

Synthetic vs. Real Training Data for Visual Navigation

Lauri Suomela · German F. Torres · Sasanka Kuruppu Arachchige
Harry Edelman · Joni-Kristian Kämäräinen

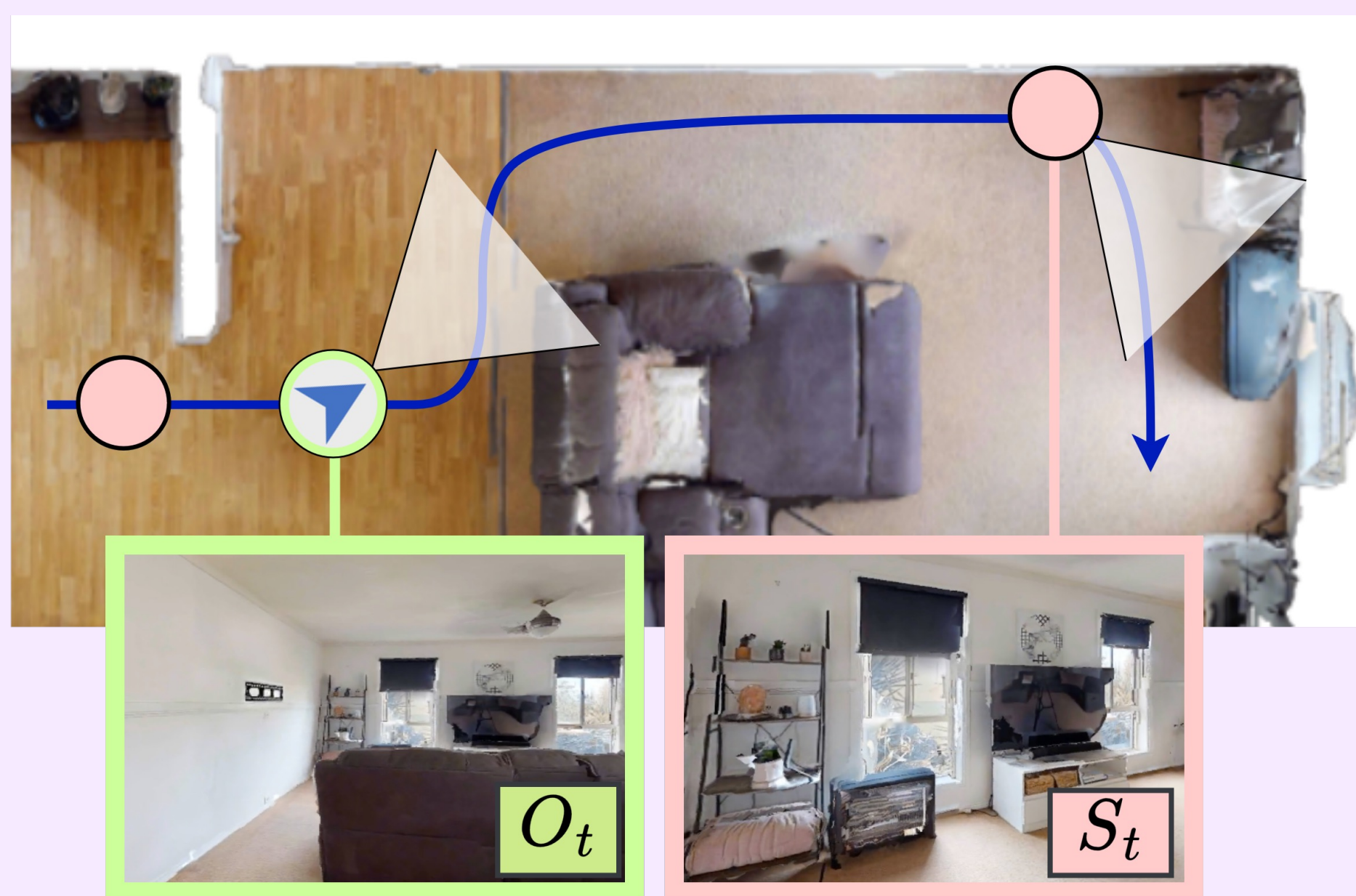
1 Simulate or collect real data?

How well does a policy trained in simulation compare to one trained with real data?



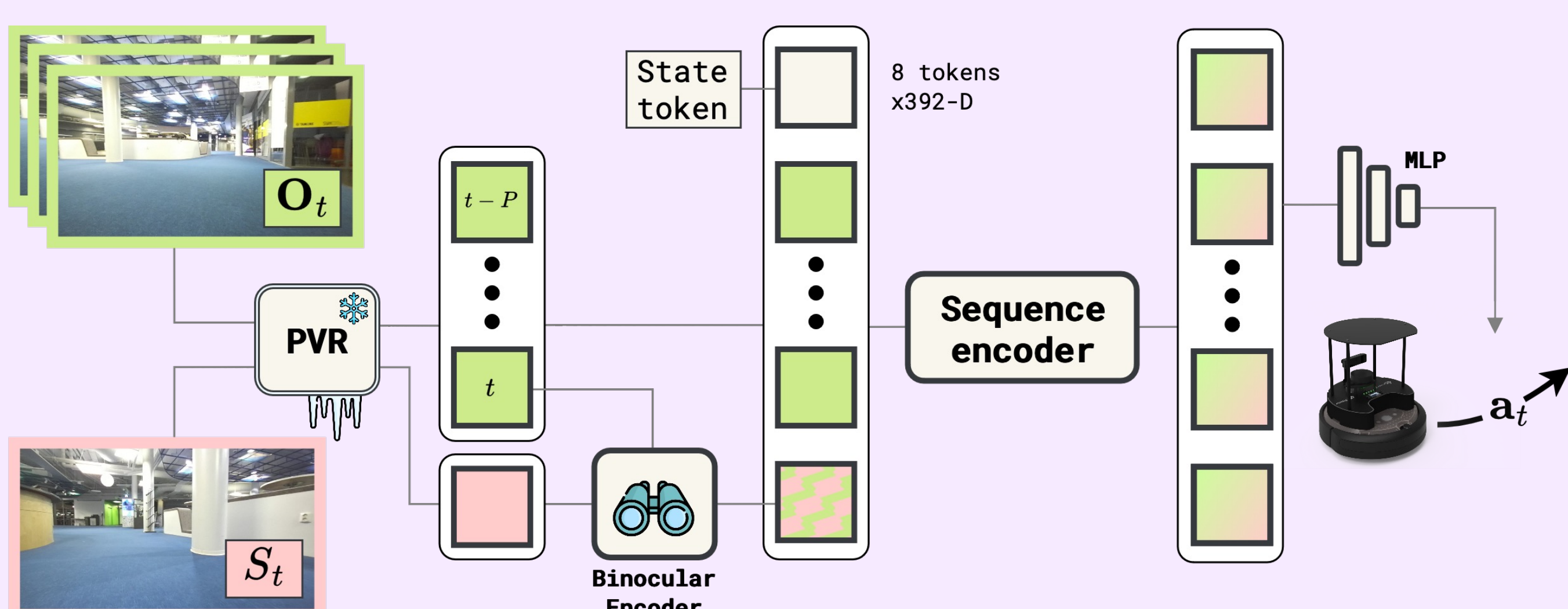
2 Topological navigation

We conduct the study