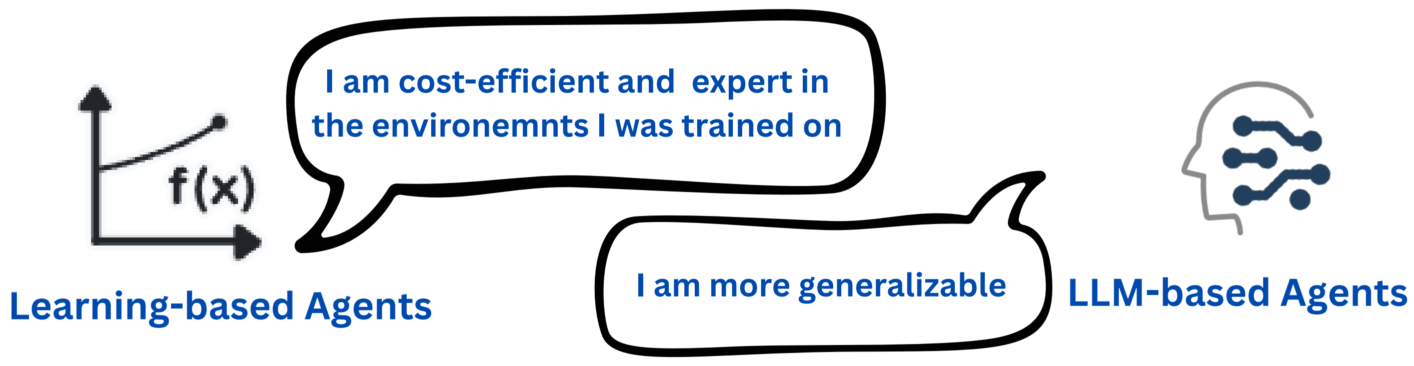
Advancing Object-Goal Navigation through LLM-Enhanced Object Affinities Transfer

Mengying Lin*, Shugao Liu*, Dingxi Zhang, Yaran Chen, Zhaoran Wang, Haoran Li✉️, Dongbin Zhao
mlin365@gatech.edu

Research Questions

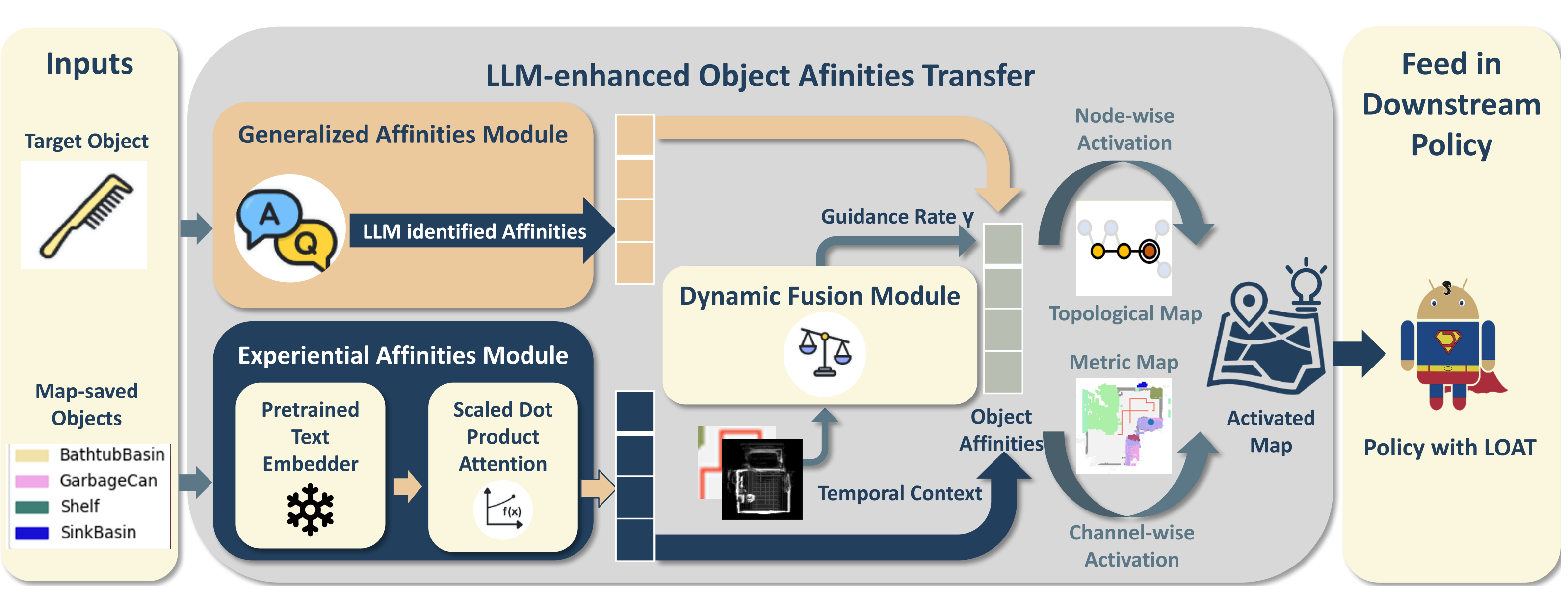
In the world of object-goal navigation, there are two main types of agents:



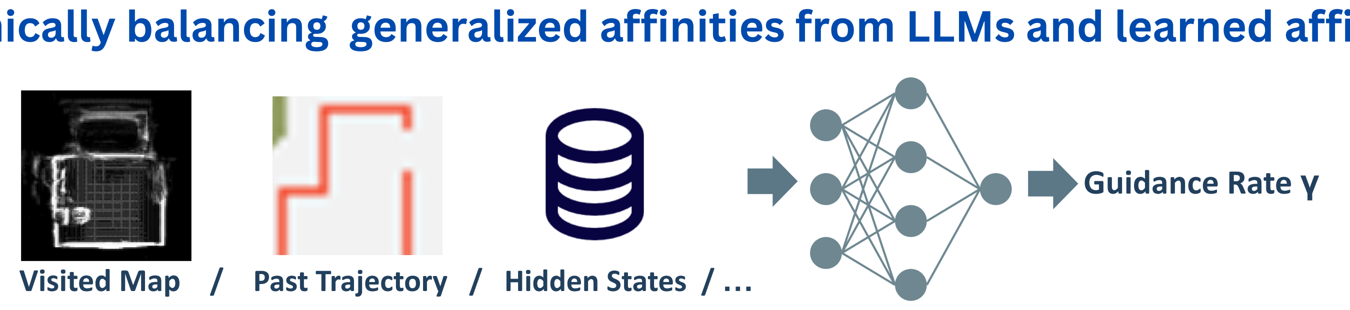
Can we get LLM-level smarts without the LLM-level bill while maintaining specificity from training data?

- What enables the generalization of LLM?
  --→ Priors about object affinities
- How can we benefit from both such priors and specificity from data?
  --→ Integrate the priors into learning-based systems
- How can we do so seamlessly without extra architecture engineering?

LLM-Enhanced Object Affinities Transfer

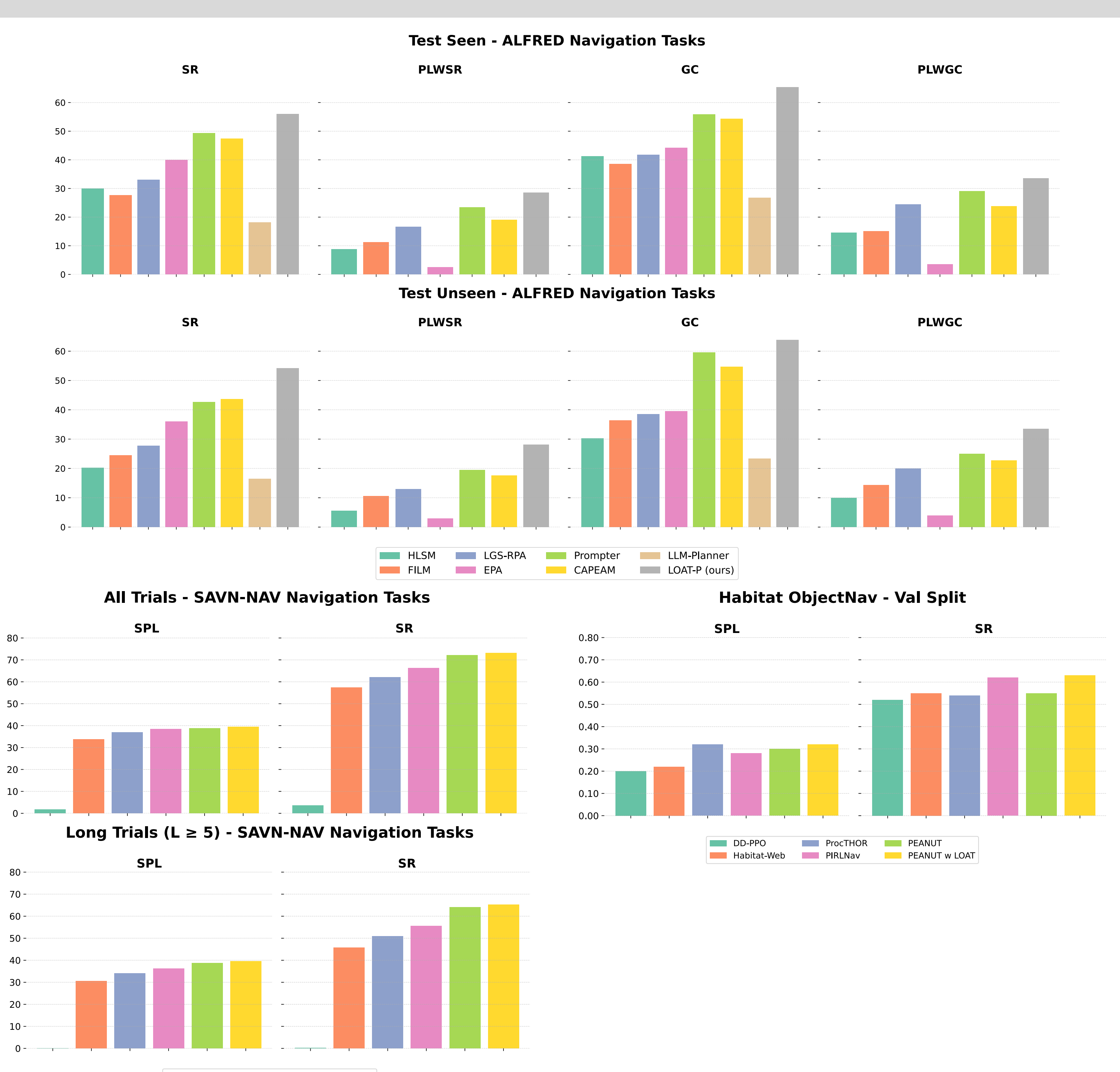


Dynamically balancing generalized affinities from LLMs and learned affinities.



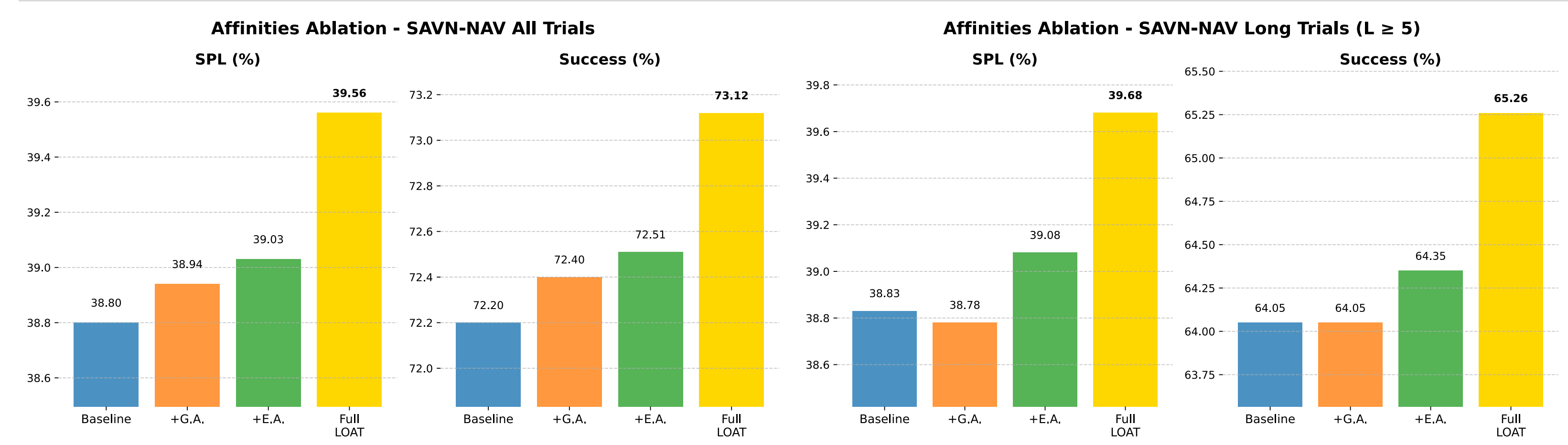
Feed available historical contexts into the dynamic fusion module.

Benchmark Results



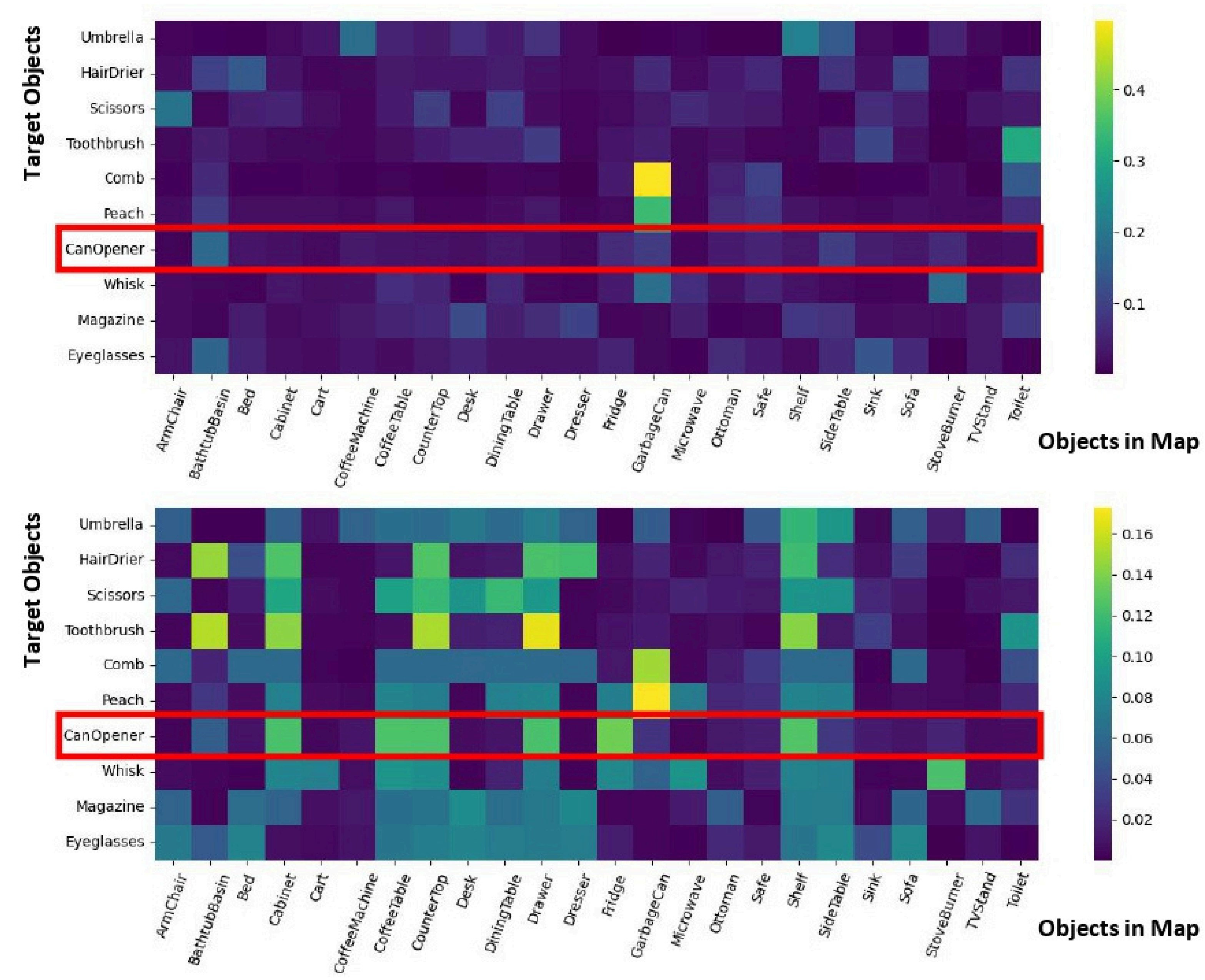
LOAT module boosts the performance of both metric-map based and graph-based polices in all the benchmarks.

