

К вопросу о zero-shot построении семантического графа сцены для визуальной локализации робота

Tuesday, 20 May 2025 14:39 (12 minutes)

Разработан zero-shot алгоритм генерации семантического графа сцены (SG) для одиночных RGB-изображений, предназначенный для визуальной локализации роботов. Пайплайн RAM++ → YOLO-World → Mobile-SAM извлекает маски объектов; далее пиплр C-расширение формирует граф контактов, а jina-CLIP-v1 выбирает предикаты по косинусному сходству без дообучения. На Visual Genome достигнут mR@100 = 8.9 % (задача PredCls) – уровень IMP (2017). Подготовлен открытый обзор 22 SOTA-статей. Пайплайн был протестирован на новом датасете Sber Office.

Primary author: ANDRYUSHENKO, Solomon (MIPT)

Presenter: ANDRYUSHENKO, Solomon (MIPT)

Session Classification: 20-Машинное обучение и нейросети

Track Classification: Машинное обучение и нейросети