

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию Ананьева Владислава Валерьевича

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

Актуальность темы

Эффективность обучения глубоких нейронных сетей для анализа изображений во многом определяется объемом, разнообразием и качеством размеченных данных. Получение экспертных аннотаций, особенно в прикладных областях, связано с существенными временными и финансовыми затратами. При недостатке таких данных снижаются устойчивость и обобщающая способность моделей, а при выраженном дисбалансе классов ухудшается качество распознавания редких, но практически значимых объектов.

Традиционные методы аугментации, которые обычно включают геометрические и колориметрические преобразования, помогают расширить вариативность уже имеющейся выборки, однако не формируют новые семантически и топологически корректные образцы. В связи с этим разработка генеративных методов синтеза обучающих данных является актуальной задачей машинного обучения и компьютерного зрения. Особый интерес представляют генеративные модели, в частности семейство диффузионных моделей, способные обеспечить возможность управляемой генерации изображений с учетом особенностей целевого предметного домена.

Цифровая патология является примером области, в которой данная задача приобретает особую сложность. В данной предметной области приходится работать с гигапиксельными изображениями гистологических препаратов, обработка которых осложняется многоуровневой структурой данных, ограничениями вычислительных ресурсов, различиями в условиях подготовки таких изображений и связанным с ними доменным сдвигом. Кроме того, целевые структуры на таких изображениях могут быть представлены в незначительном количестве, что усиливает проблему классового дисбаланса при обучении нейросетевых моделей.

Диссертационная работа В. В. Ананьева посвящена решению описанной научно-технической задачи. Автором разработаны метод адаптивного синтеза изображений с визуальным обуславливанием, метод доменной специализации латентных диффузионных моделей и алгоритм локального контекстно-ориентированного встраивания объектов. Их программная реализация объединена в модульный комплекс, предназначенный для

формирования обучающих наборов данных при работе с изображениями высокого разрешения. Тема исследования является актуальной и способна внести вклад в развитие методов обучения глубоких нейронных сетей, генеративного моделирования и программных средств анализа изображений.

**Достоверность и степень обоснованности научных положений,
выводов и рекомендаций, сформулированных в диссертации**

Представленные в диссертации положения имеют достаточное теоретическое и экспериментальное обоснование. Построение метода опирается на математическую модель латентной диффузии, архитектуры диффузионных трансформеров и известные подходы к параметрически эффективной адаптации нейронных сетей. В работе последовательно описаны алгоритмы, их программная реализация и организация вычислительного процесса для обработки гигапиксельных изображений.

Экспериментальная часть выполнена на клинических выборках цифровой патологии. Корректность оценки качества синтеза подтверждается применением нескольких групп метрик: измеряются близость распределений реальных и сгенерированных данных, сохранение доменно-специфичной семантики и разнообразие образцов. Проверка проводится в признаковых пространствах нескольких независимых энкодеров. Практическая состоятельность предложенных решений дополнительно демонстрируется в задачах сегментации и классификации: синтетические данные улучшают показатели прикладных моделей.

Выводы автора согласуются с материалами проведенных экспериментов. В диссертации выполнен анализ отечественных и зарубежных работ по теме, а результаты исследования представлены на профильных научных мероприятиях и отражены в публикациях, индексируемых в Scopus и Web of Science.

**Новизна исследования, полученных результатов, выводов и рекомендаций,
сформулированных в диссертации**

Научная новизна диссертации состоит в получении следующих результатов:

1. Разработан метод адаптивного синтеза гистологических изображений высокого разрешения, использующий латентный диффузионный трансформер с визуальным условием. В состав метода включен механизм KV-компрессии, направленный на снижение вычислительной сложности операций внимания при обработке изображений высокого разрешения.
2. Предложен метод доменной специализации латентных диффузионных моделей, разделяющий адаптацию на этап прогрессивного обучения для формирования морфологических представлений и этап настройки стиливых характеристик целевого домена с использованием низкоранговых адаптеров LoRA.

3. Разработан алгоритм локального контекстно-ориентированного встраивания объектов. Совмещение донора и фонового фрагмента осуществляется в латентном пространстве и сопровождается автоматическим колориметрическим согласованием. Алгоритм не включает дополнительные обучаемые модули и не требует ручной разметки.
4. Создана гибридная методология оценки синтезируемых данных. Она объединяет оценку статистической согласованности распределений, сохранения предметно-специфичной семантики и генеративного разнообразия в нескольких независимых признаковых пространствах. Ее применение показало, что синтетические данные повышают IoU и специфичность модели семантической сегментации при снижении дисперсии оценок, а в задаче классификации обеспечивают рост F1-меры, чувствительности и специфичности.

Значимость для науки и практики полученных автором результатов

С теоретической точки зрения работа развивает подходы к условному синтезу медицинских изображений. Существенным результатом является формализация двухэтапной доменной специализации латентных диффузионных моделей, позволяющей отделить настройку морфологии от настройки стиливых особенностей исходных данных.

