ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ
на выпускную квалификационную работу

Авшалумова Михаила Сергеевича
«Разработка мультимодальных нейросетевых методов распознавания изображений на основе трансформерных архитектур»

Направление подготовки/специальность
09.04.01 Информатика и вычислительная техника
Направленность (профиль) подготовки
Методы и технологии искусственного интеллекта

Выпускная квалификационная работа посвящена решению важной задачи в области компьютерного зрения — повышению эффективности распознавания объектов на изображениях по произвольным текстовым запросам. Актуальность темы определяется растущим интересом к мультимодальным системам и необходимостью обеспечения их надежности и точности работы в условиях неопределённости и вариативности входных данных.

В рамках выполнения ВКР разработан метод, предполагающий трансформацию исходных пользовательских запросов в более релевантную форму с использованием вспомогательной языковой модели. Обучение данной модели осуществляется по парадигме обучения с подкреплением, где функция вознаграждения задаётся реакцией целевой модели распознавания. Такой подход позволяет гибко адаптировать взаимодействие между текстовым вводом и визуальной моделью, что, как подтверждено в экспериментальной части, приводит к существенному улучшению точности распознавания.

Работа имеет логичную и последовательную структуру: представлено обоснование актуальности, сформулированы цели и задачи исследования, подробно описана архитектура решения, проведён анализ полученных результатов с их сопоставлением с существующими методами на нескольких датасетах, включая собственные. Это свидетельствует о практической применимости и универсальности предложенного подхода.

К числу достоинств работы можно отнести успешное объединение языковых и визуальных моделей с применением механизма обратной связи, грамотное использование методов обучения с подкреплением, а также ясное и чёткое изложение материала, сопровождаемое наглядными иллюстрациями и примерами. Кроме того, значимым вкладом являются оригинальные решения и использование собственных данных для валидации предложенного метода.

В качестве замечаний к работе можно выделить следующее: 1) эксперименты можно было бы провести с использованием большего количества современных методов распознавания объектов на изображениях с открытым словарем, 2) большие языковые модели для переформулирования текстового запроса весьма медленно работают и

следовало рассмотреть возможности их ускорения, 3) в тексте встречаются отдельные англоязычные термины, требующие пояснения для неспециалистов.

В целом, работа выполнена на высоком научно-техническом уровне. Студент продемонстрировал развитые исследовательские навыки, высокий уровень самостоятельности, способность к критическому анализу и обобщению информации, а также умение ясно и логично излагать материал. Студент уверенно ориентируется в предметной области, владеет современным научным аппаратом и способен к дальнейшей профессиональной и исследовательской деятельности.

Выпускная квалификационная работа заслуживает оценки «отлично» и полностью соответствует требованиям, установленным к ВКР студентов МФТИ. Рекомендую присвоить Авшалумову Михаилу Сергеевичу квалификацию магистра по направлению 09.04.01 «Информатика и вычислительная техника» по профилю подготовки «Методы и технологии искусственного интеллекта».

Научный руководитель:

Юдин Дмитрий Александрович,
кандидат технических наук, заведующий
лабораторией интеллектуального транспорта
МФТИ - НКБ ВС, Научно-образовательный
центр когнитивного моделирования МФТИ,
e-mail: yudin.da@mipt.ru

/ Юдин Д.А.
19.06.2025

Подпись руки Юдина
заверяю:
Администратор канцелярии
Административного отдела
О.А. Кораблева

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
КАНЦЕЛЯРИЯ