МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«Национальный исследовательский ядерный университет «МИФИ»
(НИЯУ МИФИ)

УТВЕРЖДАЮ
Ректор НИЯУ МИФИ
В.И. Шевченко

«12» января 2026 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ»
на диссертационную работу
Ушакова Егора Николаевича
«Метод условной генерации изображений в диффузионных моделях со структурированным текстовым условием»
представленную на соискание ученой степени кандидата физико-математических наук по специальности 2.3.5 — «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»

Диссертационная работа Ушакова Егора Николаевича посвящена разработке метода условной генерации изображений на основе диффузионных моделей, в котором неструктурированное текстовое описание преобразуется в компактное структурированное условие и используется для управления процессом генерации. Разработка и экспериментальная проверка предложенного метода выполнены на примере медицинских изображений.

Целью работы является разработка и экспериментальная валидация метода условной генерации синтетических изображений на основе диффузионных моделей со структурированным заданием условия, формируемым из текстового описания, а также создание программной реализации, обеспечивающей повышение качества генерации и практической полезности синтетических данных.

Актуальность темы

Генеративное моделирование в настоящее время является одним из активно развивающихся направлений машинного обучения и анализа данных. Особое место в нем занимают диффузионные модели, продемонстрировавшие высокое качество генерации изображений и устойчивость обучения. При этом для широкого класса прикладных задач требуется не безусловный синтез данных, а управляемая генерация объектов, обладающих заранее заданными свойствами.

Одним из существенных факторов, определяющих качество условной генерации, является способ представления условия. Свободные текстовые описания могут содержать неоднозначные формулировки, избыточную информацию, отрицания, ссылки на контекст и другие фрагменты, не имеющие непосредственного отображения в генерируемом объекте. В результате возрастает

сложность интерпретации условия моделью и снижается воспроизводимость соответствия синтезируемого изображения исходному описанию.

Указанная проблема особенно выражена при работе с медицинскими данными. Радиологические отчеты содержат сложную предметную семантику, характеризуются вариативностью формулировок и одновременно требуют точного соответствия изображения описываемым признакам. Кроме того, медицинские наборы данных характеризуются высокой стоимостью экспертной разметки, дисбалансом классов и ограничениями, связанными с доступом к данным пациентов. В связи с этим разработка метода преобразования исходного текстового описания в структурированное представление, пригодное для непосредственного использования в диффузионной модели, является актуальной научной задачей.

Основные результаты работы

1. Разработан метод формирования структурированного текстового условия для диффузионной генерации изображений, основанный на разложении исходного неструктурированного описания на семантически значимые компоненты.
2. Разработаны алгоритмы формирования списка клинических находок, включающие выделение релевантной части текстового описания, нормализацию формулировок, обработку отрицаний и неопределенности, связывание признаков с их атрибутами, объединение дублирующихся элементов и разрешение конфликтов.
3. Для алгоритмов, входящих в состав предложенного метода, получены аналитические оценки вычислительной и пространственной сложности. Сформулированы и доказаны соответствующие теоремы. Показано, в частности, что использование структурированного условия вместо исходного текста не изменяет асимптотическую сложность обучения и генерации используемой латентной диффузионной модели.
4. Разработаны архитектура и программная реализация комплекса условной генерации изображений, объединяющего подготовку исходных данных, формирование структурированного текстового условия, обучение и применение диффузионной модели, а также средства экспериментальной оценки получаемых синтетических данных.
5. Выполнена экспериментальная проверка разработанного метода на рентгенограммах органов грудной клетки из набора MIMIC-CXR. При сохранении одной и той же архитектуры генеративной модели использование структурированного списка находок вместо исходной секции FINDINGS обеспечило улучшение распределительных метрик FID и KID. В downstream-эксперименте также показано преимущество синтетических данных, сформированных с использованием структурированного условия.
6. Проведен отдельный комплекс экспериментов, направленных на оценку сходства синтетических и обучающих изображений, возможности определения принадлежности объекта к обучающей выборке и риска повторной идентификации пациента. Полученные результаты не выявили увеличения исследованных рисков при использовании предложенного представления условия.
7. Дополнительно исследована переносимость принципа структурированного задания условия на маммографические изображения набора VinDr-Mammo. Полученные результаты показывают возможность использования синтетических изображений в качестве дополнения к реальным данным при обучении прикладных моделей.

Научная новизна

Научная новизна диссертационной работы заключается в разработке нового метода структурированного представления текстового условия для задачи диффузионной генерации изображений.

В отличие от непосредственного использования свободного текстового описания, предложенный метод предусматривает предварительное выделение семантически значимых компонентов, их нормализацию и формирование компактного условия фиксированной структуры. Это позволяет исключать нерелевантные фрагменты исходного текста и в явной форме представлять признаки и связанные с ними атрибуты.

