

УТВЕРЖДАЮ:

Директор Автономной
некоммерческой
организации высшего образования
«Университет Иннополис»

Вандюков Д.В.

 2026 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертационную работу Ананьева Владислава Валерьевича

«Метод и программные средства адаптивного синтеза изображений высокого разрешения на основе диффузионных моделей при работе с гигапиксельными изображениями», представленную к защите на соискание ученой степени кандидата технических наук по специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»

Актуальность темы диссертационного исследования

Диссертационная работа Ананьева Владислава Валерьевича посвящена разработке метода, алгоритмов и программных средств адаптивного синтеза изображений высокого разрешения на основе латентных диффузионных моделей для компенсации дефицита и неоднородности обучающих данных при автоматизированном анализе гигапиксельных изображений в условиях доменного сдвига и классового дисбаланса.

Современные методы обработки изображений с применением глубоких нейронных сетей, получившие широкое распространение в различных предметных областях, сталкиваются с фундаментальными проблемами: высокой вычислительной сложностью, потребностью в крупных размеченных наборах данных и значительными затратами на их экспертную аннотацию. Особенно остро эти ограничения проявляются при анализе биомедицинских гигапиксельных изображений (например, Whole-Slide Images), характеризующихся многоуровневой иерархической структурой, ковариантным сдвигом, обусловленным вариативностью аппаратных средств оцифровки, и экстремальным классовым дисбалансом, при котором целевые структуры занимают менее 5 % площади.

В последние годы для преодоления дефицита аннотированных данных активно развиваются генеративные методы, среди которых особое место занимают вероятностные диффузионные модели, обеспечивающие высокое качество и стабильность синтеза. Ключевым

элементом управления генерацией выступает механизм обусловливания, однако традиционное текстовое описание неприменимо для строгого контроля пространственного расположения сложных морфологических структур. В связи с этим разработка метода адаптивного синтеза на основе диффузионных трансформеров с визуальным обусловливанием, а также создание модульного программно-алгоритмического комплекса для обработки гигапиксельных изображений представляется современным, перспективным и актуальным направлением исследований.

Оценка содержания диссертации и автореферата

Диссертационная работа включает введение, четыре главы, заключение, список литературы из 74 наименований, список сокращений и условных обозначений, а также пять приложений с листингами исходного кода. Общий объем диссертации составляет 140 страниц, изложение сопровождается 9 рисунками и 9 таблицами. Структура диссертации соответствует общепринятым требованиям, каждая глава содержит обобщающие выводы, а итоговое заключение суммирует основные результаты.

Автореферат диссертации написан лаконично и грамотно, материал изложен в логической последовательности, объем автореферата соответствует установленным требованиям. В автореферате четко отражены основные положения, выносимые на защиту.

Первая глава посвящена анализу предметной области: рассмотрены проблемы обработки гигапиксельных изображений, включая ковариантный сдвиг и дисбаланс классов, проведен сравнительный анализ генеративных архитектур (GAN, VAE, диффузионные модели), обоснован переход к латентным диффузионным трансформерам с визуальным обусловливанием и представлена обобщенная функциональная схема базовой модели.

Во второй главе формализуется концептуальное ядро исследования: представлены математические модели диффузионного процесса в латентном пространстве, метод пространственного сжатия матриц внимания (KV-компрессия), метод доменной специализации на основе прогрессивного обучения и адаптеров LoRA, а также алгоритм локального контекстно-ориентированного встраивания объектов с колориметрическим согласованием в пространстве CIELAB.

Третья глава описывает программную реализацию разработанных методов: модульную архитектуру комплекса, алгоритмическое обеспечение прикладных сценариев синтеза, методы аппаратно-программной оптимизации при работе с гигапиксельными изображениями, а также интерфейсы для интеграции с конвейерами обучения моделей классификации и сегментации.

Четвертая глава содержит результаты вычислительных экспериментов по верификации предложенных решений. Проведена оценка качества синтеза на основе многофакторной методологии, учитывающей статистическую согласованность, семантическую достоверность и генеративное разнообразие. Экспериментально подтверждено, что использование синтетических данных, полученных разработанными алгоритмами, позволяет повысить средние значения метрик качества и снизить их дисперсию при обучении прикладных моделей сегментации и классификации в условиях классового дисбаланса.

Новизна диссертационного исследования

Диссертантом получены следующие новые результаты:

1. Предложен метод адаптивного синтеза гистологических изображений высокого разрешения, новизна которого заключается в гибридном объединении архитектуры латентных диффузионных трансформеров (DiT), механизмов визуального обусловливания и алгоритмов сжатия матриц внимания (KV-компрессии). Формируемый таким сочетанием вычислительно эффективный архитектурный фундамент, в отличие от стандартных монолитных решений, является модульным: он обеспечивает строгий контроль над морфологической достоверностью генерируемых структур и позволяет независимо применять компоненты генеративного конвейера при пакетной обработке гигапиксельных изображений.

2. Предложен метод доменной специализации латентных диффузионных моделей, новизна которого заключается в двухэтапной декомпозиции процесса адаптации: применении прогрессивного обучения с фиксацией оптического увеличения для усвоения морфологии и использовании библиотеки низкоранговых адаптеров (LoRA) для изолированной коррекции стиля. В противовес доминирующей парадигме text-to-image, полный отказ от текстового обусловливания позволяет надежно удерживать геометрию структур базовым визуальным механизмом, в то время как стиль каждого конкретного набора данных кодируется отдельным компактным адаптером. Такое решение делает систему доменной адаптации масштабируемой, позволяя параметрически переключать визуальные характеристики синтеза без полного переобучения базовой модели.

3. Разработан алгоритм локального контекстно-ориентированного встраивания объектов, новизна которого определяется совмещением целевого объекта с фоном непосредственно в латентном пространстве признаков, дополненным автоматизированным гистограммным подбором фонового фрагмента и колориметрическим согласованием (в пространстве CIELAB). Отличие от существующих методов инпейнтинга заключается в достижении бесшовной интеграции объекта без использования дополнительных обучаемых архитектурных модулей и пространственной разметки.

4. Предложена гибридная методология оценки качества синтеза, новизна которой состоит в одновременном учете структурно-семантических свойств и генеративного разнообразия образцов. На основе данной методологии и созданного модульного программно-алгоритмического комплекса экспериментально доказана применимость синтетических данных для повышения качества обучения прикладных нейросетевых моделей в условиях экстремального классового дисбаланса.

Степень обоснованности и достоверности научных положений, выводов и рекомендаций

Диссертант корректно использует известные научные методы обоснования полученных результатов, выводов и рекомендаций. Для формализации математических моделей применяются фундаментальные принципы диффузионных процессов и трансформерных архитектур, выбор которых научно обоснован. Разработанные алгоритмы строятся на эффективных численных методах, используемых в современной практике глубокого обучения. Этапы проектирования и разработки программного обеспечения детально описаны и соответствуют стандартным требованиям.

Достоверность результатов обеспечивается использованием современных средств и методов математического моделирования, непротиворечивостью и логичностью конструируемых алгоритмов, обоснованностью используемых приближений. Адекватность

результатов, полученных с помощью вычислительных экспериментов, контролируется сравнением с данными натурных исследований и практической сходимостью метрик качества при варьировании параметров генерации. Проведенное статистическое тестирование подтверждает значимость прироста показателей прикладных моделей при добавлении синтетических данных.

Теоретическая и практическая значимость

Результаты, полученные в диссертационной работе Ананьева В.В., обладают научной значимостью. Теоретическая значимость работы заключается в развитии подходов к условной генерации изображений высокого разрешения на основе латентных диффузионных трансформеров, создании методов доменной адаптации и контекстно-ориентированного встраивания объектов, а также в разработке гибридной методологии многофакторной оценки качества синтеза. Предложенные решения могут быть адаптированы для решения задач в других предметных областях, где требуется восполнение дефицита обучающих выборок и компенсация доменного сдвига.

Практическая значимость состоит в возможности прогнозирования закономерностей и повышения качества автоматизированного анализа гистологических изображений с помощью разработанных программных средств. Модульная архитектура комплекса позволяет масштабировать его под различные прикладные задачи цифровой патологии, а интеграция синтетических данных в обучающие выборки дает выраженный регуляризирующий эффект, что подтверждено экспериментально на задачах сегментации кровеносных сосудов и классификации участков инвазии. Разработанные методические и программные решения внедрены в научную и практическую деятельность 1-го патологоанатомического отделения ФГБУ «НМИЦ АГП им. В.И. Кулакова» Минздрава России.

Публикация и апробация основных научных результатов диссертации

Список научных работ автора представляется достаточным и включает 4 печатных издания, удовлетворяющих критериям ВАК РФ, из которых 3 статьи в журналах, индексируемых международными базами Scopus и Web of Science. Получено свидетельство о государственной регистрации программы для ЭВМ.

Основные результаты по материалам диссертационной работы были представлены и обсуждались на международных научно-практических конференциях: «Анализ данных в медицине» (2023, 2024), Сеченовском международном биомедицинском саммите (2024), открытой конференции ИСП РАН (2024), Петербургском медицинском инновационном форуме (2025). Кроме того, в период выполнения диссертационного исследования автор являлся победителем конкурсного отбора на назначение стипендии Президента Российской Федерации для аспирантов и адъюнктов.

Важной отличительной чертой работы является обобщаемость полученных результатов и их потенциал для применения в иных предметных областях, помимо цифровой патологии. Предложенные методы адаптивного синтеза на основе латентных диффузионных трансформеров, алгоритмы доменной специализации и контекстно-ориентированного встраивания объектов, а также разработанный модульный программный комплекс не являются узкоспециализированными и могут быть адаптированы для широкого круга задач, связанных с дефицитом и неоднородностью обучающих данных: от анализа спутниковых снимков и

промышленного неразрушающего контроля до задач экологического мониторинга и медицинской визуализации других модальностей. Это подтверждает методологическую корректность выбранного подхода и делает результаты диссертации ценными для развития всего направления генеративного моделирования в прикладных вычислительных системах.

Общие замечания и вопросы по диссертационной работе

Диссертация Ананьева В.В. имеет ряд недостатков, которые по итогам обсуждения можно сформулировать в виде следующих замечаний и вопросов.

1. Работа носит выраженный прикладной характер, что само по себе не является недостатком. Однако научный вклад сформулирован недостаточно четко: не всегда ясно, где именно заканчивается инженерная интеграция известных решений и начинается новый метод. Например, в разделе 2.1 описывается объединение диффузионного трансформера, механизма визуального обусловливания и алгоритма KV-компрессии, но степень новизны каждой компоненты и их совокупности не выделена с достаточной строгостью.

2. Хотелось бы видеть более строгую математическую связь между синтезом данных, ростом вариативности обучающей выборки, количеством синтетических примеров и улучшением обобщающей способности прикладных моделей. В текущей версии эти зависимости описываются преимущественно на эмпирическом уровне, без формализации в виде аналитических оценок или границ применимости. Представляется целесообразным в дальнейшем расширить теоретическое обоснование обобщаемости предложенных решений.

3. Не хватает сравнения с альтернативными решениями. Поскольку заявленный результат во многом строится на выборе конкретных компонентов – диффузионной модели, visual conditioning, LoRA-адаптации, KV-компрессии и локального inpainting, – в работе желательно шире обосновать, почему именно такая комбинация является необходимой или оптимальной. Возможны альтернативы: использование VLM и фундаментальных визуально-языковых моделей, других механизмов условной генерации, иных способов parameter-efficient fine-tuning вместо LoRA, self-supervised representation learning как предварительного этапа, обучаемых моделей оценки качества синтетических изображений и т. д. В диссертации присутствует сравнительный анализ GAN, VAE и диффузионных моделей, а также обоснование выбора LDM/DiT, однако для убедительности вклада хотелось бы увидеть более подробные ablation studies, показывающие вклад каждого ключевого элемента конвейера в итоговый результат (например, таблицу с последовательным отключением KV-компрессии, LoRA-адаптации или механизма визуального обусловливания).

4. Оценка качества синтетических данных требует большей строгости. Помимо агрегированных метрик качества генерации (FID, KID, CLIP-FID), желательно добавить более подробную экспертную и статистическую проверку: анализ артефактов, устойчивость результатов при варьировании начального приближения, построение доверительных интервалов для всех ключевых метрик, а также детальное исследование влияния синтетических данных именно на редкий целевой класс, например, оценку precision/recall отдельно для минорного класса до и после добавления синтезированных образцов.

5. По списку публикаций соискателя имеются дополнительные замечания. В ряде работ, опубликованных в высокорейтинговых журналах (*Biomolecules*, *Journal of Clinical Medicine*), число соавторов достигает 7–10 человек. В условиях столь широкого соавторства категорически важно четко и детально обозначить личный вклад автора диссертации в каждой

из этих публикаций, особенно в части разработки методов и программных средств, составляющих основу диссертационного исследования. В автореферате представлено общее описание вклада, однако оно не всегда позволяет однозначно выделить результаты, полученные лично соискателем, и отделить их от инженерной реализации или интерпретации данных, выполненных соавторами. Кроме того, следует отметить, что тематические рубрики журналов *Journal of Clinical Medicine* и *Biomolecules* ориентированы преимущественно на клиническую медицину и биологию и не содержат тематических рубрик, близких по научным направлениям к представляемой специальности.

Тем не менее, отмеченные недостатки, отсутствующие пояснения и возникшие вопросы к соискателю не влияют на главные теоретические и практические результаты диссертации и не меняют общую положительную оценку диссертации.

Заключение

Полученные соискателем результаты достоверны, выводы и заключения научно обоснованы. Материал диссертации изложен вполне конструктивно и содержит большой потенциал новизны в научных областях, соответствующих специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей». Оформление текста соответствует требованиям, предъявляемым к диссертационным работам, научный стиль диссертации выдержан. Основные научные результаты достаточно полно опубликованы в работах соискателя. Автореферат диссертации отражает ее основное содержание.

Диссертационная работа Ананьева Владислава Валерьевича представляет собой законченную научно-квалификационную работу, в которой содержится решение важной научной задачи, имеющей значение для развития подходов к адаптивному синтезу изображений высокого разрешения и их применению в цифровой патологии и других предметных областях. На основании детального анализа материалов диссертации соискателя можно заключить, что результаты выполненных исследований обладают значимостью для развития научной специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей».

Таким образом, диссертационная работа Ананьева Владислава Валерьевича на тему «Метод и программные средства адаптивного синтеза изображений высокого разрешения на основе диффузионных моделей при работе с гигапиксельными изображениями», соответствует критериям, установленным Положением Постановления Правительства «О порядке присуждения ученых степеней» от 24.09.2013 № 842 (в действующей редакции), предъявляемым к кандидатским диссертациям. Автор работы заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей».

Доклад по диссертационной работе Ананьева Владислава Валерьевича «Метод и программные средства адаптивного синтеза изображений высокого разрешения на основе диффузионных моделей при работе с гигапиксельными изображениями» был заслушан на расширенном заседании Ученого совета АНО ВО «Университет Иннополис». Отзыв на

диссертацию обсужден и утвержден на этом же семинаре в качестве официального отзыва ведущей организации, протокол №УС-1/2026/07/08 от «8» июля 2026 г.

Отзыв подготовил:

профессор, доктор физико-математических наук Масловская Анна Геннадьевна,
профессор Института анализа данных и искусственного интеллекта, Факультет компьютерных
и инженерных наук

«8» июля 2026 г.

___А.Г. Масловская