Практическая ценность определяется созданием программно-алгоритмического комплекса, предназначенного для генерации морфологически достоверных обучающих выборок в условиях дефицита клинических данных и доменного сдвига. Результаты сквозного эксперимента подтверждают эффективность такого расширения выборок: F1-мера классификатора возросла с 0,9279 до 0,9577. На программный модуль получено свидетельство о государственной регистрации программы для ЭВМ. Разработанный комплекс внедрен в деятельность патологоанатомического отделения ФГБУ «НМИЦ АГП им. В. И. Кулакова» Минздрава России.

Рекомендации по использованию результатов диссертации

Материалы диссертации могут быть использованы при разработке элементов систем обработки данных цифровой патологии и в исследованиях, связанных с генерацией и расширением обучающих наборов медицинских изображений. Предложенные автором методы потенциально применимы также в других задачах управляемого синтеза, где необходимо сохранять пространственную организацию объектов и адаптировать модель к изменению визуального домена.

Оценка содержания диссертации, ее завершенность

Диссертация включает введение, четыре главы, заключение, список литературы, список сокращений и приложения. Логика изложения соответствует поставленной цели и задачам исследования.

В первой главе рассмотрены особенности WSI, источники доменного сдвига и классового дисбаланса, а также современные генеративные архитектуры. По итогам обзора автор формирует функциональную схему латентной диффузионной модели с визуальным обуславливанием и конкретизирует постановку задачи.

Вторая глава содержит формализованное описание основных предлагаемых решений. Последовательно изложены метод адаптивного синтеза, метод доменной специализации и алгоритм контекстно-ориентированного встраивания объектов. Описание доведено до математической постановки и алгоритмического представления вычислительного конвейера.

Третья глава раскрывает вопросы программной реализации. Рассматриваются модульный принцип устройства разработанного комплекса, реализация прикладных сценариев генерации и методы снижения вычислительных затрат при обработке гигапиксельных изображений.

Четвертая глава посвящена экспериментальной верификации. В ней оцениваются вычислительная эффективность предложенных методов оптимизации, достоверность синтезируемых изображений и влияние синтетических данных на качество моделей сегментации и классификации. Эксперименты подтверждают положительное влияние разработанных методов на рассматриваемые прикладные задачи.

В целом диссертационную работу Ананьева В. В. можно считать завершенным научным исследованием, изложенным ясным и последовательным языком. Полученные результаты обладают научной новизной, имеют практическую значимость и соответствуют заявленной специальности.

Замечания к содержанию и оформлению диссертации

К работе есть ряд замечаний:

1. В листингах 3.1 и 3.2 приведены параметры работы алгоритмов, такие как ранг адаптера $r=4$, масштабирующий коэффициент $\alpha=4$, а также гиперпараметры $t=0,55$ и $\lambda=0,85$. Не до конца понятно, являются ли значения, выбранные по результатам проведенных экспериментов, оптимальными, и насколько важно в контексте исследования, чтобы данные параметры были оптимальными.
2. Оценка вычислительной эффективности проведена на одном ускорителе. Вопрос масштабируемости комплекса на распределенные конфигурации с несколькими GPU

раскрыт слабо, хотя архитектура заявлена как поддерживающая распределенное обучение.

3. Набор данных LCNOV (таблица 4) приведен без ссылки на источник — в отличие от DHMC[56] и NLST[58] — и нигде в тексте не охарактеризован. Поскольку именно разнородность данных между обучающими выборками мотивирует постановку задачи, краткое описание происхождения и условий оцифровки этого набора было бы уместным.
4. Так же в таблице 4 столбец «Значение метрики» - указано много цифр после запятой, но не указана их значимость. Не указан референс: с чем сравнивать цифры в каждой строке? Что здесь будет хорошее или плохое качество?
5. При поступлении достаточного количества новых данных рекомендуется провести статистические тесты на значимость результатов: дает ли применение новых методов статистически значимый прирост качества моделей?

Приведенные замечания имеют уточняющий характер. Они не ставят под сомнение основные результаты диссертации и не изменяют ее общей положительной оценки.

Заключение о соответствии диссертации критериям, установленным Положением о порядке присуждения ученых степеней

Основные результаты диссертации опубликованы в четырех научных работах, из которых три представлены в изданиях, индексируемых в международных базах Scopus и Web of Science. Кроме того, автором получено свидетельство о государственной регистрации программы для ЭВМ.

Автореферат соответствует содержанию диссертационной работы, отражает актуальность выбранной тематики, научную новизну, практическую значимость, поставленные задачи и основные результаты исследования.

Диссертация Ананьева Владислава Валерьевича «Метод и программные средства адаптивного синтеза изображений высокого разрешения на основе диффузионных моделей при работе с гигапиксельными изображениями» является завершенной научно-квалификационной работой, в которой решена актуальная научно-техническая задача, имеющая значение для математического и программного обеспечения вычислительных систем. Работа отвечает требованиям п. 9 «Положения о порядке присуждения ученых степеней», утвержденного постановлением Правительства Российской Федерации от 24.09.2013 № 842, предъявляемым к диссертациям на соискание ученой степени кандидата наук. Автор работы заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.5 — «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей».

Официальный оппонент

Шараев Максим Геннадьевич

кандидат физико-математических наук, старший преподаватель, руководитель лаборатории Центра прикладного ИИ Автономной некоммерческой образовательной организации высшего профессионального образования «Сколковский институт науки и технологий»

Дата « 27 » августа 2026 г.

Шараев Максим Геннадьевич