

Отзыв научного руководителя

на диссертационную работу Ушакова Егора Николаевича на тему:

«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».

Ушаков Егор Николаевич, 1997 года рождения, в 2021 году окончил Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский ядерный университет «МИФИ» по специальности 14.05.01 «Ядерные реакторы и материалы». В 2021-2025 годах обучался в очной аспирантуре ИСП РАН по направлению подготовки 09.06.01 «Информатика и вычислительная техника».

В период подготовки диссертации Ушаков Е. Н. работал в ИСП РАН инженером - исследователем 30 центра Цифровой экосистемы НЦМУ «Цифровой биодизайн и персонализированное здравоохранение» и занимался исследованиями в области машинного обучения, нейронных сетей и компьютерного зрения. Диссертационная работа посвящена разработке метода условной генерации изображений в диффузионных моделях со структурированным текстовым условием.

Тема исследования обладает высокой актуальностью в связи с развитием генеративного моделирования и расширением применения диффузионных моделей. В задачах условной генерации существенное значение имеет форма задания условия: свободный текст может содержать неоднозначные формулировки, отрицания, исторические и технические сведения, не имеющие непосредственного визуального соответствия. Это затрудняет интерпретацию условия моделью и снижает управляемость генерации. Данная проблема особенно актуальна при работе с медицинскими данными, содержащими сложную клиническую семантику.

В результате исследования Ушаковым Е. Н. предложен новый метод формирования структурированного текстового условия для диффузионной генерации изображений. Метод основан на разложении неструктурированного описания на семантически значимые компоненты, их нормализации и фильтрации нерелевантной информации. На медицинских данных условие реализовано в виде списка клинических находок с указанием типа признака, его статуса, локализации и степени выраженности.

Для разработанных алгоритмов получены аналитические оценки вычислительной и пространственной сложности. Показано, что вычислительная сложность формирования списка находок линейна относительно длины обрабатываемого текста, а пространственная сложность его хранения - относительно числа выделенных находок. Также доказано, что замена секции FINDINGS структурированным представлением при использовании того же текстового энкодера не изменяет асимптотическую сложность обучения и генерации латентной диффузионной модели.

В рамках работы разработаны архитектура и программная реализация комплекса условной генерации изображений со структурированным текстовым условием. Комплекс включает подготовку исходного текста и обучающих пар, формирование структурированного

условия, обучение и применение латентной диффузионной модели, а также экспериментальную оценку синтетических данных.

Экспериментальные исследования, выполненные на рентгенограммах органов грудной клетки, подтвердили работоспособность и эффективность предложенного метода. Сравнение с исходным способом задания текстового условия показало, что структурированное представление позволяет улучшить характеристики генерируемых изображений и повысить их полезность при решении прикладных задач. Полученные результаты свидетельствуют об обоснованности предложенных в диссертации решений.

Сильной стороной работы является комплексный характер экспериментальной оценки. Помимо стандартных метрик проведена оценка полезности синтетических изображений, исследованы сходство с обучающими данными, возможность определения принадлежности к обучающей выборке и риск повторной идентификации пациента. Признаки практически точного воспроизведения обучающих изображений не выявлены, а показатели определения принадлежности находились вблизи случайного уровня. Эксперимент на наборе данных из другого домена также подтвердил переносимость принципа структурированного задания условия.

Работа имеет теоретическую и практическую значимость. Теоретическая значимость состоит в формализации преобразования неструктурированного текста в структурированное условие, разработке алгоритмов и оценке их вычислительных свойств. Практическая значимость определяется созданием воспроизводимого программного конвейера, пригодного для расширения и балансировки обучающих выборок, моделирования редких состояний и тестирования алгоритмов анализа изображений.

Считаю, что диссертационная работа Ушакова Егора Николаевича «Метод условной генерации изображений в диффузионных моделях со структурированным текстовым условием» является завершенной научно-квалификационной работой, содержащей новые научные результаты в области методов, алгоритмов и программных средств условного генеративного моделирования. Работа отвечает требованиям, предъявляемым к кандидатским диссертациям по специальности 2.3.5 - «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей», а ее автор заслуживает присуждения ученой степени кандидата физико-математических наук.

Научный руководитель,
заведующий 30 центра
Цифровой экосистемы НЦМУ
«Цифровой биодизайн и
персонализированное здравоохранение»
ИСП РАН, к.ф.-м.н.

Е.А. Карпулевич

«25» августа 2026 г.