

ОТЗЫВ

официального оппонента доктора технических наук Шипко Владимира Вацлавовича на диссертацию Бережного 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. В последние годы задачи улучшения качества изображений и синтеза реалистичных данных приобрели большую значимость в связи с широким распространением автоматических и автоматизированных систем компьютерного зрения. Эффективность таких систем для видеонаблюдения, дистанционного зондирования, медицинской диагностики во многом определяется качеством входных изображений. Аппаратные шумы, артефакты и сложные искажения, включая атмосферные явления (дождь, снег, туман), ухудшают восприятие данных нейронными сетями и снижают точность аналитических задач. Классические методы обработки, линейные фильтры, различные схемы нелинейных преобразований, как правило, не обеспечивают необходимого уровня качества в условиях наличия сложных искажений, что требует разработки более эффективных и универсальных подходов. При этом современные алгоритмы глубокого обучения и, в частности сверточные и трансформерные сети с механизмами внимания, демонстрируют высокую эффективность в задачах компьютерного зрения. В таком контексте исследование новых архитектур трансформеров с улучшенным механизмом внимания и методов целевой аугментации изображений является важной научной задачей.

В силу изложенного, тематика диссертационного исследования Бережного Н.И., направленного на разработку и анализ механизмов внимания в нейросетевых архитектурах с расширением обучающих выборок посредством аугментации, представляется безусловно *актуальной* в контексте современных задач искусственного интеллекта и обработки визуальной информации.

Степень обоснованности и достоверности научных положений, выводов и рекомендаций, сформулированных в диссертации.

Основные положения диссертации подтверждены всесторонним теоретическим обоснованием и экспериментальной верификацией. При разработке моделей и алгоритмов автор использовал современные подходы теории вероятностей, математической статистики и машинного обучения, адекватно применяя их к анализу архитектур нейросетей. Для каждого предлагаемого метода проведен сравнительный анализ с альтернативными решениями: результаты тщательно исследованы как качественно, так и количественно. Выводы о влиянии модификаций внимания и аугментации на качество обработки изображений основаны на детальном анализе экспериментальных данных. Полученные при этом закономерности в целом укладываются в общепринятые представления о поведении нейросетевых систем, что свидетельствует о высокой обоснованности диссертационных выводов.

Достоверность выносимых результатов диссертации подтверждается их понятной физической трактовкой, результатами применения взаимно дополняющих теоретических и экспериментальных методов исследования, их согласованностью, совпадением в ряде случаев с известными результатами. Для проверки адекватности разработанных моделей были проведены вычислительные эксперименты с использованием реальных и синтетических наборов изображений. При этом оценивалось качество восстановления и аугментации по стандартным метрикам (например, пиковое отношение сигнал/шум, чувствительные метрики и др.). Экспериментальные данные демонстрируют улучшение качества восстановления и синтеза изображений по сравнению с базовыми прототипами. Согласованность теоретических расчетов и наблюдаемых значений метрик подтверждает достоверность полученных результатов.

Научная новизна результатов диссертации, выносимых на защиту, заключается в следующем.

1. Предложен метод структурной регуляризации механизма внимания в трансформерных нейросетях, включающий добавление стохастических (мультипликативных и аддитивных) компонентов при вычислении матриц весовых коэффициентов внимания. Это позволяет

сглаживать распределение весов и предотвращает неуправляемый рост отдельных коэффициентов в процессе обучения, повышая устойчивость сети к аномальным возбуждениям. Его дополняет способ регуляризации процесса обучения путем введения в каждом слое обучаемой матрицы масштабирующих коэффициентов. Автором показано, что предложенный подход смягчает эффект перенасыщения функций активации в блоках внимания и улучшает динамику сходимости при обучении на сложных искаженных данных.

2. Впервые предложены модификации механизмов внимания с учетом размерности входных признаков трансформерных модулей. Так, каналное внимание реализуется с учетом сжатия канальных признаков, а пространственное внимание – с добавлением дополнительной априорной информации об изображении. Благодаря таким изменениям удалось одновременно повысить качество восстановления изображений и снизить вычислительную сложность механизма внимания. На основе этих идей разработаны новые архитектуры глубоких трансформеров, адаптированные для обработки графических данных.

3. Создана специализированная архитектура нейронной сети для аугментации изображений, моделирующая атмосферные эффекты. Сеть сочетает сверточный энкодер–декодер с перекрестным механизмом внимания, обеспечивая реалистичный синтез дождя, снега и тумана при сохранении структурной целостности сцены. Для обучения предложенной модели введена составная функция потерь, учитывающая различные аспекты качества генерируемых изображений в разных погодных условиях.

4. Разработан программный комплекс алгоритмов восстановления и аугментации изображений на основе интеграции модифицированных трансформеров с моделями синтеза помех. Предложена методика совместного применения этих алгоритмов для повышения устойчивости нейросетевых решений при нехватке реальных обучающих данных. Экспериментально показано, что использование синтетических изображений в качестве дополнительных обучающих примеров улучшает результаты при решении задач восстановления, классификации и сегментации изображений.

Теоретическая значимость диссертации заключается в развитии подходов к разработке архитектур глубоких трансформеров для задач обработки изображений. Предложенные механизмы регуляризации внимания расширяют базовые модели трансформеров, добавляя стохастические и масштабирующие компоненты, что обеспечивает более устойчивое обучение нейросетей. Выводы об эффективности таких модификаций имеют обобщенный характер и могут быть применены при конструировании сетей для других задач компьютерного зрения (например, классификации или семантической сегментации). Предложенные методы аугментации позволяют формировать реалистичные обучающие выборки с учетом сложных искажений, что повышает обобщающую способность моделей при ограниченности данных и наличии помех. Представленные теоретические и экспериментальные результаты открывают возможность сравнительного анализа альтернативных подходов и выбора конкретных алгоритмов с учетом практических ограничений.

