при одновременном снижении вычислительной сложности моделей. На основе этого предложены улучшенные архитектуры нейронных сетей трансформерного типа. Проведены экспериментальные исследования, подтверждающие

возникновение положительного эффекта в задачах восстановления изображений и улучшение значений метрик качества по сравнению с базовыми прототипами.

4. Разработаны модели и алгоритмы аугментации изображений на основе специализированных архитектур нейронных сетей. Особое внимание уделено синтезированию изображений объектов в условиях атмосферных осадков, затрудняющих восприятие анализируемых сцен. Предложена новая архитектура модели трансформер, объединяющая сверточный энкодер-декодер и перекрестный механизм внимания для генерации атмосферных осадков (дождь, снег, туман), позволяющая сохранить структурную целостность сцены. Введена составная функция потерь для обучения предложенной модели в условиях различных погодных осадков, учитывающая различные аспекты качества синтеза изображений.

5. Разработан программный комплекс алгоритмов восстановления и аугментации изображений, основанный на объединении архитектур трансформеров с усовершенствованными механизмами внимания и специализированных моделей аугментации изображений и синтеза искажений. Предложена методика их совместного применения для повышения устойчивости нейросетевых моделей к различным помехам в условиях нехватки обучающих данных. Экспериментально подтверждена эффективность использования синтезированных изображений в качестве обучающих данных в задачах восстановления, классификации и сегментации.

Соответствие паспорту специальности. Полученные в диссертации результаты, характеризующиеся научной новизной, соответствуют следующим пунктам паспорта специальности 1.2.1 Искусственный интеллект и машинное обучение:

п.4. Разработка методов, алгоритмов и создание систем искусственного интеллекта и машинного обучения для обработки и анализа текстов на естественном языке, для изображений, речи, биомедицины и других специальных видов данных;

п.14. Методы и средства формирования массивов условно-реальных данных и прецедентов, необходимых для решения задач искусственного интеллекта и машинного обучения.

Теоретическая и практическая значимость. Теоретическая значимость работы заключается в развитии подходов к совершенствованию архитектур ГНС трансформерного типа, направленных на решение задач восстановления изображений, на основе структурной регуляризации механизмов внимания. Проведенные в этом плане теоретические обоснования и доказательства носят достаточно общий характер и могут быть использованы для построения ГНС с различными вариантами реализации механизма внимания при решении других задач, в том числе, задач классификации и семантической сегментации. Предложенные модели и алгоритмы аугментации позволяют повысить обобщающую способность моделей в условиях ограниченности данных, наличия сложных искажений и помех для различных задач компьютерного зрения.

Представленные теоретические и экспериментальные результаты позволяют проводить сравнительный анализ альтернативных подходов к построению алгоритмов обработки информации рассматриваемого класса и выбор конкретного алгоритма с учетом возникающих на практике ограничений.

Практическая значимость обусловлена возможностью внедрения разработанных алгоритмов в прикладные системы компьютерного зрения, включая: автоматические системы видеонаблюдения и мониторинга; аэрокосмическую съемку в сложных погодных условиях; медицинскую томографию (для подавления шумов и артефактов на снимках); системы обработки изображений в мобильных устройствах.

Алгоритмы аугментации, в частности, предложенная в работе архитектура WeatherTransformer позволяют генерировать синтетические обучающие данные, моделирующие реальные условия съемки без необходимости их ручного сбора, что особенно актуально при обучении нейронных сетей на малых или несбалансированных выборках данных.

Результаты работы использованы при выполнении в ФГБОУ ВО «Воронежский государственный университет» научно-исследовательских работ в период 2022-2025 годов,

связанных с обработкой изображений специального назначения (НИЧ-21009, НИЧ-23019), в которых автор являлся непосредственным исполнителем, а также в учебном процессе вуза.

Положения и результаты, выносимые на защиту. На защиту выносятся следующие результаты и положения.

1. Структурная регуляризации механизма внимания в трансформерных блоках нейронных сетей может осуществляться путем внесения мультипликативной или эквивалентной ей аддитивной стохастической составляющей при вычислении весов внимания, что проявляется в сглаживании соотношения весов внимания для снижения возможности их неконтролируемого роста в процессе обучения.

2. Применение отдельно обучаемой матрицы масштабных коэффициентов в качестве мультипликативной составляющей при вычислении матриц внимания в рамках стандартного механизма позволяет снижать воздействие возникающих аномалий в виде существенно превалирующих весов внимания, ситуаций насыщения активационной функции и включает дополнительные возможности регулирования весовых коэффициентов внимания.

3. Повышение качества восстановления изображений в стандартных архитектурах глубоких нейронных сетей может быть достигнуто на основе предложенных модификаций, реализующих добавление аддитивной стохастической составляющей в виде выборочных оценок дисперсионных характеристик признаков, вычисляемых в предшествующих сверточных слоях, и обучаемых матриц масштабных коэффициентов.

4. Снижение вычислительной сложности модулей внимания в трансформерах без существенных потерь качества восстановления изображений может быть достигнуто за счет использования предложенного алгоритма канального сжатия, что позволяет эффективно учитывать как пространственные, так и каналные зависимости, а также масштабировать архитектуру сети для входных изображений высокого разрешения.

5. Эффективный и малозатратный подход к аугментации изображений с целью учета факторов, негативных для восприятия сцен и, прежде всего, атмосферных осадков (дождь, снег, туман), может быть достигнут за счет их переноса из эталонного и включения в модифицируемое изображение на основе использования предложенной модели двухвходового трансформера, объединяющего сверточный энкодер-декодер и перекрестный механизм внимания.

Степень достоверности результатов работы. Результаты исследований, сформулированные в диссертационной работе, основаны на теоретических и экспериментальных методах исследований, взаимно дополняющих друг друга, и согласуются между собой. Указанные результаты получены с использованием комплекса теоретических, вычислительных и экспериментальных методов. Проведенные исследования основываются на строго формализованных постановках задач, апробированных подходах глубокого обучения и алгоритмах обработки изображений. Все разработанные модели, механизмы внимания и алгоритмы аугментации подвергались тестированию в контролируемых условиях на синтетических и реальных датасетах с использованием общепринятых метрик качества.

