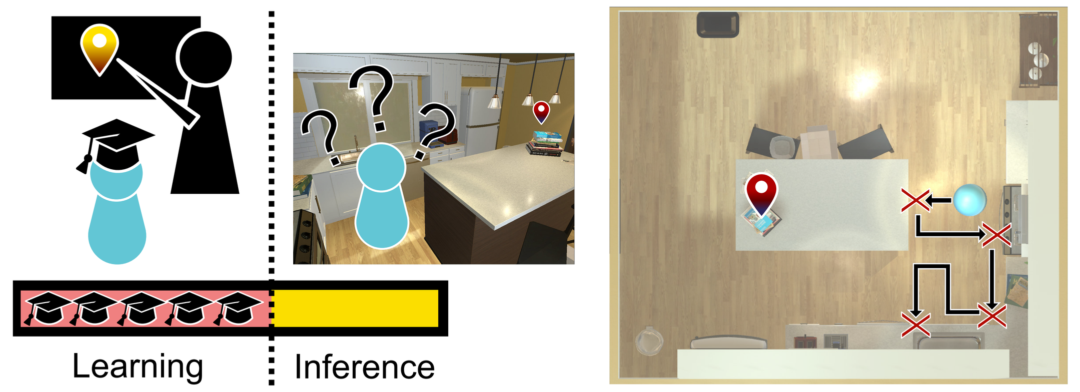
Learning to Learn How to Learn: Self-Adaptive Visual Navigation Using Meta-Learning

Mitchell Wortsman[1], Kiana Ehsani[2], Mohammad Rastegari[1,3], Ali Farhadi[1,2,3], Roozbeh Mottaghi[1]

[1] PRIOR @ Allen Institute for AI   [2] University of Washington   [3] XNOR.AI

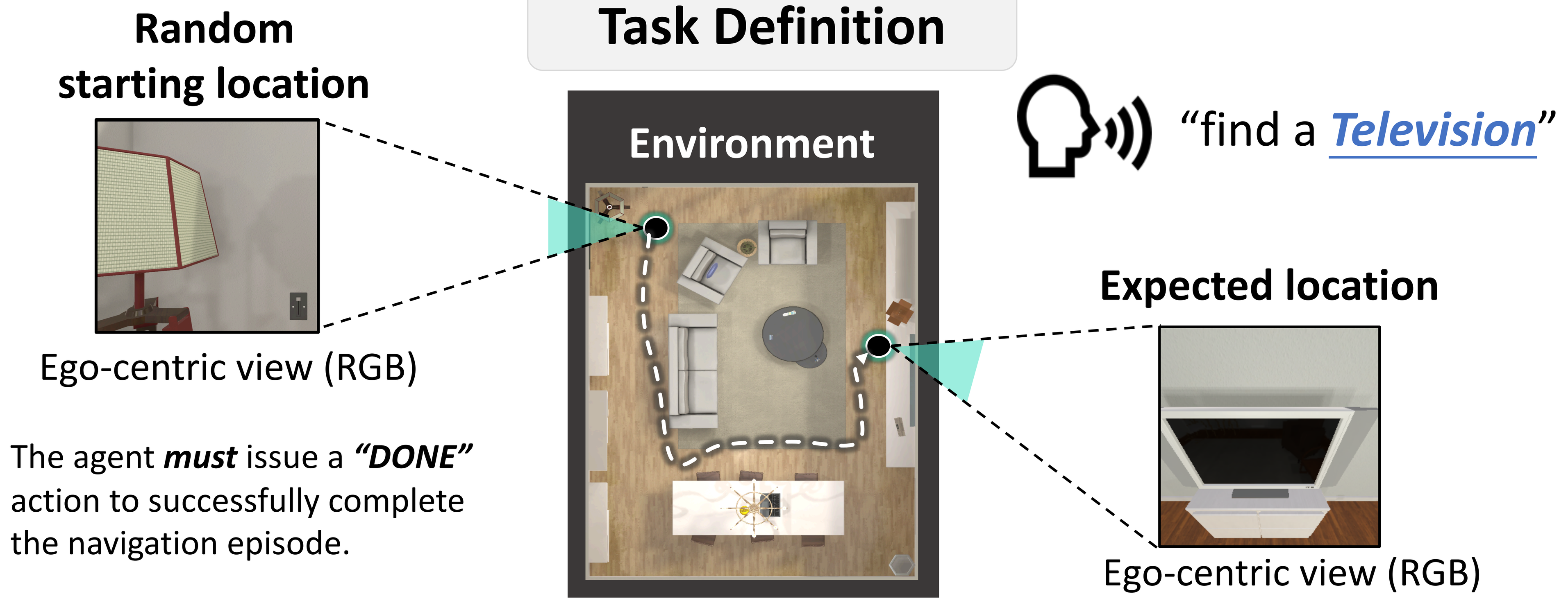
MOTIVATION: In reality there is no clear distinction between training and inference: **We learn as we perform.**



