

УТВЕРЖДАЮ

Проректор по стратегическому развитию,
науче и инновациям

кандидат сельскохозяйственных наук,

доцент

Е.В. Скрипникова

« 25 » _____ 2025г.

ОТЗЫВ

ведущей организации на диссертацию Бережнова Никиты Игоревича
«Совершенствование механизмов внимания в глубоких нейронных сетях –
трансформерах в задачах восстановления и аугментации изображений»,
представленную на соискание ученой степени кандидата технических наук по
специальности 1.2.1. Искусственный интеллект и машинное обучение

Актуальность темы диссертации. Развитие технологий компьютерного зрения в настоящее время предъявляет повышенные требования к обработке визуальных данных, поступающих на вход автоматизированных систем анализа изображений. Практическое применение подобных систем в условиях реального функционирования часто сталкивается с проблемой низкого качества входных данных, обусловленного воздействием различных дестабилизирующих факторов: шумами различной природы, атмосферными явлениями, техническими несовершенствами регистрирующей аппаратуры, ошибками при передаче данных по каналам связи.

В таких условиях первостепенную важность приобретают методы предварительной обработки, направленные на восстановление качества визуальной информации и компенсацию искажений. Особую актуальность представляют алгоритмы, способные устранять нелинейные, пространственно-неоднородные и комбинированные типы искажений, для борьбы с которыми классические подходы на базе линейной и нелинейной фильтрации демонстрируют ограниченную эффективность.

Дополнительным фактором, ограничивающим эффективность нейросетевых подходов к восстановлению изображений, является нехватка качественных размеченных обучающих данных, представляющих достаточное разнообразие типов и уровней искажений. Формирование таких датасетов путем сбора реальных данных является трудоемким и зачастую технически затруднительным процессом. Решение данной проблемы связано с применением методов аугментации обучающих выборок, включающих как классические эвристические преобразования, так и генеративные модели, способные синтезировать реалистичные искаженные изображения.

Генеративные подходы на основе глубоких нейронных сетей открывают возможности для создания синтетических изображений, имитирующих широкий спектр искажений: различные типы шумов, атмосферные явления (дождь, снег, туман), цифровые артефакты. Интеграция процессов восстановления и аугментации в единую систему обработки позволяет повысить устойчивость и обобщающую способность разрабатываемых алгоритмов. Однако существующие генеративные модели сталкиваются с проблемами контролируемости синтеза, доменного смещения и вычислительной сложности обучения.

Таким образом, актуальность исследований, направленных на совершенствование архитектур трансформеров путем модификации механизмов внимания для снижения вычислительной сложности и повышения устойчивости к переобучению, а также на разработку эффективных методов аугментации данных для расширения обучающих выборок, не вызывает сомнений. Диссертационная работа Бережнова Н.И. посвящена решению указанного круга задач и представляется весьма актуальной как с научной, так и с практической точки зрения.

Структура диссертации. Представленная на отзыв диссертация «Совершенствование механизмов внимания в глубоких нейронных сетях – трансформерах в задачах восстановления и аугментации изображений» изложена на 161 странице; состоит из введения, четырех глав, заключения, двух приложений, списка использованных источников, включающего 108

наименований научных трудов на русском и иностранных языках; содержит 28 рисунков и 16 таблиц.

Научная новизна результатов работы определяется следующим.

1. Выполнен анализ современных подходов к решению задач восстановления и аугментации изображений. Показано, что классические методы обработки изображений обладают рядом ограничений, обусловленных высокой чувствительностью к выбору параметров алгоритмов, а основным фактором, ограничивающим эффективность нейросетевых алгоритмов, является недостаточный объем обучающих данных. Этот вывод подчеркивает важность применения методов аугментации для расширения обучающих выборок и моделирования различных типов шумов и искажений.

2. Предложены и теоретически обоснованы новые подходы к структурной регуляризации механизма самовнимания, направленные на повышение устойчивости трансформерных нейросетей к переобучению. В частности, разработан метод включения стохастической компоненты в процесс вычисления внимания и показано, что такая регуляризация сглаживает доминирующие веса внимания и предотвращает неконтролируемый рост отдельных коэффициентов в ходе обучения. Помимо этого, предложен и обоснован способ регуляризации обучения трансформеров на основе введения обучаемой матрицы масштабирования коэффициентов внимания. Установлено, что использование данного способа позволяет в ряде случаев вывести активационную функцию механизма внимания из состояния насыщения, в целом повышая эффективность обучения.