Практическая значимость результатов диссертации обусловлена возможностью их внедрения в прикладные системы обработки изображений. Разработанные алгоритмы применимы в автоматических системах видеомониторинга и управления в сложных метеоусловиях, при аэрокосмической съемке для подавления атмосферных шумов, а также в медицинской визуализации для уменьшения артефактов на рентген- и КТ-снимках. Кроме того, предложенный подход к автоматической генерации синтетических изображений с погодными эффектами позволяет автоматически расширять тренировочные выборки реалистичными сценами без ручного сбора данных, что особенно актуально при обучении нейросетей на малых или несбалансированных наборах данных. Созданные программные решения могут быть интегрированы в мобильные и встроенные устройства для повышения качества фотосъемки при затрудненных условиях съемки. В целом, результаты работы обеспечивают практическое обоснование использования новых моделей и алгоритмов в различных прикладных задачах компьютерного зрения.

Основные результаты диссертационной работы докладывались и обсуждались на 4 международных научных и научно-технических

конференциях. Результаты работы опубликованы в достаточной степени: общее число публикаций – 8, в том числе 4 статьи в изданиях, рекомендованных ВАК.

В целом, следует отметить, что представленная диссертация является завершенной работой, в которой достаточно подробно и логично изложены постановка и решение рассматриваемых задач, перечислены особенности применяемых при исследовании подходов и методики получения результатов. Автореферат соответствует содержанию диссертации.

Вместе с тем, диссертация Бережнова Н.И. имеет следующие недостатки.

1. В диссертации не всегда достаточно детально приведены используемые гиперпараметры моделей и состав обучающих выборок (архитектурные параметры, объемы и состав датасетов). Уточнение этих данных повысило бы прозрачность и воспроизводимость описанных экспериментов.

2. В автореферате представлены результаты сравнения предлагаемых архитектур с современными моделями восстановления изображений (Restormer, SwinIR, MIRNet, SUNet), однако для усиления экспериментальной части можно рекомендовать более подробное и единообразное сопоставление предлагаемых решений с расширенным набором актуальных архитектур на стандартных датасетах. В частности, для моделей восстановления изображений было бы полезно дополнительно включить сравнение с такими широко используемыми алгоритмами, как DnCNN или Uformer и MIRNet-v2, а для модели WeatherTransformer — сопоставление результатов синтеза с архитектурами, применяемыми для генерации погодных эффектов (например, CycleGAN-подобными моделями или RainNet). Такое расширение позволило бы подчеркнуть преимущества предложенных подходов и повысить объективность экспериментальной оценки.

3. Экспериментальная проверка архитектур и механизмов генерации атмосферных искажений ограничена узким набором условий. Генерация осадков производится по одному эталонному примеру, не исследуются вариации плотности осадков или смешанные погодные

сценарии. Также архитектура сверточных блоков фиксирована (например, сверстка 3×3 , ReLU), и не исследовано влияние глубины или ширины сети на итоговое качество. Расширение экспериментов позволило бы повысить универсальность и обоснованность сделанных выводов.

4. В диссертации не проведен анализ вычислительной эффективности и требований к ресурсам разработанных алгоритмов при реализации на различных аппаратных платформах в том числе на мобильных устройствах. Учет этого аспекта важен для оценки практической реализуемости методов.

5. В диссертации не рассмотрена возможность применения разработанных моделей и алгоритмов к обработке последовательностей изображений (видеоданных). Учет коррелированности видеоданных и зависимость во времени моделируемых помех и искажений позволил бы значительно расширить практическую значимость полученных научных результатов.

Тем не менее, указанные замечания и недостатки в целом не снижают теоретическую и практическую значимость научных результатов, положений, выводов и рекомендаций, сформулированных в диссертации, и носят в большей степени рекомендательный характер.

Вывод. В целом диссертационная работа Бережнова Н.И. представляет собой завершённое научное исследование, в котором решена актуальная задача совершенствования моделей глубокого обучения для обработки и синтеза изображений в системах компьютерного зрения на основе методов и технологий прикладного искусственного интеллекта. Полученные результаты обладают достаточной новизной, обоснованностью и практической значимостью. Работа соответствует основным положениям и критериям, предъявляемым к кандидатским диссертациям (пункты 9, 10, 11, 13, 14 «Положения о присуждении ученых степеней», утвержденного постановлением Правительства РФ № 842 в редакции от 16.10.2024). Содержание работы соответствует специальности 1.2.1 Искусственный интеллект и машинное обучение, а ее автор заслуживает присуждения ученой степени кандидата технических наук.

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

Официальный оппонент, доктор технических наук,
доцент 42 кафедры робототехнических комплексов и систем воздушного базирования 4 факультета беспилотной авиации
ФГКВОУ ВО «Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж) Министерства обороны Российской Федерации

«16» декабря 2025 г.



Шипко Владимир Вацлавович

Шифр и наименование научной специальности, по которой защищена докторская диссертация:

6.2.11 Военная электроника, аппаратура комплексов военного назначения

Подпись Шипко В.В. заверяю

Помощник начальника строевого отдела
ВУНЦ ВВС «ВВА» (г. Воронеж)



И.В Антонов

«16» декабря 2025 г.

Почтовый адрес: 394064, г. Воронеж, ул. Ст. Большевиков, 54а.

Телефон: +7-950-776-84-96

e-mail: shipko.v@bk.ru