Traditional navigation approaches freeze the model during inference.

We introduce a self-adaptive agent for visual navigation (SAVN). SAVN learns how to adapt via self-supervised interaction with the environment.

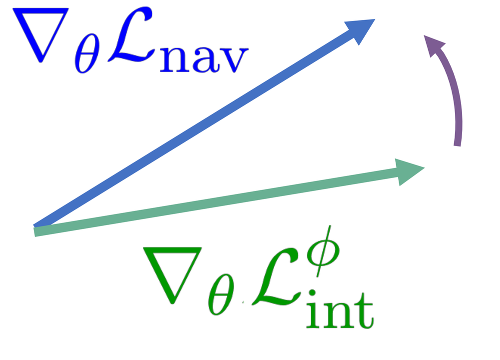
Task Definition



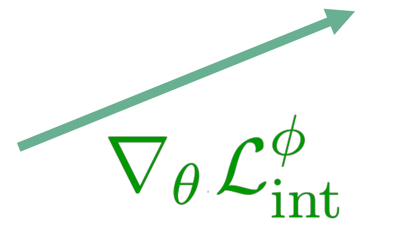
The agent ***must*** issue a ***"DONE"*** action to successfully complete the navigation episode.

Goal

- We learn a self-supervised interaction loss $\mathcal{L}_{\text{int}}^{\phi}$ to help minimize the supervised navigation loss $\mathcal{L}_{\text{nav}}$.
- During training we maximize the similarity between the gradients we receive from $\mathcal{L}_{\text{int}}^{\phi}$ and $\mathcal{L}_{\text{nav}}$ – we may then continue "learning" when there is no supervision.

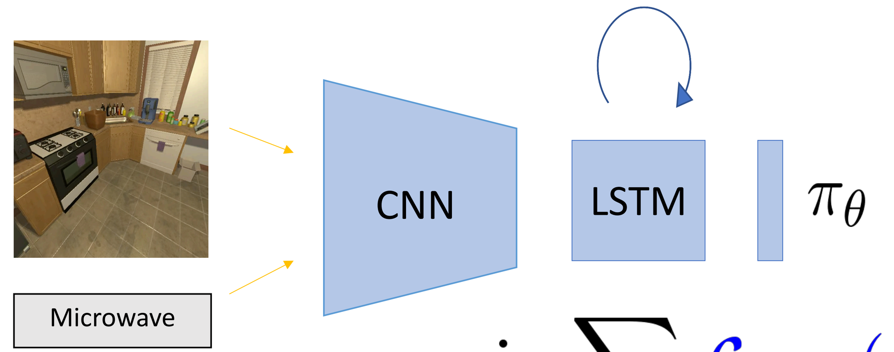


Training (first-order approx.)



Inference

Learning



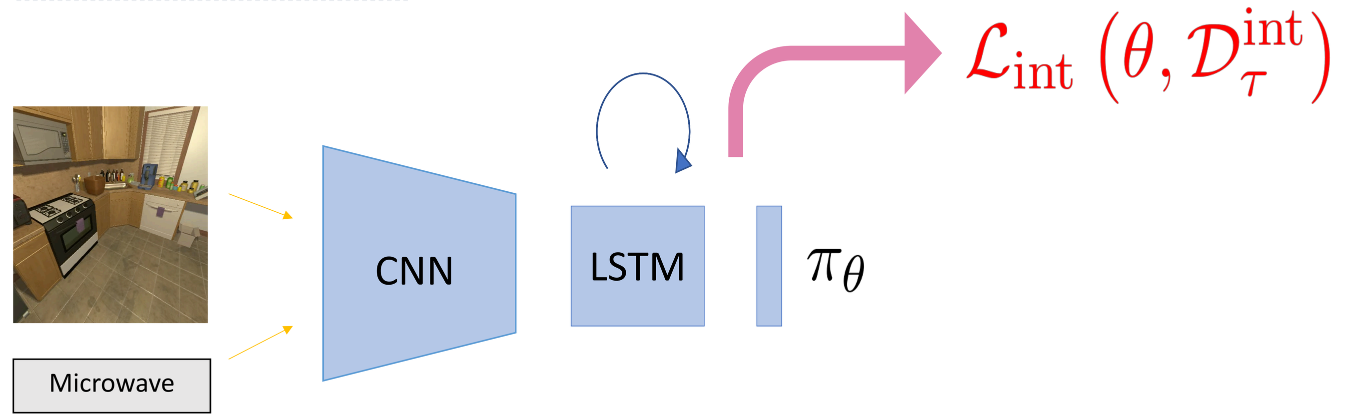
$$\min_{\theta} \sum_{\tau} \mathcal{L}_{\text{nav}}\left(\theta, \mathcal{D}_\tau\right)$$

Objective: Learn the model parameters which minimize the supervised navigation loss.

π_θ	The ***agent's policy:*** A distribution over the actions an agent should take next.
$\mathcal{L}_{\text{nav}}$	A traditional ***supervised navigation loss***, e.g. A3C Loss [1].
θ	Parameters of the model.
$\mathcal{D}_\tau$	Navigation trajectory.

Learning to Learn

We rely on the meta-learning approach detailed by the MAML algorithm [2].



$$\min_{\theta} \sum_{\tau} \mathcal{L}_{\text{nav}}\left(\theta - \nabla \mathcal{L}_{\text{int}}\left(\theta, \mathcal{D}_\tau^{\text{int}}\right), \mathcal{D}_\tau^{\text{nav}}\right)$$

Objective: Succeed at navigation after *adapting* to the environment with a gradient from the *self-supervised interaction loss*.

$$\mathcal{D}_\tau = \left(\mathcal{D}_\tau^{\text{int}}, \mathcal{D}_\tau^{\text{nav}}\right)$$

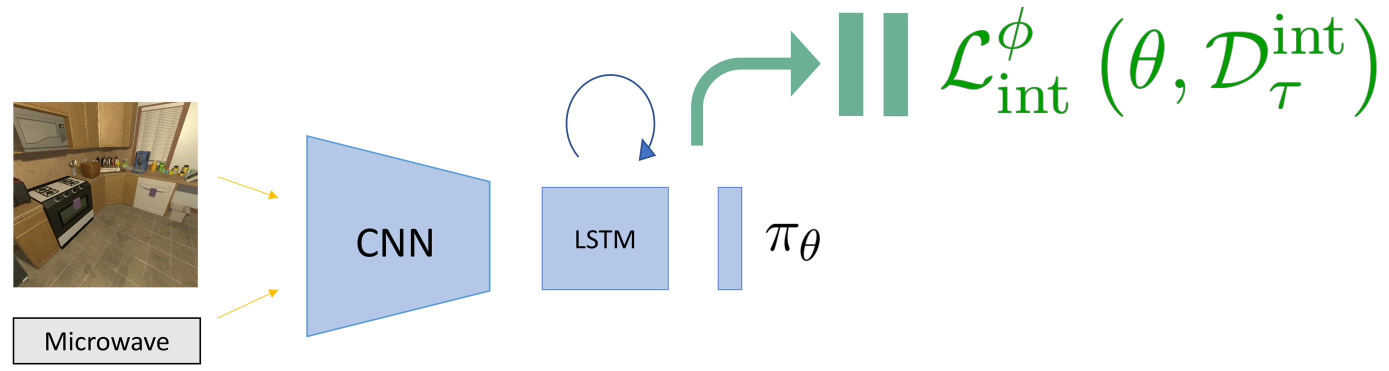
The trajectory is decomposed into an ***interaction phase*** and ***navigation phase.***

$\mathcal{L}_{\text{int}}$	Self-Supervised ***interaction loss.***

$$\theta - \nabla \mathcal{L}_{\text{int}}\left(\theta, \mathcal{D}_\tau^{\text{int}}\right)$$

Adapted model parameters.

Learning to Learn how to Learn



$$\min_{\theta, \phi} \sum_{\tau} \mathcal{L}_{\text{nav}}\left(\theta - \nabla \mathcal{L}_{\text{int}}^{\phi}\left(\theta, \mathcal{D}_\tau^{\text{int}}\right), \mathcal{D}_\tau^{\text{nav}}\right)$$

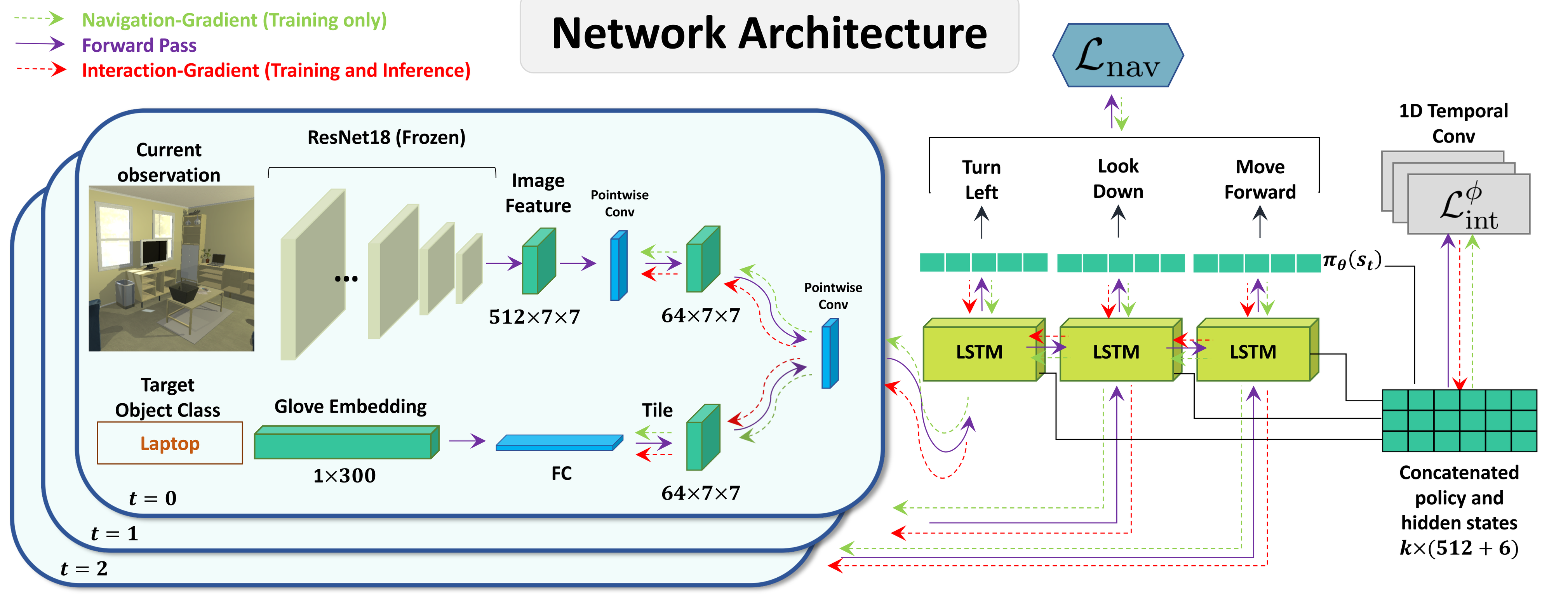
Objective: Learn the best *self-supervised interaction loss* and *model parameters* such that we succeed at navigation after adapting to the environment with a gradient step from the learned loss.

$\mathcal{L}_{\text{int}}^{\phi}$	Self-Supervised ***interaction loss*** computed via a neural network ***parameterized by*** ϕ.

$$\theta - \nabla_\theta \mathcal{L}_{\text{int}}^{\phi}\left(\theta, \mathcal{D}_\tau^{\text{int}}\right)$$

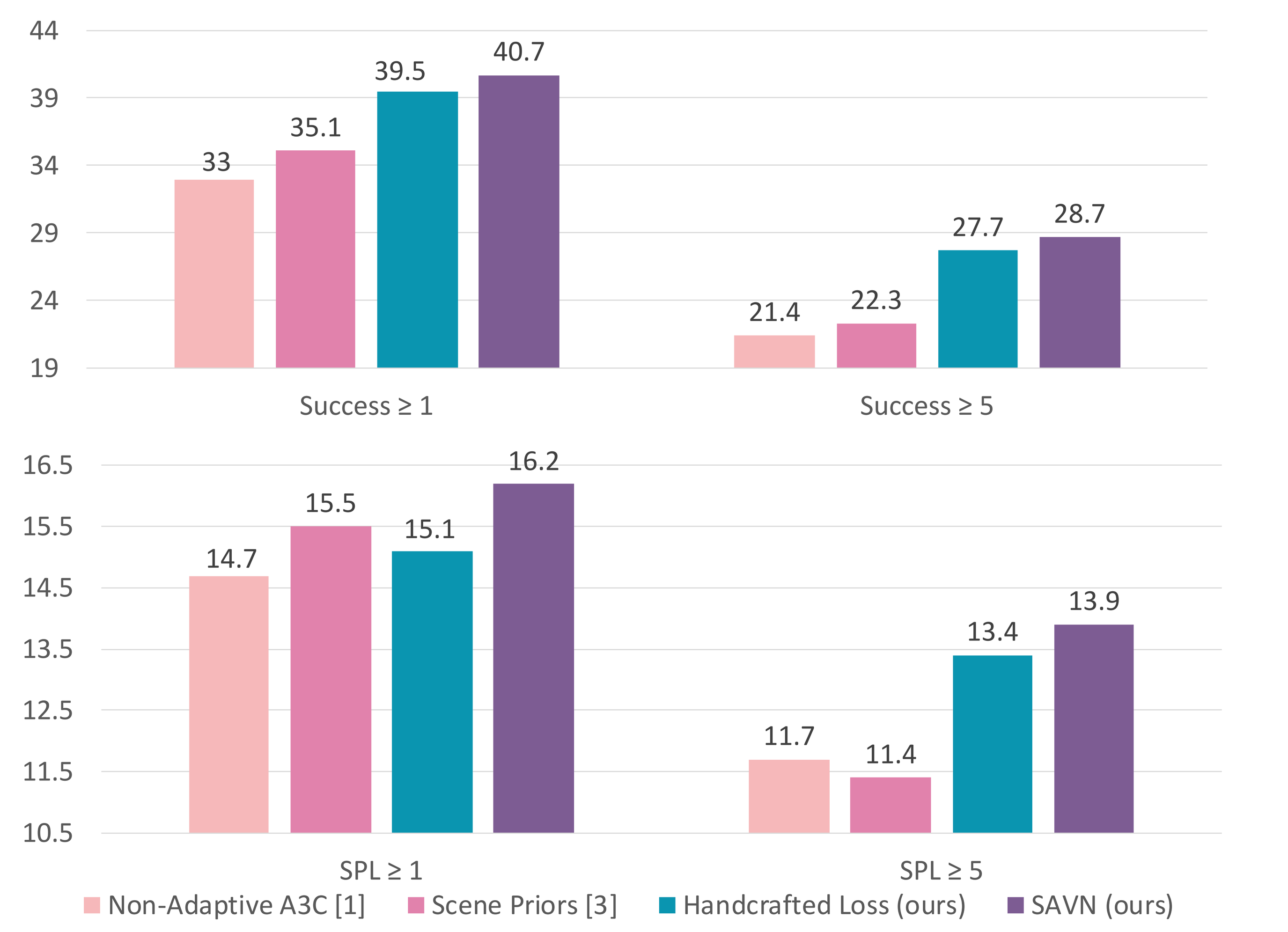
Adapted model parameters.

Network Architecture



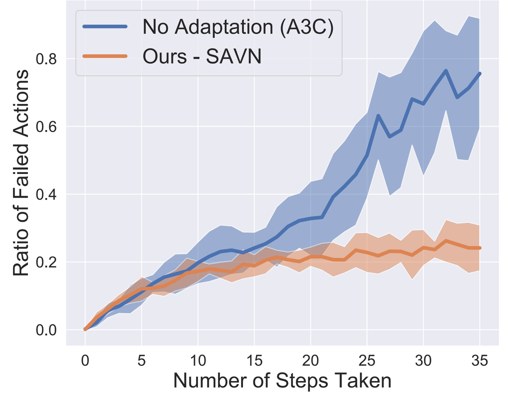
Experiments

≥ *k* = episodes where the optimal trajectory is at least *k* steps.



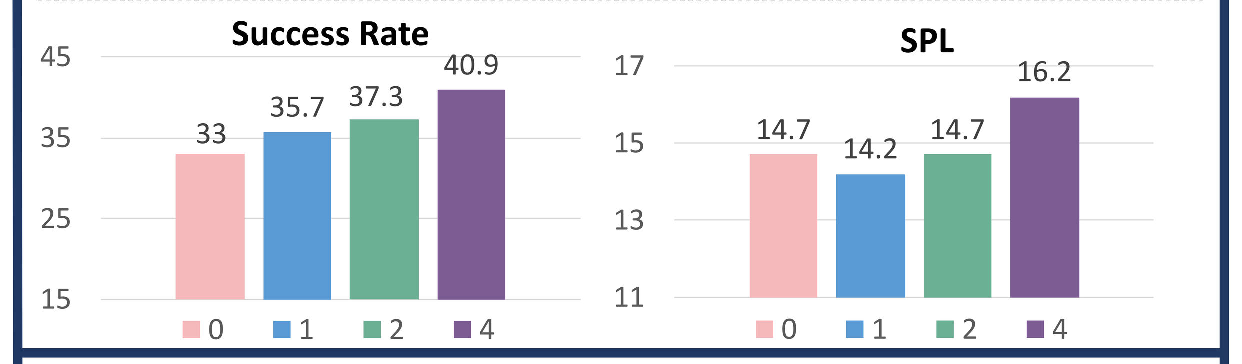
Experiments conducted using AI2-THOR [4], an interactive and near photo-realistic environment for training AI agents. We use:

- 120 scenes (80 for training, 20 for validation, 20 for testing).
- Scenes equally split between Kitchen, Living Room, Bedroom, and Bathroom.
- 18 target object classes.
- 1000 test episodes.



SAVN has fewer failed actions (e.g. bumping into walls) at test time.

Performance increases as more interaction-gradients are taken during training/testing.



References

[1] Volodymyr Mnih et al. Asynchronous methods for deep reinforcement learning. In *ICML*, 2016.
[2] Chelsea Finn et al. Model-agnostic meta-learning for fast adaptation of deep networks. In *ICML*, 2017.
[3] Wei Yang et al. Visual semantic navigation using scene priors. In *ICLR*, 2019.
[4] E. Kolve et al. AI2-THOR: An Interactive 3D Environment for Visual AI. *arXiv*, 2017.