Автономная некоммерческая организация
«Институт Искусственного интеллекта»

121170, Москва, Кутузовский пр-т, 32, к. 1, 4.В.08
E-mail: info@airi.net, Web: airi.net
ОГРН 1207700493978 ИНН 7730261209 КПП 773001001

# РЕЦЕНЗИЯ
## на выпускную квалификационную работу

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

Данная работа посвящена актуальной задаче в области компьютерного зрения — улучшению качества распознавания объектов на изображениях по произвольным текстовым запросам. Актуальность темы обусловлена широким распространением мультимодальных систем распознавания изображений и растущей потребностью в их точной и адаптивной работе в условиях неопределённости входных данных.

В работе предложен метод, в рамках которого исходные пользовательские запросы преобразуются в более подходящую форму с помощью вспомогательной языковой модели. Эта модель обучается с использованием методов обучения с подкреплением, где в качестве сигнала вознаграждения выступает отклик целевой модели распознавания. Подобная схема позволяет гибко настраивать взаимодействие между пользовательским вводом и архитектурой компьютерного зрения, что, как показано в экспериментах, существенно повышает точность распознавания.

Работа имеет чёткую структуру, включает описание актуальности задачи, формулировку целей и подхода, описание архитектуры решения и подробный анализ экспериментальных результатов. Приведены сравнения с базовыми методами на нескольких наборах данных, в том числе собственных, что подтверждает универсальность и прикладную ценность предложенного метода.

К числу достоинств работы можно отнести:

- интеграцию языковых и визуальных моделей с обратной связью;
- грамотное применение обучения с подкреплением;
- ясное изложение материала и достаточное количество иллюстраций и примеров;
- наличие оригинальных наработок и собственных данных для валидации результатов.

Критические замечания носят скорее стилистический и редакторский характер. В тексте встречаются отдельные англицизмы, а также термины, требующие пояснения для читателя без специальной подготовки. В дальнейшем автору можно рекомендовать больше внимания уделить вопросам интерпретируемости работы моделей в условиях реального применения.

Тем не менее, отмеченные недочёты не снижают общей научной ценности и технического уровня работы. Рекомендуется оценка "отлично" и присвоение выпускнику квалификации "магистр".

**Рецензент:**

Шах - 20.06.2025

Шахуро Владислав Игоревич, к. ф.-м. н.,
старший научный сотрудник АНО «Институт Искусственного Интеллекта»

Подпись рецензента, сотрудника АНО «Институт искусственного интеллекта»,
удостоверяю
HR Директор – начальник отдела
АНО «Институт искусственного интеллекта» Звонарева М.И.