Ablation over Sources of Affinities



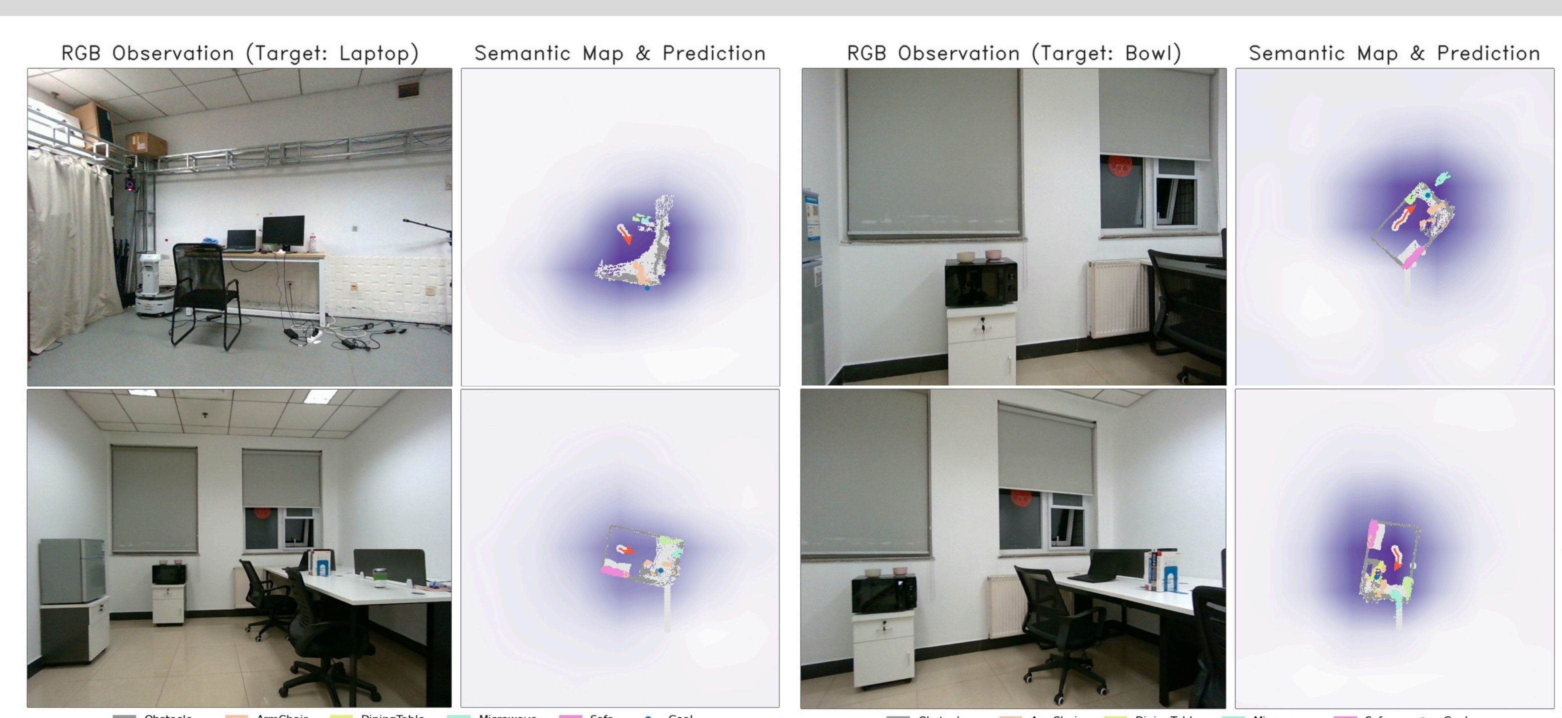
The dynamic combination of two sources of object affinities achieves the best performance.

OOD Category Evaluation



Without the generalized affinities provided by LLMs, the policy cannot reliably infer the locations of unseen objects, whereas incorporating such priors leads to more desirable predictions.

Real World Evaluation



We demonstrate successful zero-shot sim-to-real transfer of navigation policies, localizing all target objects without additional training.

Key Takeaways

- We extract object affinities from LLMs to reduce costly step-by-step queries.
- LOAT serves as a plug-and-train module that injects generalized affinities into policies without additional architectural changes.
- Combining both sources of affinities enables policies to generalize effectively across seen and unseen cases.