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

ОТЗЫВ РЕЦЕНЗЕНТА

на дипломную работу Фролова Владимира Александровича
на тему: Стабилизация обучения слотовых распутанных представлений для реальных данных

Настоящая работа посвящена актуальной и перспективной области машинного обучения — разработке и исследованию новых подходов к обучению объектно-центричным представлениям в задачах компьютерного зрения. Актуальность работы обусловлена растущей потребностью в методах, способных эффективно декомпозировать сложные визуальные сцены на семантически осмысленные компоненты без использования аннотаций. Такие подходы имеют широкое применение в машинном обучении — от анализа изображений до обучения с подкреплением и генеративного моделирования.

Объём работы соответствует требованиям к магистерским диссертациям и составляет 37 страниц текста. Работа состоит из следующих разделов: введения, постановки задачи, обзора смежных работ, описания предложенного метода, экспериментального исследования и заключения. Таким образом, структура работы логична и соответствует стандартам научного исследования. Диссертация является завершённым научным исследованием, в котором реализованы все необходимые этапы.

Сфера исследования относится к объектно-центричному обучению представлениям с упором на задачу сегментации объектов без учителя. В работе рассматриваются современные архитектуры, такие как Slot Attention и Slot Mixture Module (SMM), и предлагаются модификация последней, а также модификации архитектуры модели, объединенные в новом методе обучения объектно-центричным представлениям, названном RAPID. В рамках экспериментальной части демонстрируется превосходство RAPID над такими моделями, как DINOSAUR и DINO SMM, на трёх наборах данных: COCO, PASCAL VOC и ClevrTex. Кроме того, проведены эксперименты по анализу эффективности предлагаемых модификаций, которые демонстрируют эффективность каждой из них.

Основная новизна работы заключается в стабилизированной версии Slot Mixture Module, где реализована повторная оценка апостериорных коэффициентов после каждого обновления параметров смеси. Эта модификация является оригинальной и технически обоснованной. В то же время, идея обучаемой инициализации слотов уже рассматривалась ранее в работе “Improving Object-Centric Learning with Query Optimization” (ICLR 2023), которая не упомянута в диссертации. Хотя архитектурный подход отличается (трансформер вместо фиксированных векторов), задачи совпадают, и следовало бы указать это направление как связанное или провести сравнение. Прочие модификации представляются скорее инженерными улучшениями, аналогичные подходы уже встречались в существующей литературе.

Также в качестве замечаний можно отметить, что изложение некоторых технических разделов представляется излишне сжатым — более подробное описание архитектурных решений и визуализация результатов способствовали бы лучшему восприятию. Кроме того, расширение экспериментальной части на видеоданные позволило бы более полно оценить универсальность предложенного подхода. Тем не менее, указанные недостатки не снижают общей положительной оценки работы.

Рекомендуется оценка «отлично» с присвоением выпускнику квалификации «магистр».

Рецензент,
PhD, н.с. AIRI
Управления по исследованиям
АНО «Институт искусственного интеллекта» Мунхоева М. Л.

Подпись рецензента, сотрудника АНО «Институт искусственного интеллекта»,
удостоверяю
Бизнес-партнер по управлению персоналом
АНО «Институт искусственного интеллекта» Трехлетова Ю. А.

АВТОНОМНАЯ НЕКОММЕРЧЕСКАЯ ОРГАНИЗАЦИЯ • ОГРН 1207700493978 • ИНН 7730261209 • РОССИЙСКАЯ ФЕДЕРАЦИЯ ГОРОД МОСКВА
"ИНСТИТУТ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА"