
Локализация по графу сцены

Кондрашов Тимофей, 3 курс

Научный руководитель: Юдин Дмитрий Александрович

Ментор: Мелехин Александр Алексеевич

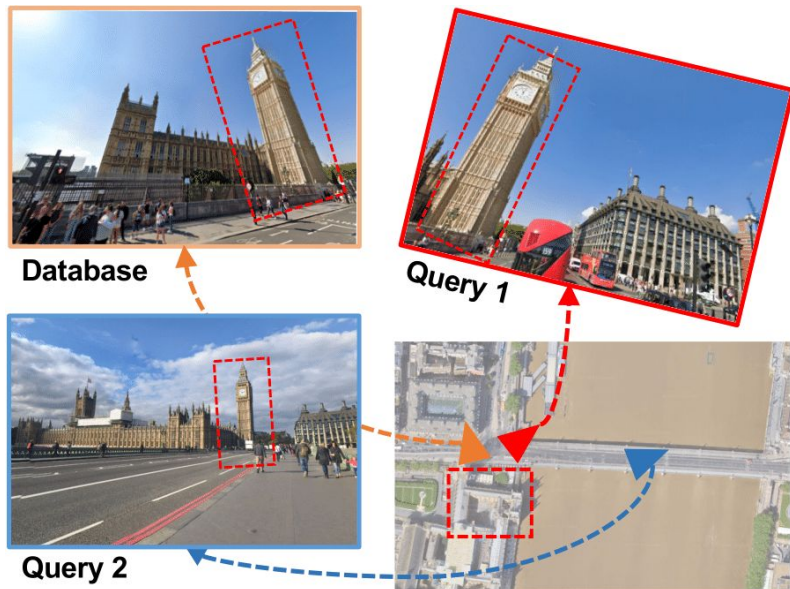
Центр когнитивного моделирования МФТИ, 18 мая, 2025

Поставленная задача



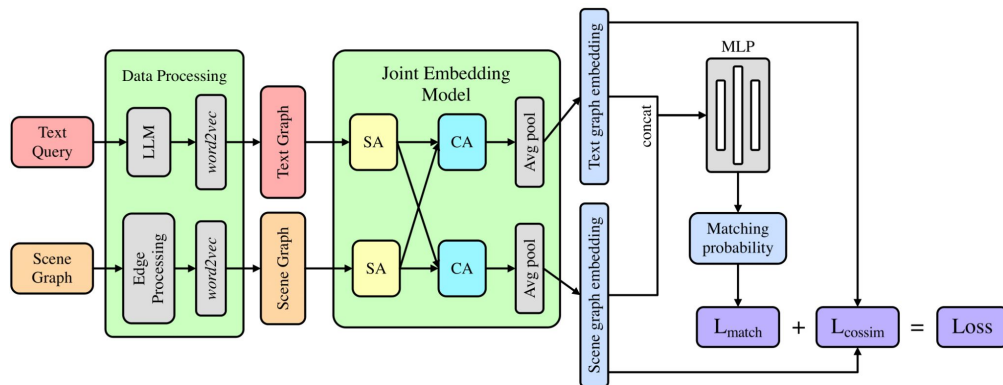
Локализация робота - это задача определения положения и ориентации робота в заданной системе координат на основе сенсорных данных и априорной информации о среде.

Встречающиеся проблемы

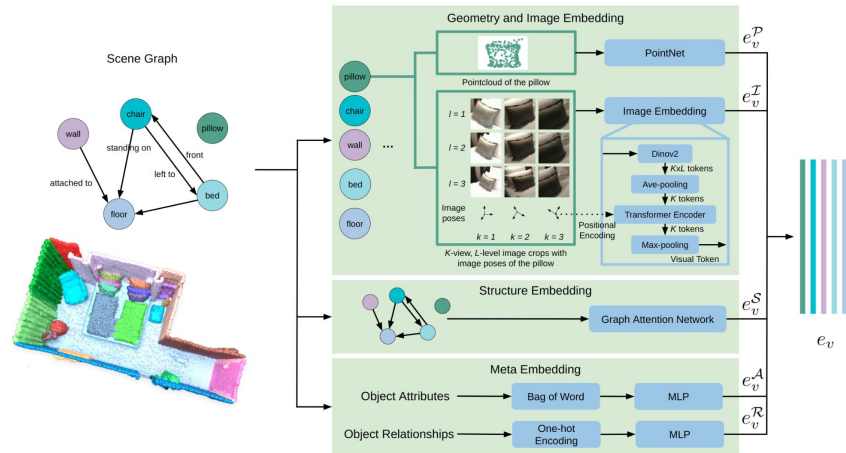
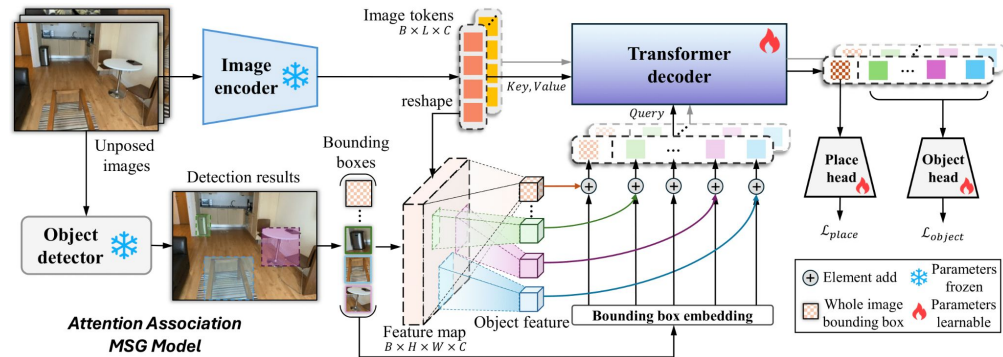


- Изменение ракурса камеры
- Нестатическое окружение
- Ограниченные вычислительные ресурсы

Срез научных работ



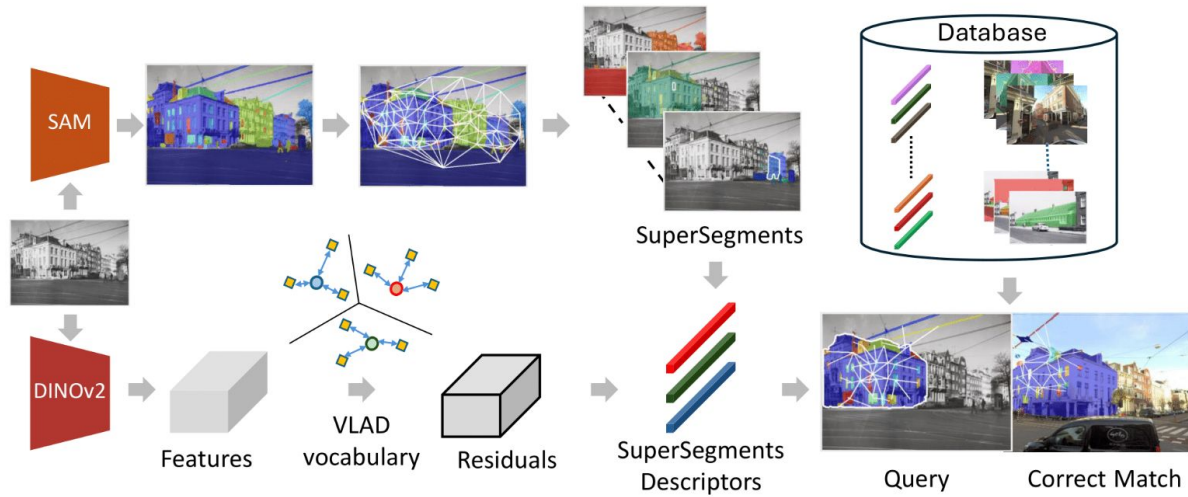
Where am I? Cross-View Geo-localization with Natural Language Descriptions. [1]



SceneGraphLoc: Cross-Modal Coarse Visual Localization on 3D Scene Graphs. [2]

Multiview Scene Graph. [3]