Корректность и воспроизводимость синтезированных архитектур трансформеров, модификаций внимания и генеративных моделей подтверждается результатами многочисленных вычислительных экспериментов, сопоставлением с базовыми методами, а также статистической обработкой результатов. Выводы, сделанные в работе, имеют обоснованную интерпретацию, совпадают в ряде частных случаев с результатами, полученными другими авторами, и согласуются с общепринятыми теориями в области машинного обучения и обработки изображений. Таким образом, полученные в ходе диссертационной работы результаты можно считать в достаточной степени обоснованными, достоверными и практически значимыми.

Апробация работы. Основные положения, выводы и рекомендации, сформулированные в диссертации, докладывались и обсуждались на ряде научных конференций различного уровня. В частности, результаты были представлены:

- на XXIII, XXIV и XXV Международных конференциях «Информатика: проблемы, методология, технологии» (г. Воронеж, 2023–2025 гг.);
- на 5-й Международной конференции «International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)» (г. Липецк, 2023 г.);
- на ежегодных научных сессиях факультета компьютерных наук ВГУ.

Публикации. По теме диссертационной работы опубликовано 8 научных работ, из них 4 статьи в изданиях, рекомендованных ВАК и 1 статья в материалах конференции, представленной в IEEE Explore (Scopus), получено 1 свидетельство о государственной регистрации программы для ЭВМ.

Личный вклад. Все выносимые на защиту результаты и положения принадлежат лично автору. В публикациях, выполненных в соавторстве с руководителем, последнему принадлежат постановка задачи и выбор направления исследований. Непосредственно соискателю принадлежат: обоснование предложенных архитектур и моделей, разработка алгоритмов, реализация программных прототипов, постановка и проведение экспериментов, анализ и интерпретация результатов.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения и списка литературы. Объем работы составляет 161 страницу основного текста, включая два приложения, 28 рисунков и 16 таблиц. Список использованных источников содержит 108 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, сформулированы цель и задачи работы, показаны научная новизна, теоретическая и практическая значимость, описаны методы исследования, результаты, выносимые на защиту, степень их достоверности и апробации.

В первой главе проводится анализ современных методов и алгоритмов восстановления и аугментации изображений. Рассматриваются классические подходы к восстановлению (линейная и нелинейная фильтрация, частотные преобразования) и современные нейросетевые методы, включая сверточные архитектуры и трансформерные модели. Значительное внимание уделяется алгоритмам аугментации изображений как эвристическим, так и генеративным моделям (GAN, VAE, диффузионные модели). Отдельно выделяется направление синтеза изображений с моделированием погодных эффектов, рассматриваемых как фактор, затрудняющий восприятие сцен нейронными сетями. Анализ существующих методов показал ограниченность классических алгоритмов при наличии сложных и пространственно-неоднородных искажений, а также выявил высокую зависимость нейросетевых подходов от качества и объема обучающих данных. Это послужило обоснованием необходимости комплексного подхода, объединяющего методы восстановления и синтеза изображений в рамках единой системы обработки информации.

В главе предлагается общая схема проведения исследований, направленная на создание алгоритмов восстановления изображений и аугментации данных, состоящая из следующих этапов.

1. Исследование и модификация механизмов внимания в трансформерных архитектурах для задач восстановления изображений в условиях сложных искажений.
2. Разработка комплекса алгоритмов аугментации данных на основе как эвристических методов, так и генеративных моделей, включая синтез изображений с погодными эффектами.
3. Интеграция алгоритмов восстановления и аугментации в универсальную схему обработки изображений, устойчивую к шумам и способную увеличивать набор обучающих данных без ручного их сбора.
4. Экспериментальное сравнение синтезированных архитектур с современными аналогами по стандартным и перцептивным метрикам, а также оценка их работоспособности на реальных данных.

В конце главы сформулированы выводы о необходимости комплексного подхода, объединяющего нейросетевые методы восстановления и целенаправленную аугментацию

данных как основу для построения эффективных алгоритмов компьютерного зрения в условиях ограниченности обучающих выборок и вариативности сложных шумовых воздействий.

Во второй главе обосновываются и исследуются возможные способы модификации механизма внимания (самовнимания, self-attention) в трансформерах для задач восстановления изображений. Подробно рассматриваются особенности вычисления механизма в тензорном потоке и его взаимосвязь с типами искажений изображения. Предлагаются и теоретически обосновываются способы регуляризации весов внимания на этапе обучения для повышения устойчивости моделей к шумам и пространственно-неоднородным артефактам.

В классической архитектуре трансформера входное изображение разбивается на патчи, которые после линейной проекции представляются в виде последовательности векторов признаков. Для каждого патча вычисляются три матрицы: запросов, ключей и значений соответственно: $Q \in \mathbf{R}^{n \times d}$, $K \in \mathbf{R}^{n \times d}$, $V \in \mathbf{R}^{n \times v}$, где n – число пикселей изображения, используемое для расчета матрицы внимания, а d – размер признакового пространства. Тогда можно вычислить взвешенные механизмом внимания выходы следующим образом:

$$O = \text{attn}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d}} \right) V \in \mathbf{R}^{n \times v},$$

$$Q = \begin{pmatrix} q_1^{(1)} & \dots & q_d^{(1)} \\ \vdots & \vdots & \vdots \\ q_1^{(n)} & \dots & q_d^{(n)} \end{pmatrix}, K = \begin{pmatrix} k_1^{(1)} & \dots & k_d^{(1)} \\ \vdots & \vdots & \vdots \\ k_1^{(n)} & \dots & k_d^{(n)} \end{pmatrix}, V = \begin{pmatrix} v_1^{(1)} & \dots & v_v^{(1)} \\ \vdots & \vdots & \vdots \\ v_1^{(n)} & \dots & v_v^{(n)} \end{pmatrix}$$

Для задачи восстановления изображений также применяются многоголовые варианты self-attention, что позволяет учитывать как локальные, так и глобальные зависимости.

Одним из предложенных в работе подходов является внесение стохастических составляющих в матрицу весов внимания для снижения эффектов переобучения и сглаживания распределения весов с целью предотвращения их неконтролируемого роста. Рассмотрены два базовых варианта: мультипликативная и аддитивная регуляризация.