in the context of topological navigation



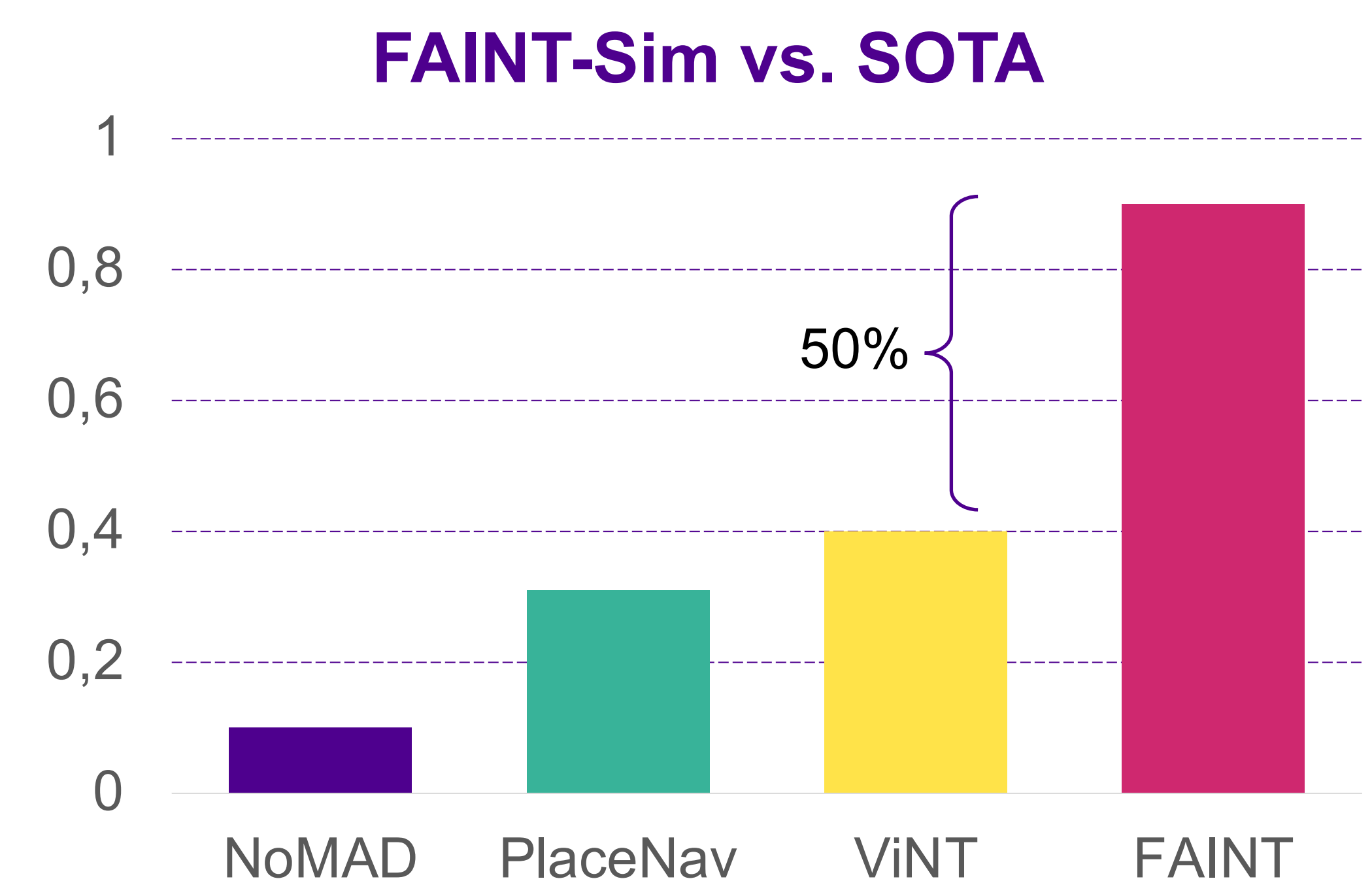
