10417/10617 Intermediate Deep Learning: Autonomous Navigation

Russ Salakhutdinov

Machine Learning Department
Carnegie Mellon University

Carnegie Mellon University

MACHINE LEARNING DEPARTMENT

Talk Outline

- Modular Visual Navigation using Active Neural Mapping
- Active Neural Localization: Towards Deep SLAM
- MineRL NeurIPS Competition

Navigation Tasks

Known goal location

- Require efficient navigation to the goal
- Tasks
  - Pointgoal [1, 2, 3]
  - Language Instructions describing path to goal [4]

[1] Anderson et al. *arXiv:1807.06757*, 2018.
[2] Mirowski et al. In *NeurIPS*, 2018.
[3] Savva et al. arXiv:1712.03931, 2017.
[4] Anderson et al. In *CVPR*, 2018.

Navigation Tasks

Known goal location

- Require efficient navigation to the goal
- Tasks
  - Pointgoal [1, 2, 3]
  - Language Instructions describing path to goal [4]

Unknown goal location

- Require exhaustive exploration
- Tasks
  - Exploration: Maximize explored area [5]
  - Object/Area Goal [3, 6, 7]
  - Semantic Goal Navigation [8]
  - Embodied Question Answering [9, 10]

[1] Anderson et al. *arXiv:1807.06757*, 2018.
[2] Mirowski et al. In *NeurIPS*, 2018.
[3] Savva et al. arXiv:1712.03931, 2017.
[4] Anderson et al. In *CVPR*, 2018.
[5] Chen et al. *ICLR*, 2019.
[6] Lample et al. In *AAAI*, 2017.
[7] Mirowski et al. *ICLR*, 2017.
[8] Chaplot et al. *AAAI*, 2018.
[9] Gordon et al. *CVPR*, 2018.
[10] Das et al. *CVPR*, 2018.

Desirable Characteristics of a Navigation model

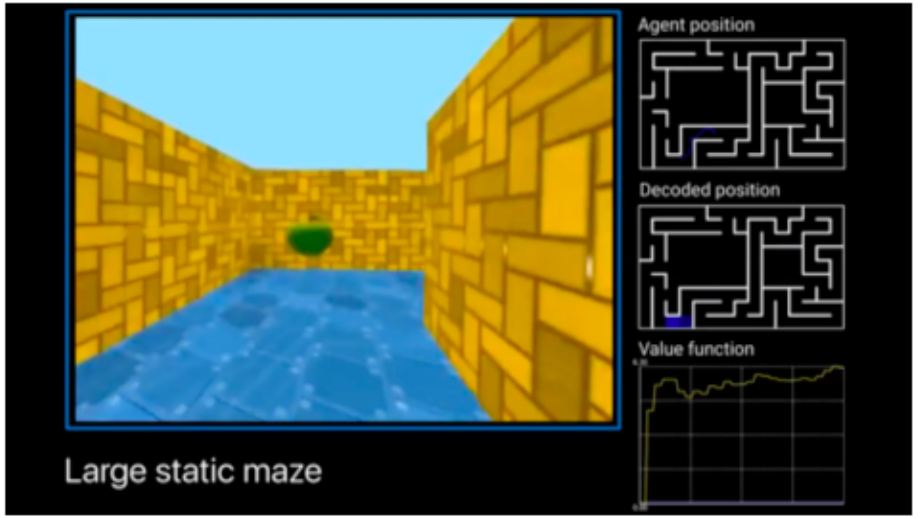
- Effective at both types of Navigation tasks:
  - Known goal location (Pointgoal) and
  - Unknown goal location (Exploration)
- Generalization: domains, task, goals
- Sample efficiency

Limitations of Classical SLAM

- Generalization
  - Robustness to environment conditions [Maddern et al. 2016]
  - Robustness to dynamic objects [Zou and Tan, 2012]
  - Failure cases of keypoint tracking [Cadena et al. 2016]
- Passiveness
  - Unable to decide the actions taken by the agent in order to map the environment or localize as accurately and efficiently as possible.

Deep RL?

[Lample & Chaplot, 2016]



[Mirowski et al. 2017]

Limitations of “end-to-end” Deep RL

- Ineffective at long-term planning
- Sample inefficiency
- Poor transferability

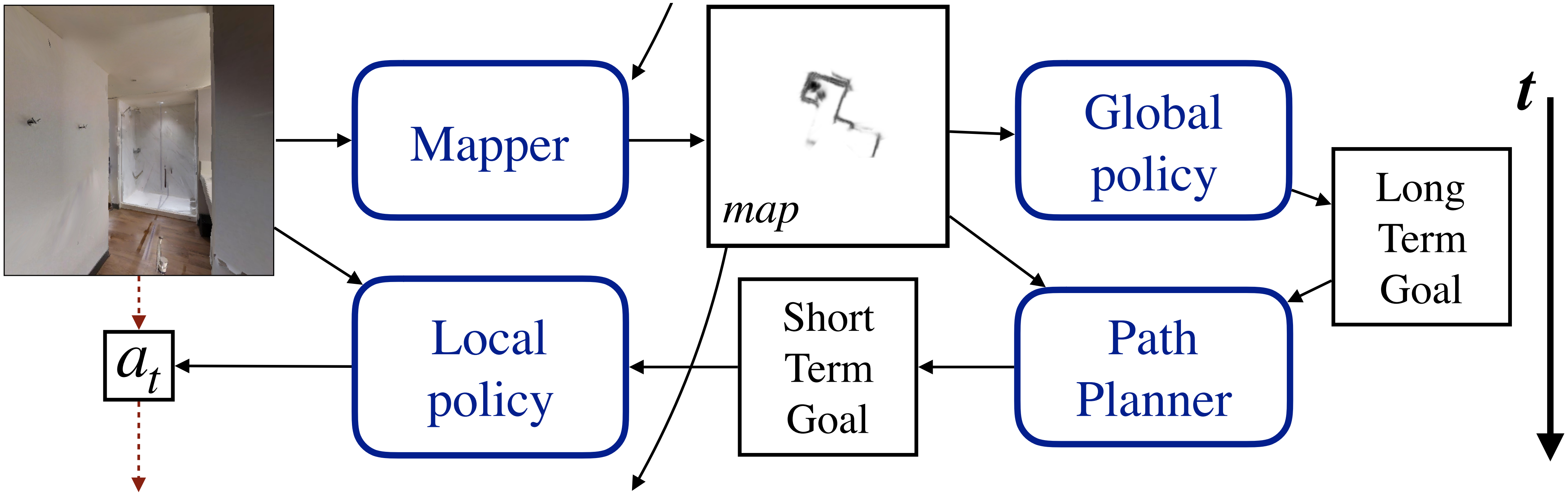
Active Neural Mapping

- Modular hierarchical navigation model that leverages the strengths of both learning-based and classical methods
- Efficient and exhaustive exploration, accurate long-term planning, domain and task generalization

- Won the **CVPR 2019 Habitat Challenge 2019** for PointGoal Navigation for both RGB or RGBD tracks.

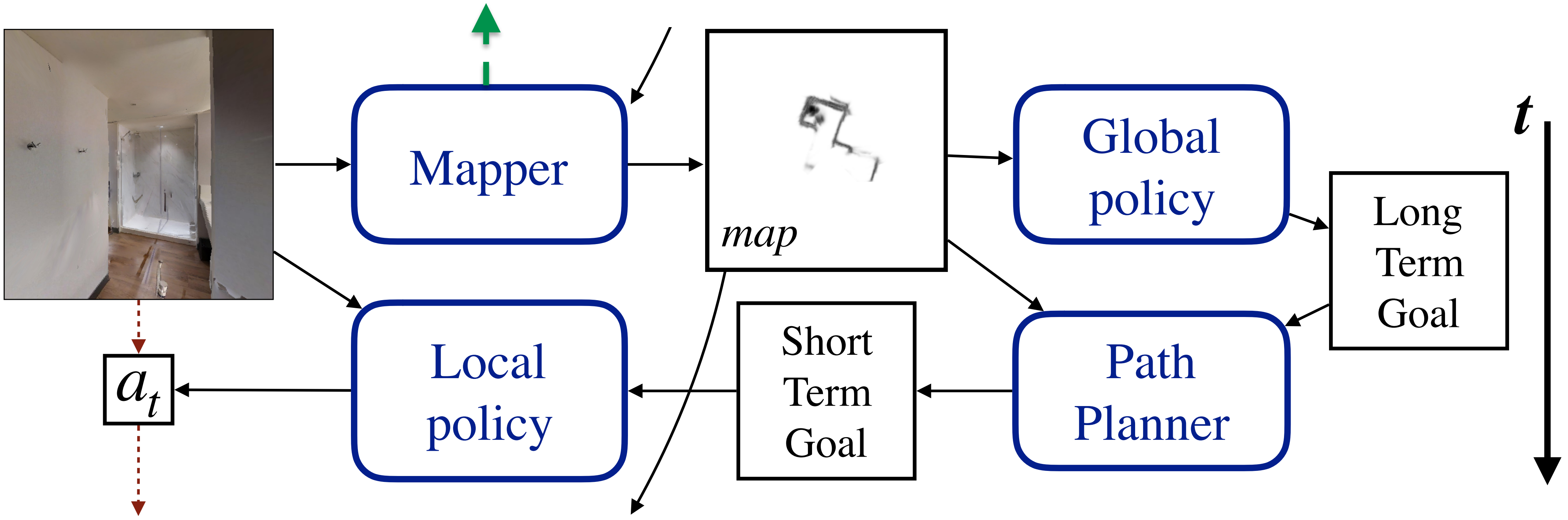
Devendra Chaplot, Saurabh Gupta, Abhinav Gupta, Ruslan Salakhutdinov,
Modular Visual Navigation using Active Neural Mapping, 2019