В работе доказана следующая **основная теорема о внесении стохастической составляющей** (подробное доказательство приведено в статье автора [1]). Пусть имеется дискретная случайная величина $\mathbf{h}^{(i)}$ с распределением $P(h)$ на интервале $[0,1]$, вносящая независимое случайное воздействие на элементы строки выхода $o^{(i)}$ взвешенной функцией softmax матрицы внимания в виде:

$$o_j^{(i)} = \sum_{p=1}^n \left(\frac{h_p^{(i)} \alpha_p^{(i)} v_j^{(p)}}{\sum_{k=1}^n h_k^{(i)} \alpha_k^{(i)}} \right), o_j^{(i)} = \alpha^{(i)} v_j = \sum_{p=1}^n \alpha_p^{(i)} v_j^{(p)}, i = \overline{1, n}, j = \overline{1, v}.$$

$$O = \begin{pmatrix} o_1^{(1)} & \dots & o_v^{(1)} \\ \vdots & \vdots & \vdots \\ o_1^{(n)} & \dots & o_v^{(n)} \end{pmatrix} = AV = \begin{pmatrix} \alpha_1^{(1)} & \dots & \alpha_n^{(1)} \\ \vdots & \vdots & \vdots \\ \alpha_1^{(n)} & \dots & \alpha_n^{(n)} \end{pmatrix} \begin{pmatrix} v_1^{(1)} & \dots & v_v^{(1)} \\ \vdots & \vdots & \vdots \\ v_1^{(n)} & \dots & v_v^{(n)} \end{pmatrix}$$

где O – матрица выходов слоя самовнимания, каждая строка которой соответствует своему пикселю (после линейной проекции), в которой осуществляется взвешивание компонентов с одинаковыми номерами (j) для всех пикселей с номерами $i = \overline{1, n}$.

Пусть также исходные веса внимания в любых двух позициях имеют соотношение $\alpha_t^{(i)} > \alpha_s^{(i)}$. Тогда соотношение условных математических ожиданий величин $h_t^{(i)}, h_s^{(i)}$, осуществляющих коррекцию исходных весов внимания определяется неравенством:

$$c_t^{(i)} = M_{a,h} \left[\frac{h_t^{(i)}}{\sum_{k=1}^n h_k^{(i)} \alpha_k^{(i)}} \right] < c_s^{(i)} = M_{a,h} \left[\frac{h_s^{(i)}}{\sum_{k=1}^n h_k^{(i)} \alpha_k^{(i)}} \right].$$

Отсюда следует, что внесение мультипликативной стохастической составляющей обладает сглаживающим эффектом по отношению к неконтролируемому росту весовых коэффициентам:

$$\hat{o}_j^{(i)} = M_a[o_j^{(i)}] = M_h \left[\sum_{p=1}^n \left(\frac{h_p^{(i)} \alpha_p^{(i)} v_j^{(p)}}{\sum_{k=1}^n h_k^{(i)} \alpha_k^{(i)}} \right) \right] = \sum_{p=1}^n c_p^{(i)} \alpha_p^{(i)} v_j^{(p)}.$$

Следствие 1. Аналогичные неравенства могут быть получены для случая, когда используются значения случайной величины $\mathbf{h}^{(i)} \in [0,1]$, принадлежащие множеству континуум, для которой задана непрерывная плотность распределения $P_h(h_p^{(i)})$, $p = \overline{1, n}$.

Следствие 2. Как частный случай рассматривается классический метод DropKey, где для i -й строки выхода стохастическая составляющая: $h_p^{(i)} = \{1, \Pr(\mathbf{h}^{(i)} = 1) = d_{key}, h_p^{(i)} = \{0, \Pr(\mathbf{h}^{(i)} = 0) = 1 - d_{key}$ осуществляет исключение весов внимания в соответствии с законом Бернулли.

Следствие 3. Пусть при вычислении самовнимания осуществляется добавление аддитивной составляющей $h_p^{(i)} \geq 0$, которая представляет значение случайной величины $\mathbf{h}^{(i)}$, вносящей независимое случайное воздействие на элементы строки матрицы внимания в позиции p перед выполнением нормировки в softmax. Пусть для определенности $\mathbf{h}^{(i)}$ является случайной величиной, заданной на интервале $[0,1]$ с распределением $P(h) = P_h(h_p^{(i)})$, $p = \overline{1, n}$, $i = \overline{1, n}$. Пусть также исходные веса внимания в двух позициях (веса до внесения аддитивной составляющей) имеют соотношение $\alpha_t^{(i)} > \alpha_s^{(i)}$. Оценим в этих условиях соотношение условных математических ожиданий величин $h_t^{(i)}, h_s^{(i)}$, осуществляющих коррекцию исходных весов внимания. Исходя из свойств экспоненциальной функции, для весов $\tilde{\alpha}_p^{(i)}$ имеет место соотношение:

$$\tilde{\alpha}_p^{(i)} = \frac{\exp(h_p^{(i)}) \exp(q^{(i)} k_p / \sqrt{d})}{\sum_{p=1}^n \exp(h_p^{(i)}) \exp(q^{(i)} k_p / \sqrt{d})} = \frac{\tilde{h}_p^{(i)} \alpha_p^{(i)}}{\sum_{p=1}^n \tilde{h}_p^{(i)} \alpha_p^{(i)}}, \quad \tilde{h}_p^{(i)} = \exp(h_p^{(i)}).$$

Таким образом, полученное воздействие сводится к ранее рассмотренному мультипликативному.

Еще один подход к модернизации механизма SA, предлагаемый и обоснованный в диссертации, основан на введении отдельно обучаемой матрицы S для регулирования масштабных коэффициентов скалярных произведений, вычисляемых до преобразования softmax. Обучение элементов этой матрицы и весов внимания должно осуществляться отдельно. Математически подобное преобразование можно представить путем операции поэлементного перемножения матриц следующим образом:

$$R_s = S \otimes \frac{1}{\sqrt{d}} QK^T.$$

В диссертации и статье автора [1] показано, что в рамках стандартного механизма такое преобразование позволяет снижать воздействие возникающих аномалий в виде существенно превалирующих элементов самовнимания и включает дополнительные возможности регулирования весовых коэффициентов внимания в ситуации насыщения активационной функции.

Помимо этого, в работе рассмотрен подход, основанный на использовании для аддитивной коррекции стандартной меры внимания на величину относительной

среднеквадратичной ошибки, оптимальной в классе линейных оценок $\tilde{x}_{j|i} = \tilde{x}_j(x_i) = \tilde{M}[x_j / x_i] = r_{j,i}x_i$, $r_{j,i} = r_{i,j}$ пикселя x_i относительно пикселя x_j , которая имеет вид:

$$\delta_{j|i} = \frac{(x_i - r_{i,j}x_j)^2 + \varepsilon}{(\sigma^2 - \sigma^2 r_{i,j}^2) + \varepsilon},$$

где $r_{i,j}$ – заданное или обучаемое значение коэффициента корреляции одного элемента изображения относительно другого для однородного случайного поля; ε – малое положительное значение, неспособное повлиять на результаты вычислений. Здесь для коррекции внимания фиксируется нарушение степени линейной связи между пикселями, свойственное реальным изображениям.

В третьей главе осуществляется синтез и исследование алгоритмов восстановления изображений на базе моделей-трансформеров с усовершенствованными механизмами внимания. Предлагаются две архитектуры: с модифицированным канальным вниманием и с пространственным вниманием. Приводятся результаты численных экспериментов, демонстрирующие преимущество предложенных моделей по стандартным метрикам PSNR, SSIM и FID, а также по вычислительной сложности по сравнению с базовыми архитектурами.

Для архитектуры трансформера с модифицированным канальным вниманием был разработан модуль CBSA (Channel Bottleneck Self-Attention) (рис 1), реализующий механизм сжатия канальной информации для снижения вычислительной сложности.

В CBSA каждый нормализованный тензор I признаков размерами $H \times W \times C$ проходит следующие этапы обработки.

Этап 1: Канальное сжатие и разделение на Q, K, V . Применяется точечная свертка (с ядром 1×1) для извлечения объединенного тензора \tilde{X} :

$$\tilde{X} = \text{Conv}_{1 \times 1}(I), \quad \tilde{X} \in \mathbf{R}^{H \times W \times 3C'}, \quad C' = \frac{C}{L}$$

Эта операция реализует как проекцию на Q, K, V , так и одновременно уменьшает число каналов в L раз. Далее применяется depthwise-свертка:

$$\tilde{X}^{dw} = \text{DWConv}_{3 \times 3}(\tilde{X}),$$

где каждая из $3C'$ компонент обрабатывается независимым 3×3 фильтром. После чего производится разбиение тензора по каналам на Q, K, V :

$$\tilde{X}^{dw} = [Q', K', V'], \quad Q', K', V' \in \mathbf{R}^{H \times W \times C'}.$$

Этап 2: Преобразование в матричную форму. Для применения механизма SA, тензоры преобразуются в матрицы:

$$Q, K, V \in \mathbf{R}^{h \times d \times N}, \quad N = H \cdot W, \quad d = \frac{C'}{h},$$

где h – число голов внимания, а d – размерность канального подпространства для одной головы.

Этап 3: Механизм самовнимания. В классической постановке механизм self-attention формализуется ранее приведенными соотношениями, но в формулах ниже он расписан для нескольких голов внимания, при этом переставлены местами канальная и пространственная размерности.

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right), \quad A \in \mathbf{R}^{h \times d \times d}.$$

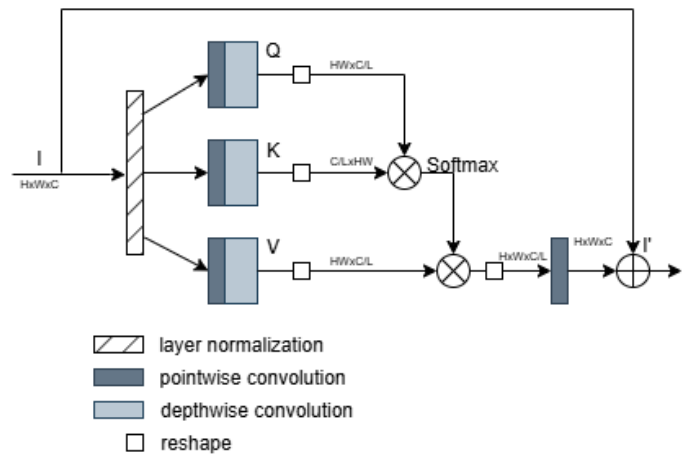


Рисунок 1 – Модуль CBSA

Полученные веса используются для взвешивания значений:

$$O = AV \in \mathbf{R}^{h \times d \times N}.$$

Этап 4: Обратное преобразование и восстановление каналов. Происходит преобразование результата O в пространственный тензор: $Z' \in \mathbf{R}^{H \times W \times C'}$, после чего применяется pointwise-свертка, вместо классических полносвязных слоев с увеличением канальной информации в L раз: $Z = \text{Conv}_{|x|}(Z') \in \mathbf{R}^{H \times W \times C}$.

Изначально вычислительную сложность глобального механизма внимания для двух операций матричного умножения без учета голов внимания можно рассчитать следующим образом: $f(N) = O(N^2C + NC^2)$. В результате приведенных преобразований с коэффициентом сжатия L итоговая вычислительная сложность становится равной:

$$f_r(N) = O(N^2C / L + NC^2 / L),$$

т.е. происходит уменьшение вычислительной сложности механизма самовнимания в L раз.

Общий вид реализованной модели ГНС с внедренным модулем CBSA представлен на рис.2. Анализ полученных результатов показал, что предложенная модель CBSA достигает качества прототипа (модели Restormer) при меньшем числе параметров (~1.5x меньше).

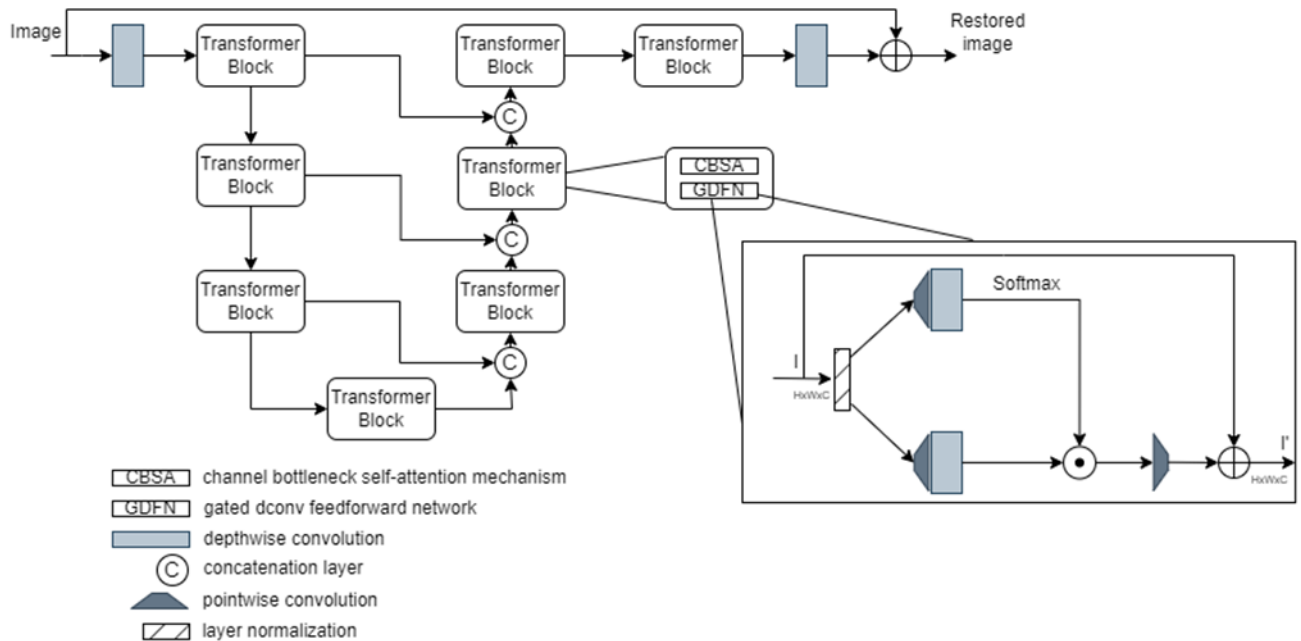


Рисунок 2 – Архитектура модели-трансформер со сжатием канального внимания

Варьирование параметра L позволяет подбирать целесообразные значения, обеспечивающие уменьшение вычислительной сложности при незначительном снижении качества восстановления изображений, что показано в таблице 1 на примере обработки изображений, полученных в плохих погодных условиях. Проведенные эксперименты на датасетах SIDD, ImageNet и погодных изображениях показали, что оптимальным выбором является $L = 2$, обеспечивающий баланс между качеством и скоростью обучения.

Таблица 1 – Качество восстановления изображения в зависимости от коэффициента L

Величина сжатия	PSNR	SSIM	FID
$L=1$	26,6	0,93	74,42
$L=2$	26,6	0,92	74,75
$L=4$	26,5	0,92	75,46
$L=8$	26,0	0,90	76,46

Вторая предлагаемая модель основана на базовой архитектуре SwinIR с локальным механизмом внимания, модифицированным различными методами структурной регуляризации, обоснованными в главе 2. На рис 3 показана схема внесенных архитектурных изменений.

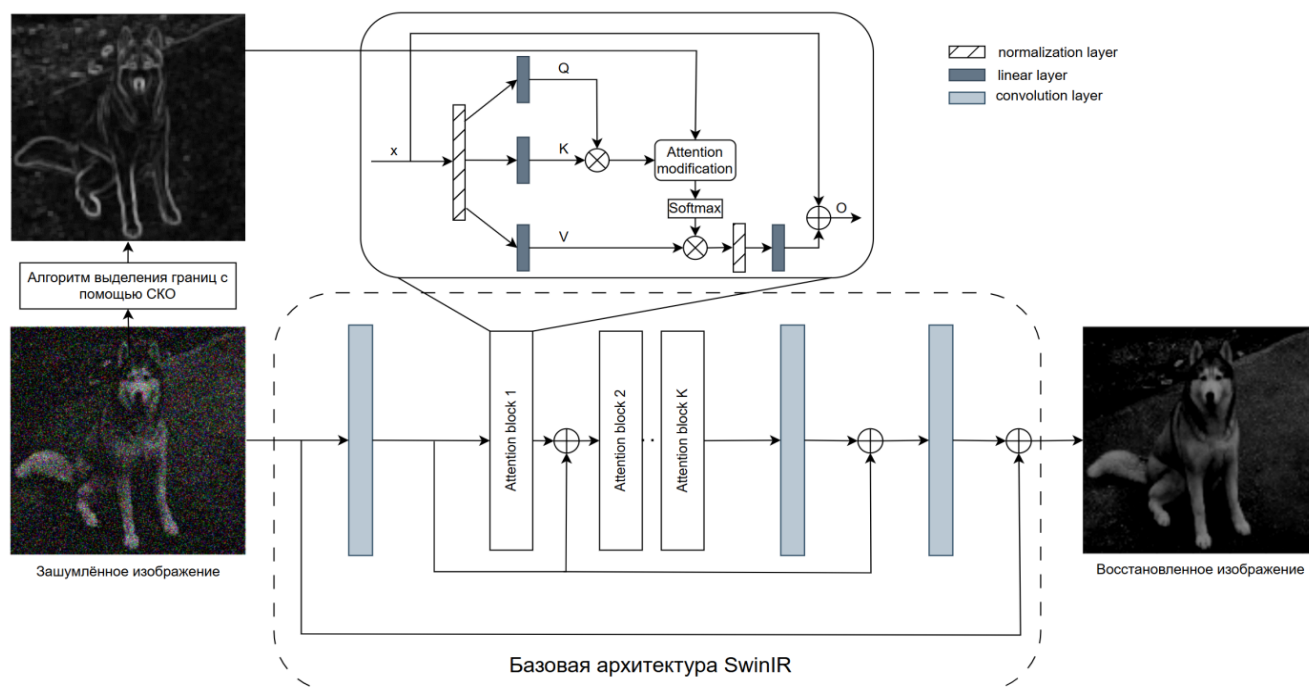


Рисунок 3 – Архитектура модели с модификацией пространственного внимания

Установлено, что в качестве стохастических составляющих, вносимых в механизм самовнимания, целесообразно использовать локальные выборочные оценки дисперсионных и корреляционных характеристик пространственных карт входных признаков слоя. При этом все описанные выше теоретические обоснования для структурной регуляризации остаются действенными.

В работе предлагается провести дисперсионный анализ изображения и построить добавочный тензор D для корректировки механизма внимания. Выполняются следующие действия.

- 1) Проводится отображение цветного изображения I в формат оттенков серого.
- 2) Выполняется скользящее сканирование изображения I_c окном размером $k \times k$ с единичным шагом и добавлением соответствующего паддинга для сохранения размерности.
- 3) Оценивается выборочная дисперсия в скользящих окнах $k \times k$.
- 4) Для полученной матрицы $I_{std} = (\sigma_{ij})$ с целью устранения выбросов и уменьшения влияния шумов выполняется гауссовское размытие с маской 3×3 .
- 5) Формируется тензор $D = I_{sp} \cdot I_{sp}^T$ для внесения в механизм SA.

