

# Scene Graph Construction

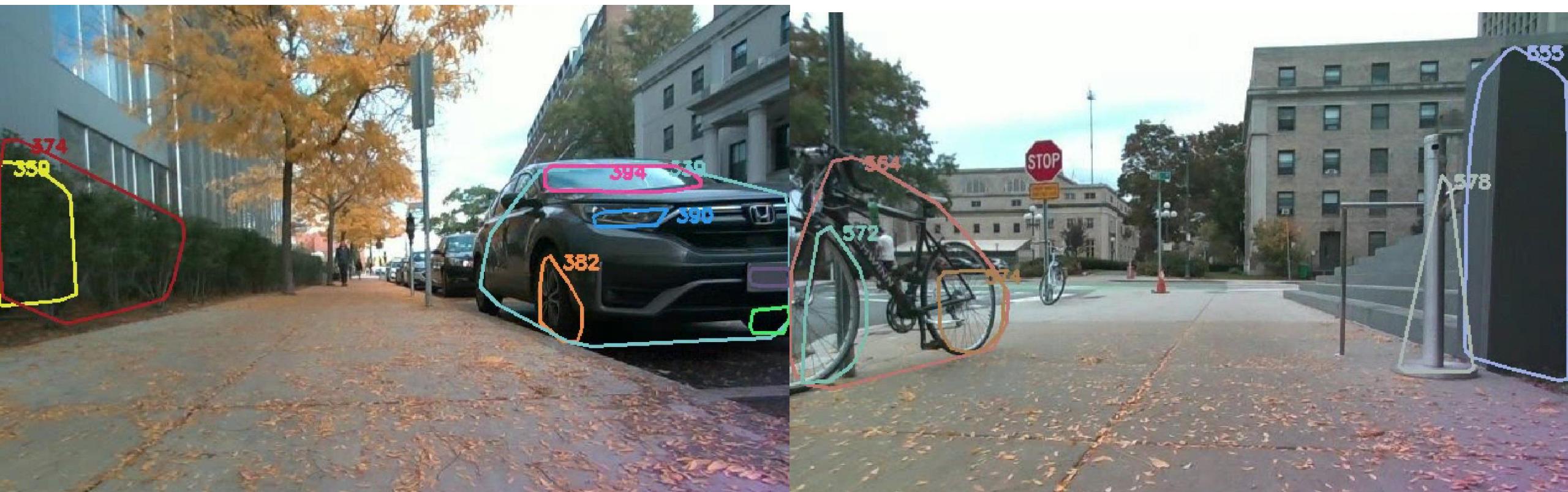
Andryushenko Solomon, MIPT

Cognitive Modeling, 2025



# Roman

- Launched on the prebuilt dataset in a docker container.



# Ram++ & YOLO & Mobile-SAM-v2 pipeline

- Created a full pipeline.



YOLO



MSAM

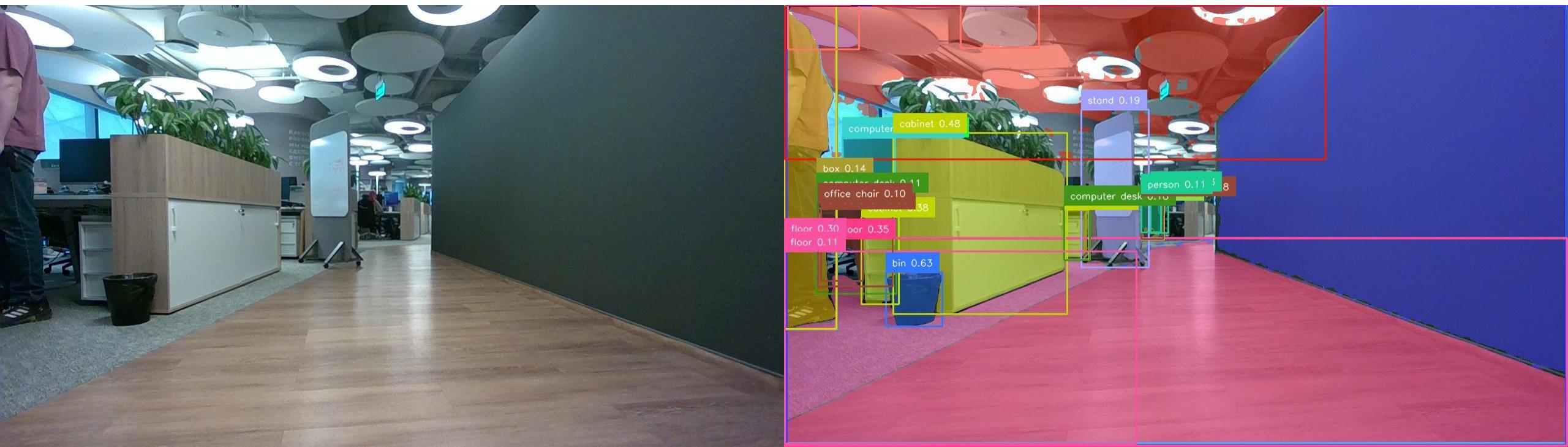


Couch, dog, painting, dining table, ...

RAM++

# Sber Dataset

- Made a run on a new sber dataset (780 RGBD images) using Rerun.io
- Forwarded the results to a teammate (works on objects tracking).



# Graph Construction

- Finding objects with “touching” masks.
- Blurring / Cutting the initial image for each two masks.
- Generating CLIP embeddings.

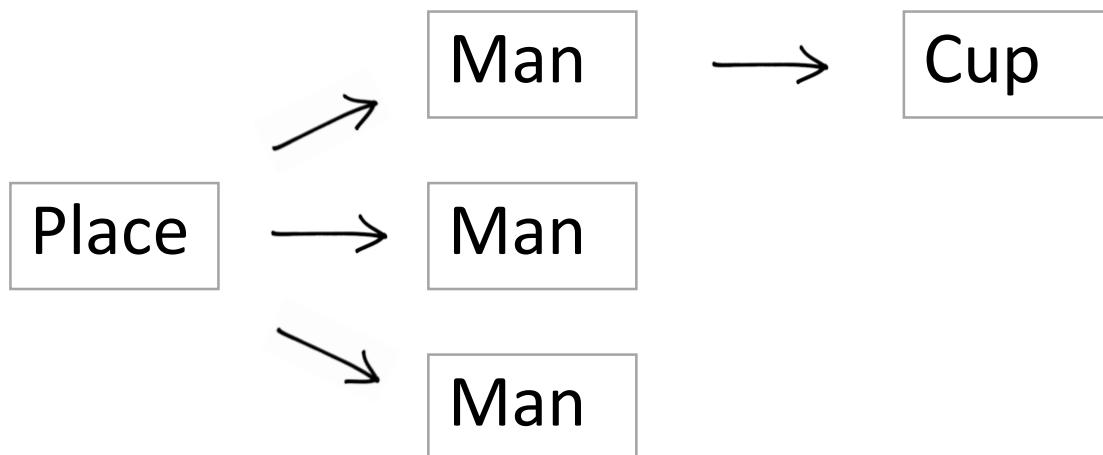


Find a place with a dining table on the floor.

# Query Graph

- Finds objects and relations between them in a query.
- Builds a graph.

Find a place where were 3 men and one of them was holding a cup.



Thank you for your attention