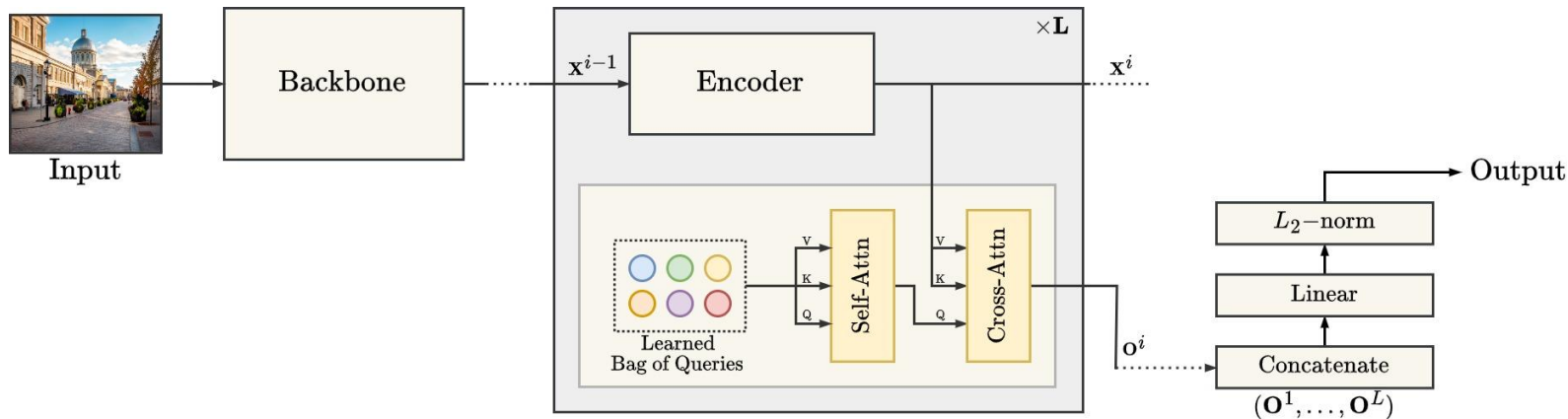
SegVLAD



SegVLAD pipeline. [4]

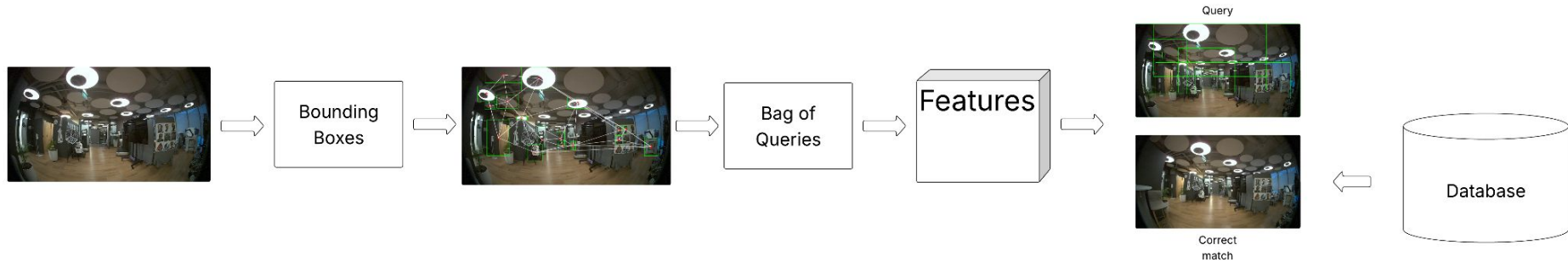


Visual Features for Place Recognition: Bag of Queries

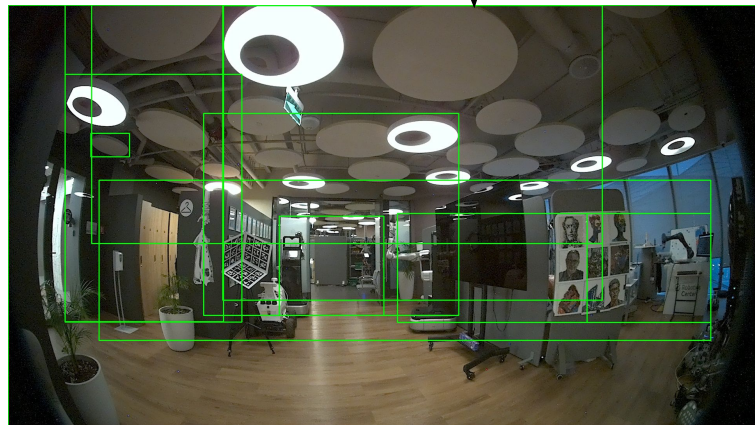
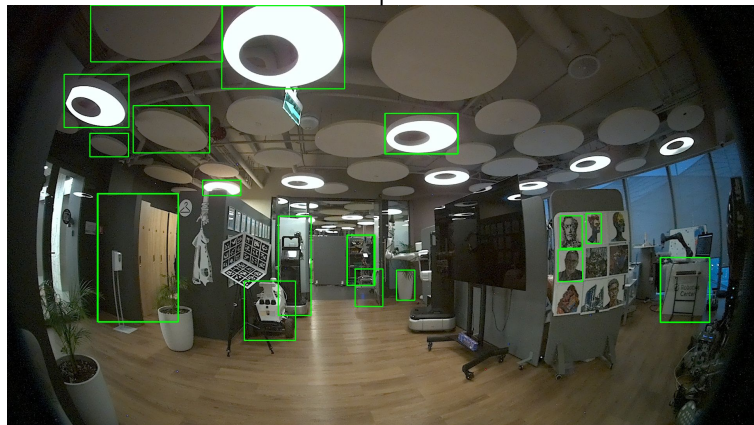
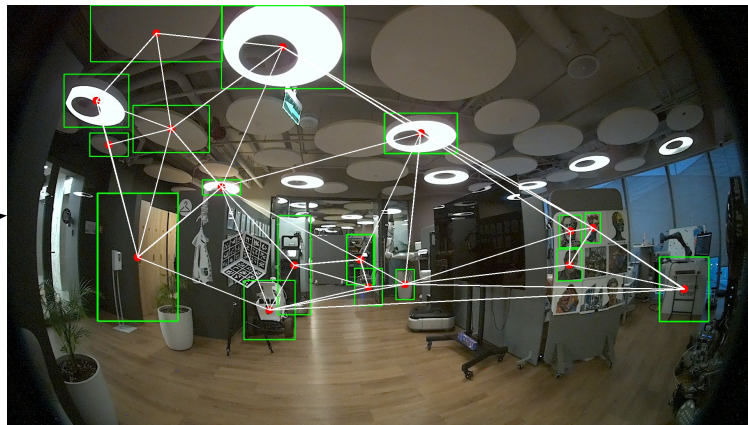


Bag of Queries pipeline [5]

SegVLAD + BoQ = SegBoQ



СуперСегменты в SegBoQ



Демонстрация работы SegVoQ

Запрос



Top 1



Top 2



Top 3

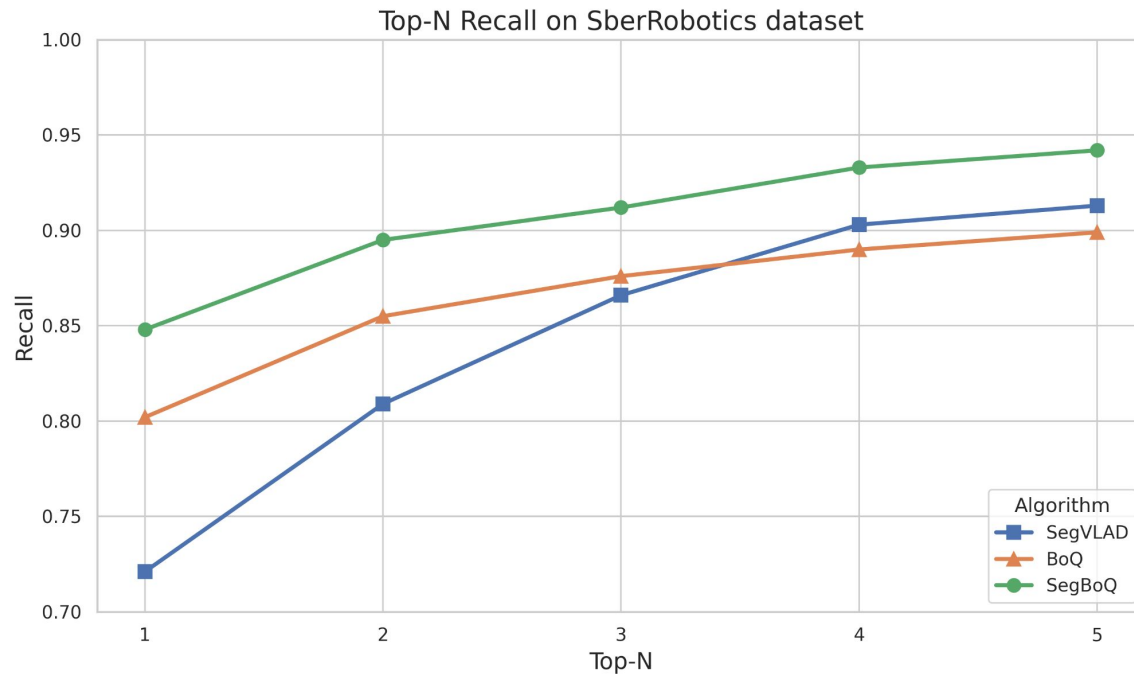


Параметры SegBoQ

Метод	Кол-во ближайших сегментов	Кол-во сегментов на картинку	Радиус суперсегментов	Метод агрегации	Источник масок	Recall@		
						@1	@2	@3
SegBoQ	30	10	1	Sum	Sam	81.3%	87.3%	89.2%
SegBoQ	30	5	1	Sum	Sam	83.6%	86.6%	88.5%
SegBoQ	30	10	1	Max	Sam	80.0%	83.9%	88.0%
SegBoQ	30	5	1	Max	Sam	81.1%	86.5%	89.6%
SegBoQ	30	10	1	Sum	RYS*	83.3%	89.5%	90.7%
SegBoQ	30	5	1	Sum	RYS*	82.3%	88.6%	90.4%
SegBoQ	30	10	1	Max	RYS*	84.8%	89.5%	91.2%
SegBoQ	30	5	1	Max	RYS*	83.8%	89.4%	92.5%
SegBoQ	30	10	1	Mean	RYS*	79.1%	86.2%	90.4%
SegBoQ	30	10	1	SoftMax	RYS*	79.6%	86.6%	90.6%
SegBoQ	30	10	1	Top3	RYS*	83.8%	88.5%	90.9%
SegBoQ	30	30	1	Max	RYS*	83.1%	89.2%	91.1%
SegBoQ	50	10	1	Max	RYS*	84.8%	89.5%	91.2%
SegBoQ	50	10	1	Sum	RYS*	82.3%	88.4%	90.2%
SegBoQ	30	10	2	Max	RYS*	82.0%	86.3%	89.4%
SegBoQ	30	10	2	Sum	RYS*	80.4%	84.5%	87.8%

*RAM++ & YOLO-World & Mobile-SAM-v2

Результаты тестирования



References

- [1] Ye, Junyan, et al. "Where am I? Cross-View Geo-localization with Natural Language Descriptions." arXiv preprint arXiv:2412.17007 (2024).
 - [2] Miao, Yang, et al. "SceneGraphLoc: Cross-Modal Coarse Visual Localization on 3D Scene Graphs." European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024.
 - [3] Zhang, Juexiao, et al. "Multiview Scene Graph." Advances in Neural Information Processing Systems 37 (2025): 17761-17788.
 - [4] Garg, Kartik, et al. "Revisit Anything: Visual Place Recognition via Image Segment Retrieval." European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024.
 - [5] Ali-Bey, Amar, Brahim Chaib-draa, and Philippe Giguère. "BoQ: A place is worth a bag of learnable queries." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
-