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 Изместьевой Ульяны Андреевны
на тему: «Применение генеративных моделей для аугментации обучающих данных в задачах сегментации изображений»

Выпускная квалификационная работа посвящена разработке автоматической системы генерации синтетических обучающих данных для задач компьютерного зрения. Основной идеей является использование диффузионных моделей, в частности Stable Diffusion v2, для создания новых изображений и соответствующих им масок объектов и карт глубины. Предложен оригинальный метод, который позволяет не просто изменять отдельные изображения, но и трансформировать целые наборы данных, сохраняя их статистические свойства. Важной частью исследования стало создание нового лосса — Regularized Inpainting Contrastive Loss (RIC Loss), который улучшает качество генерации объектов и предотвращает нежелательное затемнение участков вне маски.

В работе также реализовано использование модели Depth Anything v2 для оценки глубины сцены, что позволило обеспечить реалистичное размещение сгенерированных объектов в пространстве. Предложенный метод был протестирован на нескольких задачах, включая расширение набора данных InstSnowRoadDetection и обучение YOLO-моделей для сегментации дорожных дефектов.

Работа выполнена на высоком научном уровне и обладает рядом несомненных достоинств. В работе используются современные подходы, что говорит о хорошем знании современного состояния исследований в данной области. Особое внимание уделено оригинальности: предложен новый RIC-лосс, а также реализован метод расширения наборов данных, сохраняющий их статистические свойства. Работа сопровождается публично доступным кодом, что повышает её воспроизводимость и открывает возможности для дальнейшего развития. Экспериментальная часть проведена тщательно, включая тестирование на реальных задачах сегментации, что подтверждает эффективность предложенного подхода.

Замечания и направления для дальнейшей работы:

1. Текущая реализация системы добавляет только один объект на изображение, что ограничивает её применимость в сложных сценариях, требующих генерации нескольких объектов или учёта их взаимодействия.
2. Отсутствует механизм семантически обоснованного выбора позиции объекта — его расположение определяется случайным образом, что может привести к визуально некорректным результатам.

3. Несмотря на улучшение локальных метрик (CLIP Score, SSIM, PSNR), глобальные показатели качества (FID, LPIPS) ухудшились при использовании RIC-лосса, что указывает на возможное снижение визуальной достоверности и разнообразия сгенерированных изображений.
4. В перспективе необходимо автоматизировать выбор количества и расположения объектов на основе анализа сцены и карт глубины для более естественного и логически корректного размещения.
5. Требуется доработать функцию потерь или ввести дополнительную регуляризацию с целью улучшения визуального качества и разнообразия генерируемых изображений, чтобы повысить значения метрик FID и LPIPS.
6. Целесообразно провести дополнительные эксперименты на других предметных областях и с различными архитектурами моделей (например, DETR), чтобы подтвердить общую применимость предложенного метода.

Магистерская диссертация Изместьевой Ульяны Андреевны соответствует предъявленным требованиям и заслуживает оценки отлично. Считаю, что студент проявил высокий уровень самостоятельности, глубокое понимание предмета исследование и достоин присвоения квалификации магистра по направлению подготовки 09.04.01 «Информатика и вычислительная техника».

Рецензент, Кузнецов Андрей Владимирович

кандидат технических наук,
директор лаборатории FusionBrain, AIRI
АНО «Институт искусственного интеллекта»

Подпись рецензента, сотрудника АНО «Институт искусственного интеллекта»,

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