Active Neural Mapping



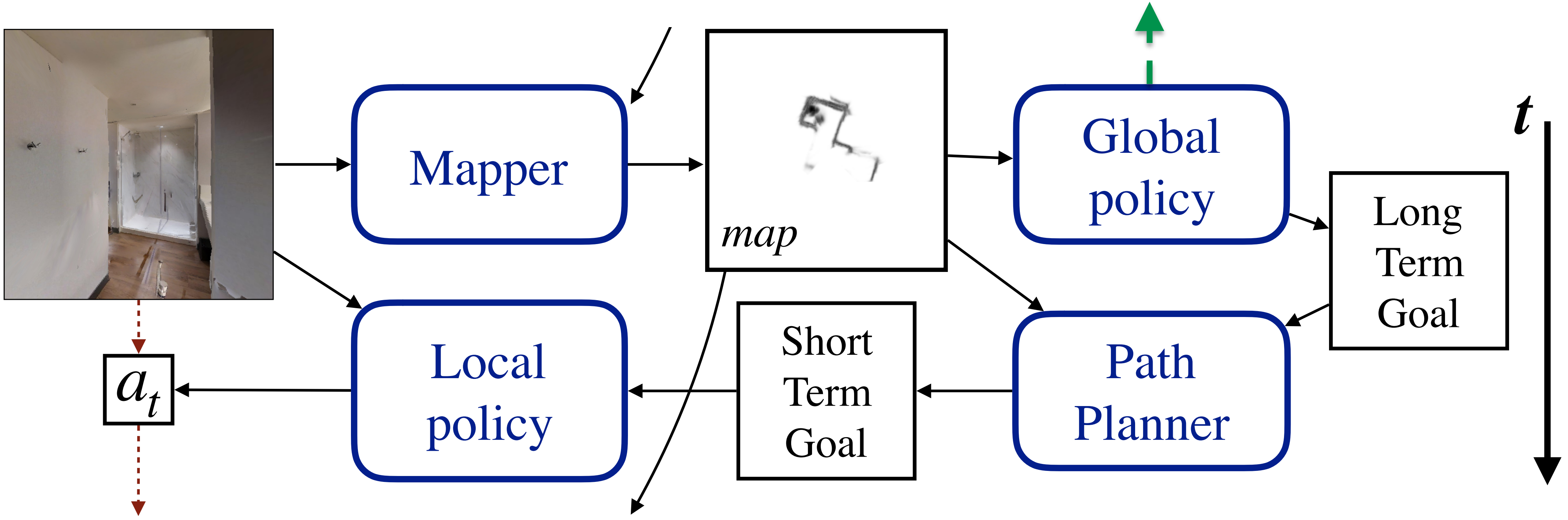
Active Neural Mapping

Updates the map based on the current observation

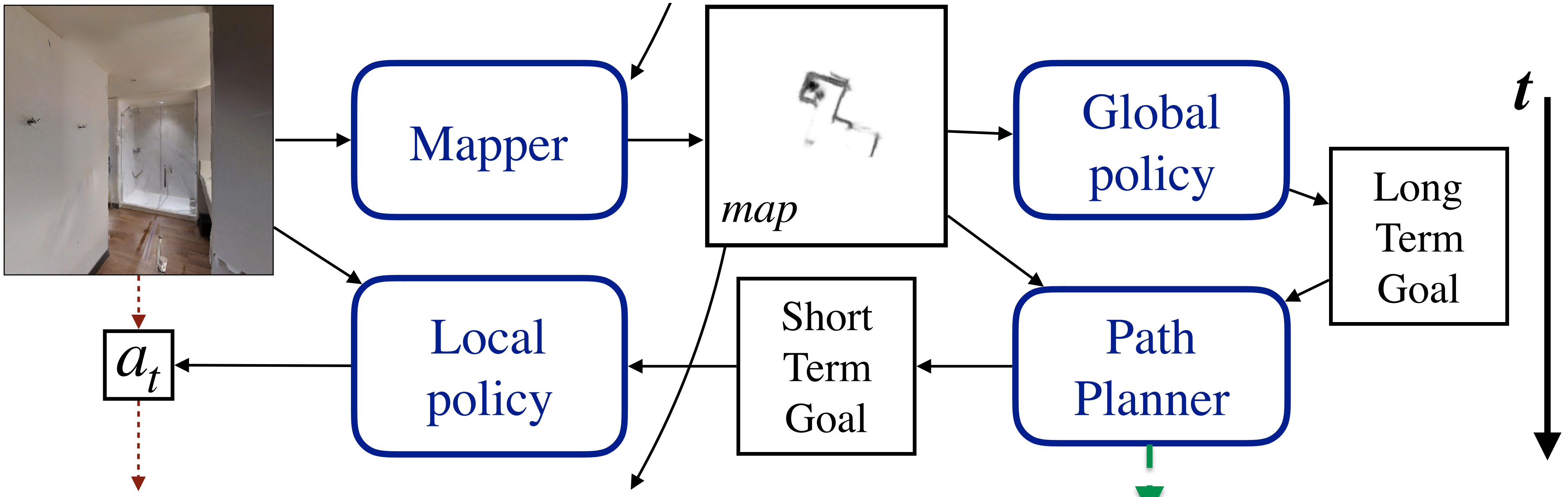


Active Neural Mapping