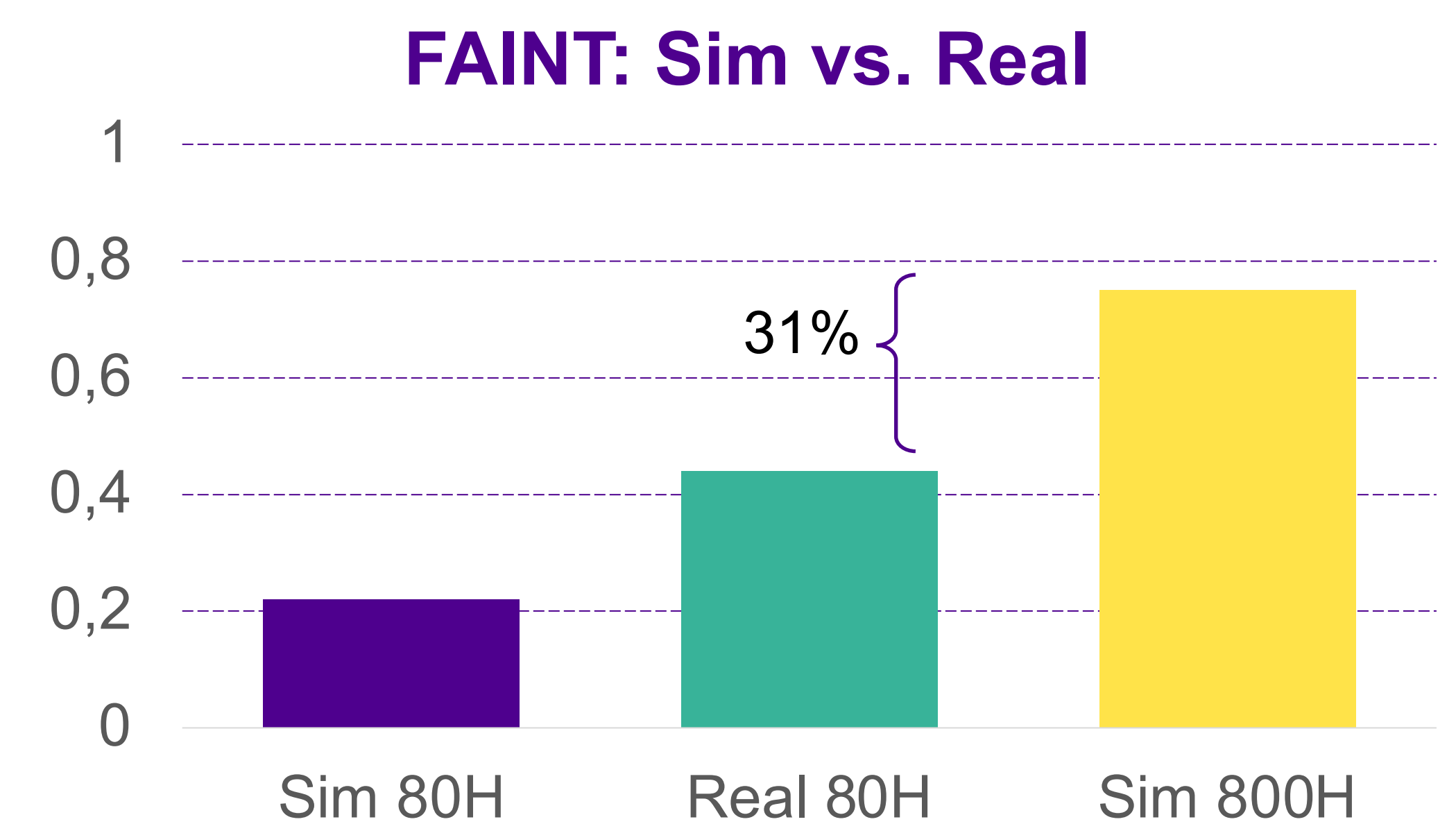
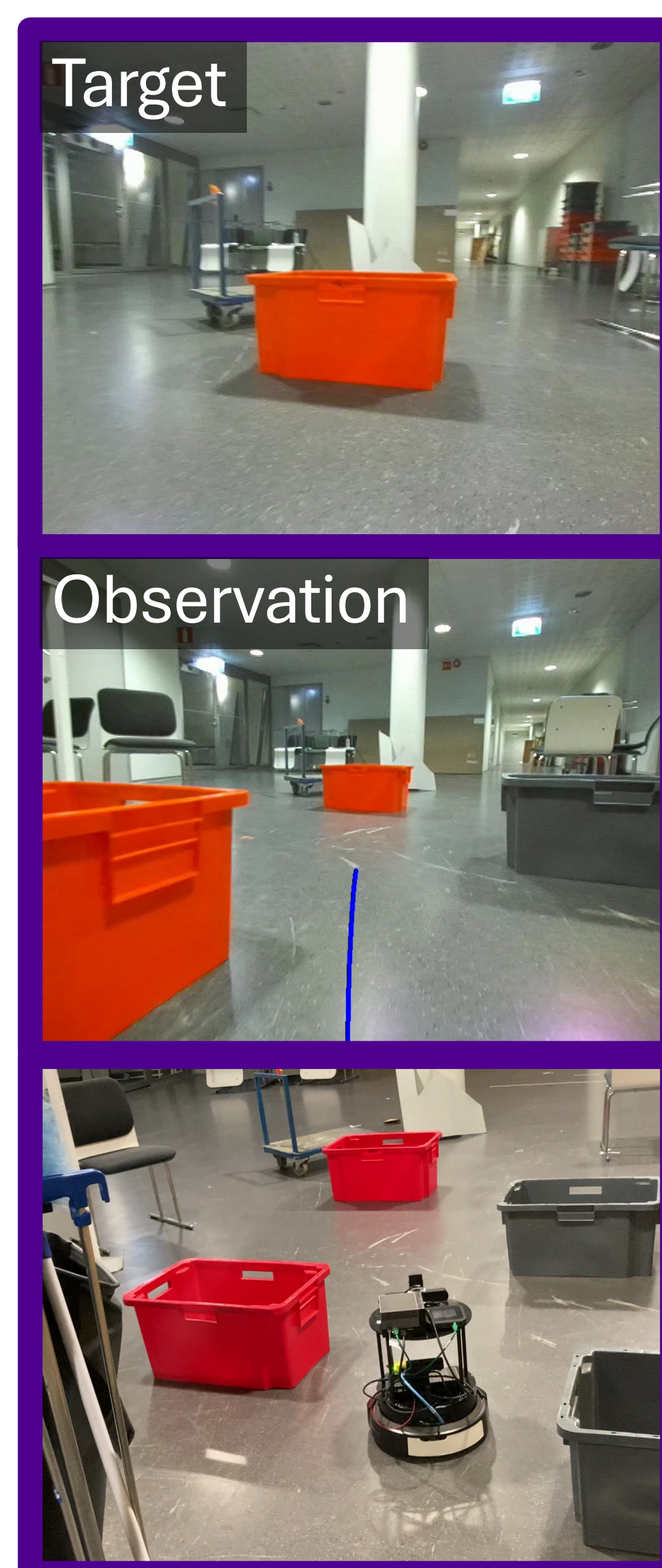
3 FAINT 🌀

A navigation policy suitable for training with real or synthetic data



4 Experiments

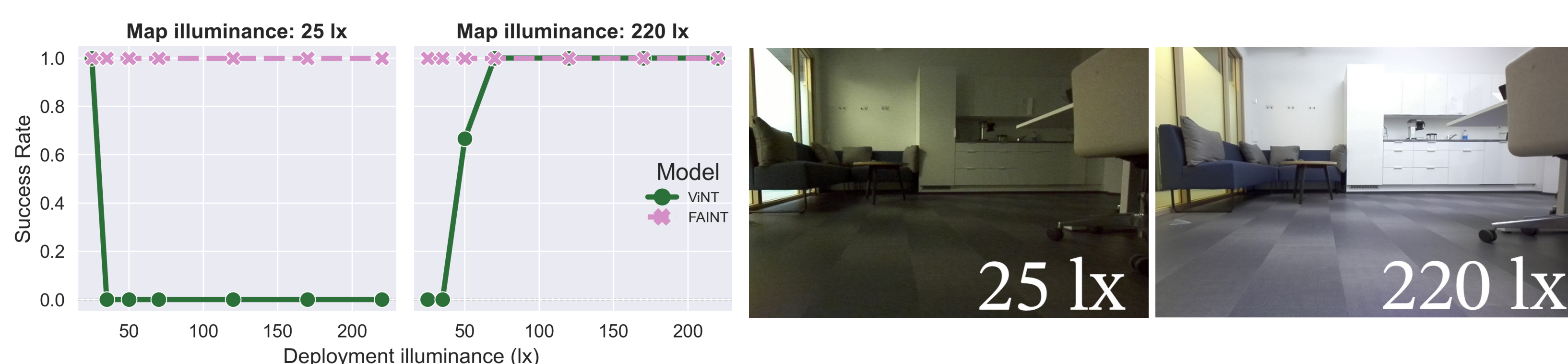
50 hours of evaluation in the real world, with **success rate** as metric



What enables Sim2Real?

Method	Encoder type	Encoder frozen	Mode	Real $n = 30$	Sim 2500
FAINT	Theia CDDSV	✓	BC	0.23	0.87
	Theia CDDSV	✓	DAGger	0.80	0.91
	Theia CDIV	✓	DAGger	0.60	0.91
	EfficientNet-B0	✓	DAGger	0.13	0.79
ViNT [8]	EfficientNet-B0	✗	DAGger	0.40	0.87

Besides Sim2Real gap, FAINT is robust to illumination change



5 Conclusion

We find that a simulation-trained policy can match its counterpart trained with real data

Factors crucial for high real-world performance:

- Visual encoders pretrained with diverse data
- On-policy learning

Future work: Combine real data with on-policy data from sim?

Contact: lauri.a.suomela@tuni.fi

<https://lasuomela.github.io/faint/>

Project page:

