

ПРОЦЕДУРА ОЦЕНКИ РОБАСТНОСТИ МЕТОДОВ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ В ТРЕХМЕРНЫХ СЦЕНАХ: АНАЛИЗ ПОДХОДОВ И МЕТРИК

Куркова Р.Е. (ИТМО), Попов М.Ф. (ИТМО)

Научный руководитель - доктор технических наук, профессор Колюбин С.А. (ИТМО)

Введение.

В современных исследованиях в области компьютерного зрения возникает все больший интерес к разработке систем, позволяющим сопоставлять трехмерные сцены и объекты с текстовыми запросами произвольной формы. Данный доклад фокусируется на разработке процедуры оценки робастности методов семантической сегментации в трехмерных сценах, с акцентом на выделении наиболее репрезентативных и показательных метрик, а также подготовке данных для тестирования моделей, включая выборку датасетов и изменение их конфигураций в среде Habitat-Sim[1].

Основная часть.

В ходе работы были проанализированы различные методы семантической сегментации, включая ConceptGraphs[2], OpenMask3D[3], Context-Aware Entity Grounding with Open-Vocabulary 3D Scene Graphs[4], OVIR-3D[5], PLA[6] и другие. Оценка робастности каждого из этих методов включала анализ их применимости к различным сценариям с использованием стандартных метрик, таких как средняя точность (AP), частотно-взвешенное IoU (f-mIoU) и другие метрики, описанные в работе. Введение метрики Grounding Success Rate дополняет анализ робастности методов семантической сегментации, фокусируясь на точности интерпретации языковых текстовых запросов в контексте создания трехмерной семантической карты. Таким образом, в рамках исследования проведен анализ различных подходов и метрик оценки робастности методов семантической сегментации.

Дальнейшее развитие работы предполагает проведение углубленного анализа преимуществ и ограничений каждой из рассматриваемых метрик в контексте их применимости к различным сценариям. Это включает анализ их робастности к изменениям в данных, таким как различия в освещении, углах обзора или наличии различных объектов в сцене.

Выводы.

Были созданы подробные рекомендации по проведению тестирования методов семантической сегментации, которые включают углубленный анализ преимуществ и ограничений каждой рассмотренной метрики, а также их применимость к различным сценариям и условиям. Результаты исследования подчеркивают важность разработки процедур оценки робастности методов семантической сегментации, особенно с условием выбора наиболее подходящих метрик и баз данных для тестирования.

Список используемых источников:

1. Manolis Savva, Abhishek Kadian, Oleksandr Maksymets, Yili Zhao, Erik Wijmans, Bhavana Jain, Julian Straub, Jia Liu, Vladlen Koltun, Jitendra Malik, Devi Parikh, & Dhruv Batra. Habitat: A Platform for Embodied AI Research. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). — 2019.

2. Gu Qiao, Kuwajerwala Alihusein, Morin Sacha, Jatavallabhula Krishna Murthy, Sen Bipasha, Agarwal Aditya, Rivera Corban, Paul William, Ellis Kirsty, Chellappa Rama, Gan Chuang, de Melo Celso Miguel, Tenenbaum Joshua B., Torralba Antonio, Shkurti Florian, and Paull Liam. ConceptGraphs: Open-Vocabulary 3D Scene Graphs for Perception and Planning. arXiv. — 2023
3. Takmaz, A., Fedele, E., Sumner, R., Pollefeys, M., Tombari, F., & Engelmann, F. OpenMask3D: Open-Vocabulary 3D Instance Segmentation. In Advances in Neural Information Processing Systems (NeurIPS). — 2023.
4. Haonan Chang, Kowndinya Boyalakuntla, Shiyang Lu, Siwei Cai, Eric Pu Jing, Shreesh Keskar, Shijie Geng, Adeeb Abbas, Lifeng Zhou, Kostas Bekris, & Abdeslam Boularios. Context-Aware Entity Grounding with Open-Vocabulary 3D Scene Graphs. In the 7th Annual Conference on Robot Learning. — 2023.
5. Shiyang Lu, Haonan Chang, Eric Pu Jing, Abdeslam Boularias, & Kostas Bekris. OVIR-3D: Open-Vocabulary 3D Instance Retrieval Without Training on 3D Data. In 7th Annual Conference on Robot Learning. — 2023.
6. Ding, R., Yang, J., Xue, C., Zhang, W., Bai, S., & Qi, X. PLA: Language-Driven Open-Vocabulary 3D Scene Understanding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2023.