Полученный тензор D используется при реализации двух вариантов регуляризации: внесение аддитивной стохастической составляющей к матрице скалярных произведений в виде дисперсий окрестности исходных пикселей; использование подхода DropKey со случайным удалением весов внимания, причем порог для удаления принимается с учетом информации об резких и текстурных областях, содержащихся в элементах D .

В таблице 2. представлены результаты тестирования для двух наборов обучающих данных: ImageNet и погодного датасета.

Таблица 2 – Результаты экспериментов для модификаций механизма самовнимания

N	Использованные модификации самовнимания	Подмножество ImageNet		Погодный датасет	
		PSNR	SSIM	PSNR	SSIM
1	Добавление матрицы выборочных дисперсий окрестности исходных пикселей	41,12	0,96	29,6	0,92
2	Адаптивный вариант DropKey	40,45	0,96	29,6	0,93
3	Стандартный DropKey	40,20	0,93	28,0	0,90
4	Адаптивный вариант DropKey + 2000 итераций обучения с отключением	41,20	0,97	29,9	0,93
5	Регулирование путем умножения на обучаемую матрицу S	40,17	0,94	29,2	0,93
6	Добавление матрицы выборочных дисперсий и регулирование путем умножения на S	41,04	0,97	29,5	0,91
7	Аддитивная коррекция на основе коэффициента корреляции	41,00	0,94	28,3	0,89
8	Исходная архитектура SwinIR	38.18	0,90	25,0	0,85

Наилучшие показатели достигнуты при комбинации адаптивного DropKey и поэтапного отключения регуляризации на финальных итерациях обучения. Для объективной оценки все предложенные модели были сопоставлены с оригинальными Restormer, SwinIR, SUNet и MIRNet. Все сети обучались на одних и тех же наборах данных с идентичными параметрами аугментации. Как следует из таблицы 2, архитектура с модифицированным пространственным вниманием во всех вариантах обеспечивает значимый прирост PSNR и SSIM на текстурных и зашумленных изображениях по сравнению с исходным прототипом (SwinIR).

В четвертой главе рассматриваются алгоритмы генерации и стилизации изображений с целью аугментации. Приводится систематизация возможных подходов: от простых эвристик до сложных генеративных архитектур. Особое внимание уделено исследованию предложенной новой модели *WeatherTransformer*, реализующей синтез атмосферных осадков с использованием модели-трансформера и перекрестного внимания.

Рассматриваются задачи: моделирование шумов и помех аддитивного и аппликативного типа, синтез реалистичных погодных условий (дождь, снег, туман), частичная стилизация изображений с использованием слоев AdaIN. Разработана модель *WeatherTransformer* с механизмом перекрестного внимания (cross-attention), позволяющая накладывать атмосферные эффекты на исходные изображения. Количественные результаты синтеза в системе выбранных показателей представлены в таблице 3.

Таблица 3 – Синтез изображений для различных погодных условий

Тип погодного осадка	PSNR	SSIM	FID
Все погодные осадки	21.88	0.78	30.47
Только дождь	22.43	0.81	29.68
Только снег	22.5	0.81	29.70
Только туман	23.34	0.82	29.23

Структурная схема модели трансформера показана на рис 4.

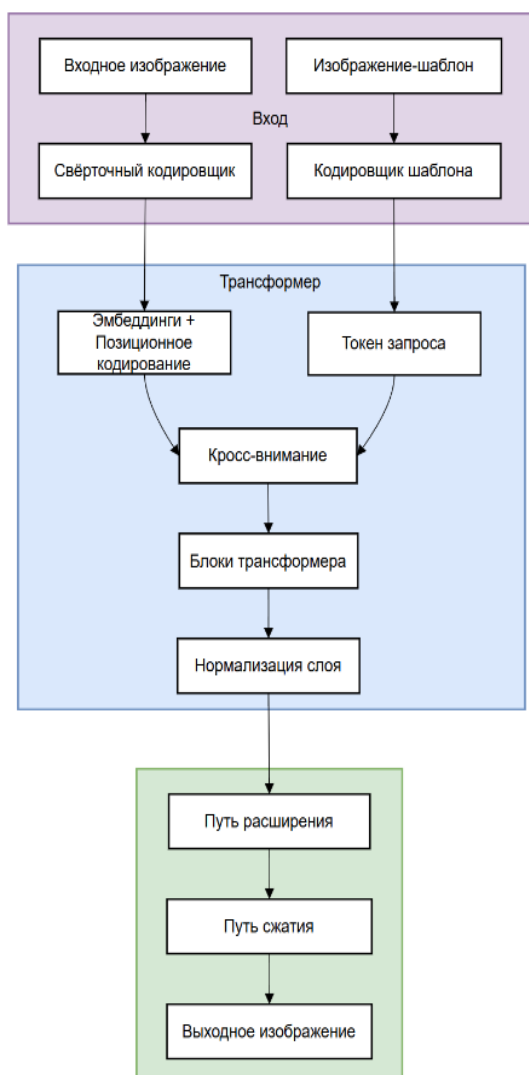


Рисунок 4 – Схема WeatherTransformer

Архитектура WeatherTransformer сочетает сильные стороны сверточных слоев (локальная инвариантность и иерархичность представлений) и трансформеров (глобальное внимание и адаптивность к контексту). Модель принимает на вход два изображения: исходное чистое изображение и изображение-шаблон с примером целевого погодного эффекта. На выходе формируется исходное изображение с включением наложенного эффекта погодного шаблона. Общая архитектура включает три основных блока: сверточный энкодер, трансформер с перекрестным вниманием и сверточный декодер. Кроме того, используется позиционное кодирование для сохранения пространственных связей между частями изображения. Для обучения модели WeatherTransformer в работе предложено использовать составную функцию потерь weatherLoss, учитывающую контентное, структурное и перцептуальное сходство (сходство восприятия):

$$L_{total} = 0.7L_{content} + 0.2L_{struct} + 0.1L_{perc} .$$

Пример синтезированных изображений представлен на рис. 5.

В конце главы описана методика применения разработанных алгоритмов в задачах сегментации, классификации и восстановления изображений, приводится структура программного комплекса восстановления и аугментации изображений (ПК ВИА), в который интегрированы все предложенные модели и алгоритмы.



Рисунок 5 – Примеры обработки изображений а) – исходные изображения, б) – изображения - шаблоны, в) – синтезированные изображения.

Общая архитектура комплекса и логика взаимодействия модулей показаны на рис 6. ПК ВИА объединяет модули восстановления, генерации искажений, синтеза погодных эффектов и анализа внимания в единую систему, а также реализует итеративную циклическую схему обучения. Это позволяет формировать синтетические пары изображений и прогрессивно улучшать качество обеих задач даже при ограниченном объеме размеченных данных. Предложена методика совместного применения разработанных средств для повышения устойчивости нейросетевых моделей к различным помехам в условиях нехватки обучающих данных. Экспериментально подтверждена эффективность использования синтезированных и искаженных изображений в качестве обучающих данных в задачах восстановления, классификации и сегментации в условиях ограниченных объемов обучающих данных.

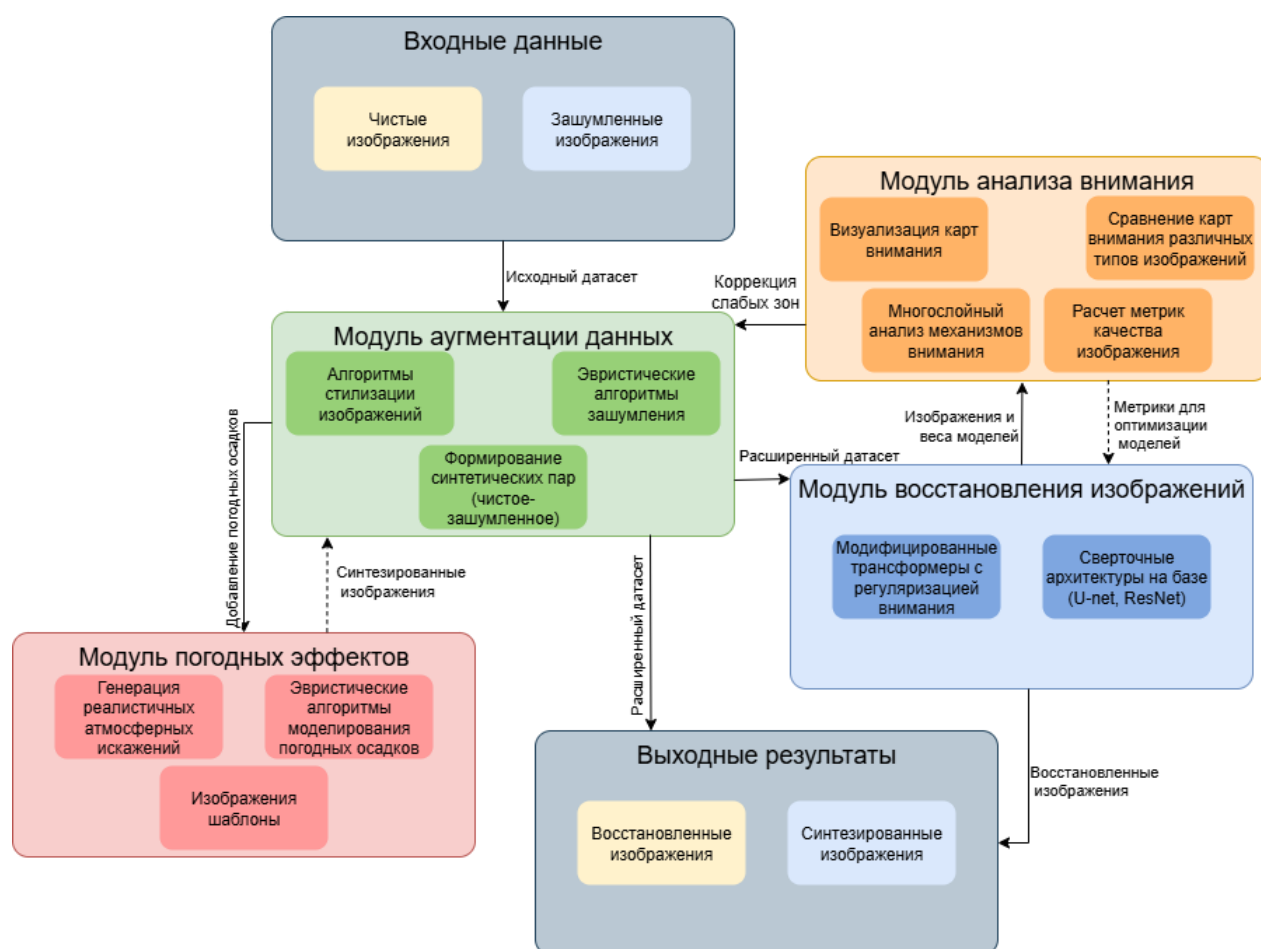


Рисунок 6 – Укрупненная структурная схема предлагаемого программного комплекса

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Проведен анализ современных подходов в задачах восстановления и аугментации изображений. Выявлены ограничения классических методов обработки изображений, обусловленные их высокой чувствительностью и зависимостью к выбору параметров восстановления. Установлено, что основным фактором, ограничивающим эффективность нейросетевых алгоритмов, является нехватка обучающих данных, что делает особенно важным использование продвинутых средств аугментации данных для расширения обучающих выборок и моделирования различных типов шумов и искажений.

2. Проанализированы современные методы аугментации данных. Выявлена необходимость использования комплексного подхода, объединяющего нейросетевые алгоритмы восстановления изображений с целенаправленной аугментацией данных для повышения качества изображений в условиях нехватки данных и высокой вариативности искажений.

3. Теоретически обоснованы новые подходы к структурной регуляризации механизма самовнимания с целью повышения устойчивости трансформерных сетей к переобучению. Предложен метод внесения стохастической составляющей в механизм внимания. Показано, что такая регуляризация сглаживает доминирующие весовые коэффициенты внимания и препятствует неконтролируемому росту отдельных весов в процессе обучения. Предложен и теоретически обоснован способ структурной регуляризации процесса обучения трансформеров, отличающийся использованием обучаемой матрицы масштабных коэффициентов. Показано, что применение подобной матрицы способствует выходу активационной функции, применяемой в механизме внимания, из области насыщения.

4. Разработаны обладающие новизной архитектуры ГНС трансформерного типа для задачи восстановления изображений, отличающиеся использованием канального механизма внимания со сжатием и уменьшением размерности внутреннего представления признаков, а также использованием модернизированного механизма пространственного внимания с внесением стохастической составляющей. Экспериментально показано, что при умеренном сжатии канальной информации качество восстанавливаемых изображений снижается незначительно, тогда как вычислительная сложность типового модуля внимания сокращается пропорционально коэффициенту сжатия. Для сети с модифицированным пространственным вниманием, опирающейся на базовую архитектуру SwinIR, в ходе экспериментов установлено, что по показателям качества внесенные различные варианты структурной регуляризации в рамках теоретически обоснованных принципов превосходят исходную модель.

5. Предложена оригинальная модель ГНС WeatherTransformer для добавления атмосферных осадков на изображения реальных сцен. Архитектура WeatherTransformer объединяет сверточный энкодер-декодер для извлечения и восстановления многоуровневых признаков с блоками перекрестного внимания, точно накладывающими шаблон погодного эффекта (дождь, снег, туман) на исходное изображение без искажения структуры сцены. При ее обучении введена специальная функция потерь, учитывающая содержательное и структурное сходство с исходным изображением, а также чувствительное сходство (ошибка восприятия) с эталонным образцом эффекта. Проведена экспериментальная оценка эффективности предложенной модели аугментации данных для различных задач компьютерного зрения. Аугментация с помощью разработанных алгоритмов позволила существенно повысить точность работы нейросетевых моделей.

6. Разработан программный комплекс ПК ВИА для восстановления и аугментации изображений, основанный на принципах циклического обучения, в котором модули восстановления и генерации последовательно улучшают друг друга на основе синтетических и реальных данных.

Рекомендации по использованию. Разработанные модели и алгоритмы улучшения качества изображений могут быть внедрены в различные системы компьютерного зрения и цифровой обработки изображений. Наиболее перспективными областями применения являются автоматизированные системы, требующие восстановления или повышения качества визуальных данных в присутствии шумов и помех: системы видеонаблюдения и безопасности (для очистки и детализации кадров); системы дистанционного аэрокосмического мониторинга (улучшение спутниковых снимков в сложных погодных условиях), медицинская визуализация (подавление артефактов и улучшение резкости снимков), мобильная фотография и постобработка изображений.

Предложенные алгоритмы аугментации данных могут использоваться для расширения тренировочных наборов при обучении нейронных сетей различных архитектур, что особенно важно в задачах, где сбор реальных зашумленных данных затруднен или объем данных ограничен. Генерация синтетических дождя или тумана на изображениях позволит подготовить систему компьютерного зрения к работе в реальных погодных условиях.

Перспективы дальнейшего развития. Результаты работы открывают ряд направлений для будущих исследований. Одним из таких направлений является дальнейшее совершенствование гибридных нейросетевых алгоритмов путем интеграции предложенных

механизмов внимания с другими современными архитектурными решениями и методами аугментации. Перспективным также является реализация подхода к совместному обучению нескольких нейронных сетей, задействованных на разных этапах восстановления и аугментации, на единых данных, что позволит более точно и согласованно настраивать их параметры, чтобы добиться лучшей производительности и эффективности.

Кроме того, важным направлением развития исследования представляется применение разработанных моделей и алгоритмов к обработке видеоданных (потокное видео). Также требуется более детальное рассмотрение вопросов регуляризации внимания и аугментации в динамических сценах, например, синтез и обработка помех с учетом их временной корреляции на соседних кадрах видеопоследовательности. Такой подход позволил бы эффективно восстанавливать видеоизображения при длительных или зависимых во времени искажениях.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК РФ

1. Бережнов, Н.И. Регуляризация механизма самовнимания в блоках трансформеров и ее применение в задачах классификации и восстановления изображений / Н.И. Бережнов, А.А. Сирота // Искусственный интеллект и принятие решений. – 2025. – №2. – С. 114-129.

2. Бережнов Н.И., Сирота А.А. Универсальный алгоритм повышения качества изображений с использованием глубоких нейронных сетей / Н.И. Бережнов, А.А. Сирота // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2022. – № 2. – С. 81-92.

3. Бережнов, Н.И. Модели глубокого обучения для синтеза изображений с включением атмосферных осадков с целью решения задач компьютерного зрения в различных погодных условиях / Н.И. Бережнов, А.А. Сирота // Вестник Воронежского гос. ун-та, Серия: Системный анализ и информационные технологии. – 2025. – № 2. – С. 89-104.

4. Бережнов, Н.И. Совершенствование механизмов внимания для архитектуры трансформер в задачах повышения качества изображений / Н.И. Бережнов, А.А. Сирота // Компьютерная оптика. – 2024. – Т. 48. – №. 5. – С. 726-733.

Публикации, индексированные в Scopus

5. Berezhnov, N.I. Understanding the attention mechanism in neural network transformer models in image restoration tasks / N.I. Berezhnov, A.A. Sirota // 5nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA). – Lipetsk: 2023. – P. 207-211.

Свидетельство о регистрации программы

6. Свидетельство о государственной регистрации программы для ЭВМ 2025682168 Российская Федерация. Программный комплекс для автоматизированного восстановления и аугментации графических данных / Н.И. Бережнов; заявитель и правообладатель Федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». – № 2025682168; заявление 07.08.2025; опублик. 21.08.2025.

Статьи и материалы конференций

7. Бережнов, Н.И. Исследование обобщающей способности методов глубокого обучения для улучшения качества изображений / Н.И. Бережнов, А.А. Сирота // XXIII Международная конференция «Информатика: проблемы, методология, технологии». – Воронеж: ИПЦ ВГУ: – 2023. – С. 510-518.

8. Бережнов, Н.И. Влияние априорной информации на механизм внимания в задаче улучшения качества изображений в моделях-трансформерах / Н.И. Бережнов, А.А. Сирота // XXIV Международная конференция «Информатика: проблемы, методология, технологии». – Воронеж: ИПЦ ВГУ: 2024. – С. 602-609.

9. Бережнов, Н.И. Модификации механизмов внимания в моделях-трансформерах в задаче восстановления изображений / Н.И. Бережнов // XXV Международная конференция «Информатика: проблемы, методология, технологии». – Воронеж: ИПЦ ВГУ: 2025. – С 483-491.