3. Разработаны новые архитектуры глубоких нейронных сетей трансформерного типа для задачи восстановления изображений. Предложенные модели отличаются использованием канального механизма внимания с принудительным сжатием внутреннего представления признаков, а также модернизированным механизмом пространственного внимания с добавлением стохастической составляющей. Для архитектуры с модифицированным пространственным вниманием (на базе базовой модели SwinIR) установлено, что различные варианты предложенной структурной регуляризации обеспечивают

более высокие показатели качества (PSNR, SSIM) по сравнению с исходной моделью.

4. Предложена оригинальная модель глубокого нейросети «WeatherTransformer», реализующая перенос и наложение атмосферных осадков на изображения реальных сцен. Архитектура WeatherTransformer сочетает сверточный энкодер-декодер с блоками перекрестного внимания, позволяющими накладывать шаблон погодного эффекта (дождь, снег, туман) на исходное изображение без искажения структуры сцены. Для обучения модели разработана составная функция потерь, учитывающая контентное и структурное сходство с оригиналом, а также перцептуальное сходство с эталонным образцом эффекта. Экспериментально показано, что применение данного метода аугментации данных существенно повышает точность работы нейросетевых моделей в различных прикладных задачах компьютерного зрения.

5. Создан программный комплекс для восстановления и аугментации изображений (ПК ВИА), реализующий принципы циклического (итеративного) обучения. В данном комплексе модули восстановления и генерации искажений чередуются и последовательно повышают качество работы друг друга на основе синтетических и реальных данных, что обеспечивает устойчивое улучшение результатов при дефиците исходных обучающих данных.

Степень обоснованности и достоверность результатов диссертации. Представленные в диссертации результаты и выводы, в целом, следует признать обоснованными и достоверными. Это подтверждается корректным сочетанием в работе взаимно дополняющих теоретических и экспериментальных методов исследования, строгой формализацией решаемых задач и использованием апробированных моделей глубокого обучения и алгоритмов обработки изображений. Все разработанные автором модели, механизмы внимания и алгоритмы аугментации реализованы и протестированы в контролируемых условиях на синтетических и реальных датасетах с использованием общепринятых метрик качества (таких как PSNR, SSIM и др.). Корректность и воспроизводимость полученных результатов подтверждается многочисленными вычислительными экспериментами, сопоставлением с базовыми методами, а

также статистической обработкой данных.

Сделанные в работе выводы имеют физическую интерпретацию, в ряде частных случаев совпадают с результатами других исследователей и не противоречат общепринятым теориям в области машинного обучения и цифровой обработки изображений. Таким образом, полученные результаты можно считать в достаточной степени обоснованными и надежными. Автор диссертации продемонстрировал уверенное владение современными методами разработки, обучения и анализа нейронных сетей (включая архитектуры трансформеров и генеративные модели), а также навыками программной реализации и экспериментального исследования алгоритмов обработки изображений.

Теоретическая и практическая значимость для соответствующей отрасли науки. Значимость результатов, полученных Н.И. Бережновым, для науки заключается в развитии подходов к совершенствованию архитектур ГНС трансформерного типа на основе структурной регуляризации механизмов внимания, что позволяет повысить их устойчивость к переобучению и улучшить качество восстановления изображений при ограниченных объемах обучающих данных.

В диссертационной работе поставлена и решена задача совершенствования механизма самовнимания с использованием обоснованных видов стохастических воздействий, которая позволяет осуществить тонкую настройку процесса обучения, максимально учитывая характер входных возмущений и особенности обрабатываемых изображений. Предложенные подходы отличаются своей универсальностью и могут быть применены к различным вариантам реализации механизма внимания при решении других задач, включая классификацию и семантическую сегментацию.

Разработанное семейство алгоритмов аугментации и синтеза изображений расширяет методологические возможности построения обучающих наборов данных для различных задач компьютерного зрения. Предложенная архитектура WeatherTransformer позволяет генерировать реалистичные изображения с атмосферными эффектами, что открывает новые возможности для моделирования сложных условий съемки.

Практическая ценность работы заключается в создании и внедрении инструментария, позволяющего решить одну из ключевых проблем разработки и применения моделей глубоких нейронных сетей для приложений компьютерного зрения – повышение их устойчивости к различным типам искажений и повышение эффективности при ограниченности обучающих данных. Разработанные алгоритмы и программный комплекс нашли практическое применение и использованы в научно-исследовательских работах, проводимых в ФГБОУ ВО «Воронежский государственный университет» в период 2022-2025 годов, связанных с обработкой изображений специального назначения.

Полученные в диссертации результаты в части разработки алгоритмов восстановления и аугментации изображений могут быть использованы в научно-исследовательских, проектных и промышленных организациях, занимающихся созданием систем компьютерного зрения, цифровой обработки изображений и интеллектуальными системами управления. Разработанные методы синтеза искажений являются особенно актуальными в ситуациях, когда сбор данных затруднен или экономически нецелесообразен, а формирование репрезентативных обучающих выборок критично для качества обучаемых моделей.

В дальнейшем результаты работы целесообразно использовать в организациях, осуществляющих фундаментальные и прикладные исследования в области компьютерного зрения и искусственного интеллекта. В частности, они могут быть применены в исследовательских проектах, выполняемых в: Белгородском государственном национальном исследовательском университете, Воронежском государственном университете, Федеральном исследовательском центре «Информатика и управление» РАН (г. Москва), Институте систем обработки изображений РАН (г. Самара), Санкт-Петербургском институте информатики и автоматизации РАН, Ярославском государственном университете им. П.Г. Демидова, Юго-Западном государственном университете (г. Курск).

Результаты работы также целесообразно использовать в образовательном процессе высших учебных заведений при подготовке бакалавров, магистров и специалистов в области искусственного интеллекта, машинного обучения и обработки изображений.

Соответствие критериям, которым должны отвечать диссертации на соискание ученых степеней.

Диссертация Бережнова Н.И. является завершенной научно-квалификационной работой, в которой содержится решение важной для компьютерного зрения и разработки систем искусственного интеллекта научной задачи совершенствования архитектур ГНС трансформерного типа на основе модификации механизмов внимания и разработки методов аугментации данных для повышения качества восстановления изображений (п.9).

Диссертация обладает внутренним единством, логически правильно выстроена, содержит новые научные результаты, полученные лично автором. Предложенные соискателем решения в достаточной степени аргументированы и сопоставлены с известными решениями, что позволяет объективно оценить их преимущества и ограничения (п.10).

Все научные результаты диссертации опубликованы в рецензируемых научных изданиях. Соискатель имеет 8 научных работ, в том числе 4 статьи в журналах, рекомендованных ВАК Российской Федерации, 1 статья в материалах конференции, представленной в IEEE Xplore и Scopus, а также 1 свидетельство о государственной регистрации программы для ЭВМ. Качество и количество публикаций соискателя по теме диссертации удовлетворяют требованиям ВАК (пп. 11, 13).

В диссертации имеются обязательные ссылки на источники заимствования материалов. В диссертации и автореферате указан вклад соискателя в работы, выполненные в соавторстве с научным руководителем (п.14).

Замечания. По сути рассмотренного, следует отметить несколько замечаний, которые, тем не менее, не влияют на общую положительную оценку работы:

1. Математическое обоснование целесообразности сглаживания весовых коэффициентов (стр.77-80) выполнено, на наш взгляд весьма приближенно. Фактически здесь анализируется оптимизационная задача общего вида и приводятся соотношения для производной целевой функции по обучаемым параметрам. Однако при этом не учитывается многослойность архитектуры модели, непонятно, какой вид целевая функция будет иметь в скрытых слоях

модели. Поэтому приведенное обоснование было бы желательно дополнить численным моделированием процесса обучения в указанном аспекте.

2. В работе предложен оригинальная модель WeatherTransformer для переноса погодных эффектов с одного изображения на другое. Данный подход сочетает сверточные и трансформерные слои и использует перекрестное внимание. Однако в литературе имеются и другие подходы к генерации погодных условий (например, генеративно-состязательные сети, диффузионные модели и пр.), которые автор рассматривает исключительно качественно. Было бы полезно включить в диссертацию экспериментальное сравнение предлагаемого метода с альтернативными моделями генерации погодных искажений по качеству и реалистичности синтеза. Это позволило бы показать преимущества выбранной архитектуры или выявить области, где она может уступать другим решениям.

3. Есть замечание по интерпретации количественных результатов. В третьей и четвертой главах приводится значительное число экспериментальных данных в виде таблиц и графиков (метрики PSNR, SSIM, FID и др.). Хотелось бы видеть более развернутый анализ этих результатов в тексте диссертации. Например, автору следовало бы детальнее прокомментировать, за счет чего достигается улучшение того или иного показателя, в каких случаях метод дает наибольший выигрыш, а где его эффективность несколько снижается. Также уместно было бы обсудить примеры неудачных случаев или ситуаций, в которых метод работает недостаточно хорошо (если таковые наблюдались) – это добавило бы объективности оценке предлагаемого подхода и могло бы наметить пути его дальнейшего совершенствования.

4. Вид некоторых математических выражений в автореферате может быть упрощен для улучшения читаемости, в частности, это касается представления формул для матриц трансформации признаков.

5. При описании приложений разработанного программного комплекса целесообразно было бы привести более подробное описание перспектив и специфики его использования в различных практико-ориентированных системах и очертить более четко круг возможно приложений, например, возможности применения при обработке медицинских изображений.

Заключение. Диссертационная работа Н.И. Бережнова «Совершенствование механизмов внимания в глубоких нейронных сетях – трансформерах в задачах восстановления и аугментации изображений» является научно-квалификационной работой, в которой содержится решение актуальной задачи повышения эффективности архитектур трансформеров на основе совершенствования механизмов внимания и разработки комплексных методов аугментации обучающих данных.

Диссертация подготовлена соискателем самостоятельно, обладает внутренним единством, содержит новые научные результаты, имеющие значение как для науки, так и для практических приложений. Работа оформлена в соответствии с требованиями, предъявляемыми к кандидатским диссертациям. Научные положения, выводы и рекомендации, сформулированные в диссертации, обоснованы и достоверны. Автореферат полностью и корректно отражает основные положения и выводы диссертации.

Диссертационная работа соответствует паспорту специальности 1.2.1. Искусственный интеллект и машинное обучение и охватывает следующие области исследования: 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».

Таким образом, с учетом изложенного выше, считаем, что диссертация на тему : «Совершенствование механизмов внимания в глубоких нейронных сетях – трансформерах в задачах восстановления и аугментации изображений» соответствует требованиям «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства РФ от 24.09.2013 №842 (ред. 16.10.2024), а ее автор, Никита Игоревич Бережнов заслуживает присуждения ученой степени кандидата технических наук по специальности 1.2.1. Искусственный интеллект и машинное обучение.

Составители отзыва дают согласие на использование своих персональных данных, включение их в документы, связанные с работой диссертационного совета, и их дальнейшую обработку.

Отзыв составил:

Профессор кафедры информационно-телекоммуникационных систем и технологий Института инженерных и цифровых технологий, доктор технических наук, профессор

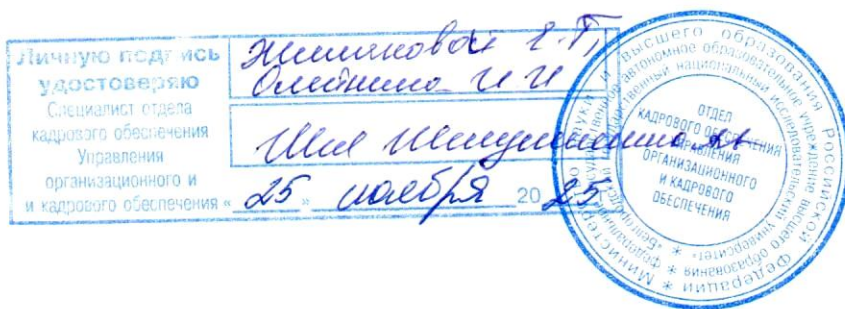
Жиляков Евгений Георгиевич

Утвержден:

на заседании кафедры информационно - телекоммуникационных систем и технологий протокол № 5 от «20» ноября 2025 г.

Заведующий кафедрой информационно- телекоммуникационных систем и технологий кандидат технических наук, доцент

Олейник Иван Иванович



Федеральное государственное автономное образовательное учреждение высшего образования «Белгородский государственный национальный исследовательский университет»

Почтовый адрес: 308015, г. Белгород, ул. Победы, 85

Телефон: +7 (4722) 30-12-11

E-mail: info@bsuedu.ru, Web-сайт: http://www.bsuedu.ru/