Новыми являются разработанные алгоритмы формирования структурированного условия и полученные для них аналитические оценки вычислительной и пространственной сложности. Существенным результатом является также доказательство того, что замена исходного текстового описания его структурированным представлением при сохранении используемого текстового энкодера и механизма передачи условия не приводит к изменению асимптотической сложности обучения и генерации латентной диффузионной модели.

Теоретическая и практическая значимость полученных результатов

Теоретическая значимость работы состоит в развитии методов условного генеративного моделирования за счет формализации преобразования неструктурированного текстового описания в структурированное условие и анализа вычислительных свойств соответствующих алгоритмов. Полученные математические представления и оценки могут быть использованы при дальнейшем исследовании методов условной генерации, способов представления мультимодальных данных и алгоритмов подготовки текстовых условий для генеративных моделей.

Практическая значимость определяется созданием программного комплекса, реализующего воспроизводимый конвейер от подготовки исходного текстового описания и формирования структурированного условия до обучения диффузионной модели, синтеза изображений и экспериментальной оценки полученных результатов. Разработанные решения могут быть использованы для формирования и расширения обучающих выборок, моделирования редко встречающихся состояний, балансировки наборов данных и экспериментальной проверки алгоритмов анализа изображений. Результаты диссертационной работы могут быть использованы при разработке программных систем условной генерации изображений, в которых исходное управляющее описание представлено в текстовой или иной слабоструктурированной форме.

Достоверность и степень обоснованности результатов

Достоверность результатов диссертационного исследования обеспечивается корректной постановкой решаемой задачи, использованием известных и апробированных математических методов и архитектур генеративных моделей, теоретическим анализом разработанных алгоритмов и проведением вычислительных экспериментов.

Для алгоритмов формирования структурированного условия получены аналитические оценки вычислительной и пространственной сложности, обоснованные доказательствами соответствующих теорем.

Экспериментальная часть построена таким образом, что при основном сравнении двух вариантов условия сохраняются неизменными архитектура Stable Diffusion 1.5, набор

изображений и основные параметры обучения и генерации. Это позволяет отделить влияние исследуемого способа представления текстового условия от влияния других факторов.

Оценка выполнена с использованием нескольких взаимодополняющих подходов, включающих стандартные метрики генеративных моделей, downstream-эксперименты и анализ характеристик конфиденциальности синтетических данных. Выводы, сформулированные автором, соответствуют представленным экспериментальным результатам.

Основные результаты работы прошли апробацию на профильных научных конференциях и отражены в четырех публикациях в изданиях, рекомендованных ВАК.

Замечания

Наряду с общей положительной оценкой работы следует отметить несколько замечаний:

1. Описание экспериментальной базы было бы полезно дополнить сведениями об объемах всех выборок, распределении классов, разбиении данных и использованной вычислительной конфигурации. Это упростило бы сопоставление условий экспериментов и оценку влияния дисбаланса классов.
2. В обзоре литературы желательно более явно связать предложенный метод с конкретными работами предшественников, указав, какие их положения и решения послужили основой исследования.
3. В качестве дальнейшего развития работы представляет интерес апробация метода на локальных выборках медицинских организаций, не представленных в открытом доступе. В дополнение к проведенным экспериментам на открытых данных это позволило бы оценить применимость метода при иных условиях получения изображений и составления текстовых описаний.
4. В дальнейшем также рекомендуется дополнить работу проверкой статистической значимости различий между сопоставляемыми вариантами условий.

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

Диссертационная работа Ушакова Егора Николаевича «Метод условной генерации изображений в диффузионных моделях со структурированным текстовым условием» является завершенной научно-квалификационной работой, в которой решена научная задача разработки и обоснования метода преобразования текстового описания в структурированное условие диффузионной генерации изображений, имеющая значение для развития математического и программного обеспечения систем обработки и анализа данных.

По актуальности, научной новизне, обоснованности результатов, теоретической и практической значимости диссертация соответствует требованиям Положения о присуждении ученых степеней, утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 (в действующей редакции), предъявляемым к кандидатским диссертациям, а ее автор, Ушаков Егор Николаевич, заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 2.3.5 — «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей».

Диссертационная работа рассмотрена на совместном семинаре кафедр «Информатика и процессы управления», «Анализ конкурентных систем» Федерального государственного

автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ», протокол № 2/26/27 от «15» сентября 2026 г.

Отзыв подготовили

Заведующий кафедрой «Информатика и процессы управления» НИЯУ МИФИ, д.т.н., проф. — Алексей Дмитриевич Модяев

Заведующий кафедрой «Анализ конкурентных систем» НИЯУ МИФИ, д.т.н., доц. — Алексей Анатольевич Артамонов

Сведения о ведущей организации:
Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский ядерный университет «МИФИ»

Адрес: 115409, г. Москва, Каширское шоссе, д. 31.