Samples a long-term goal based on the current map

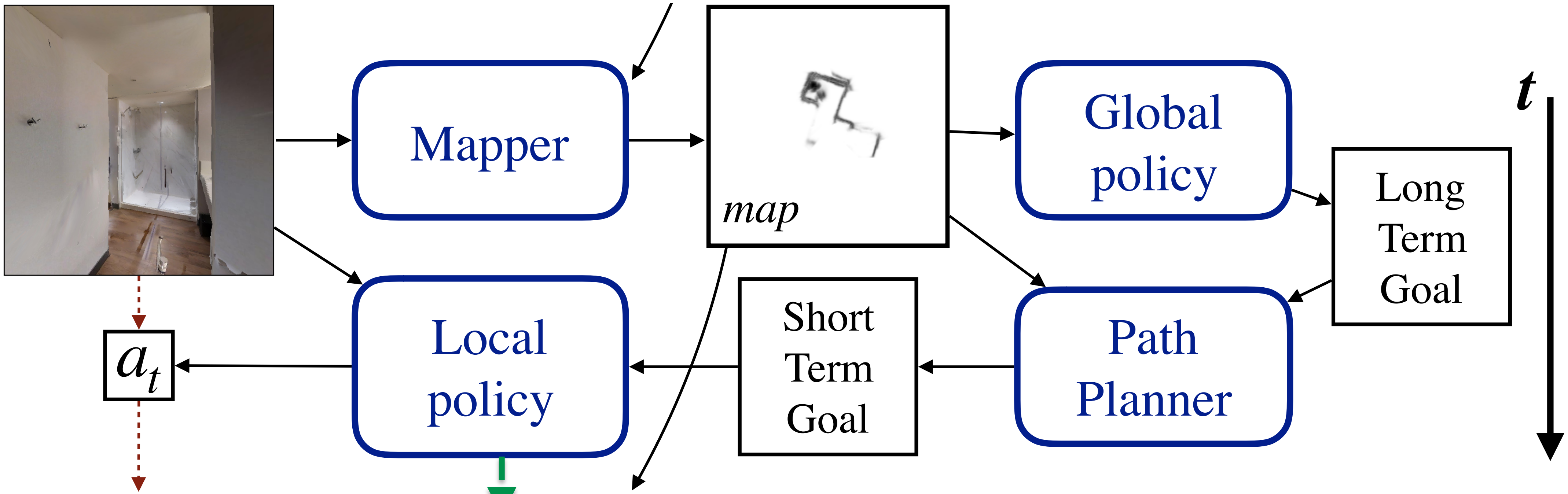


Active Neural Mapping



Plans a path to the long-term goal and computes a short-term goal on the planned path

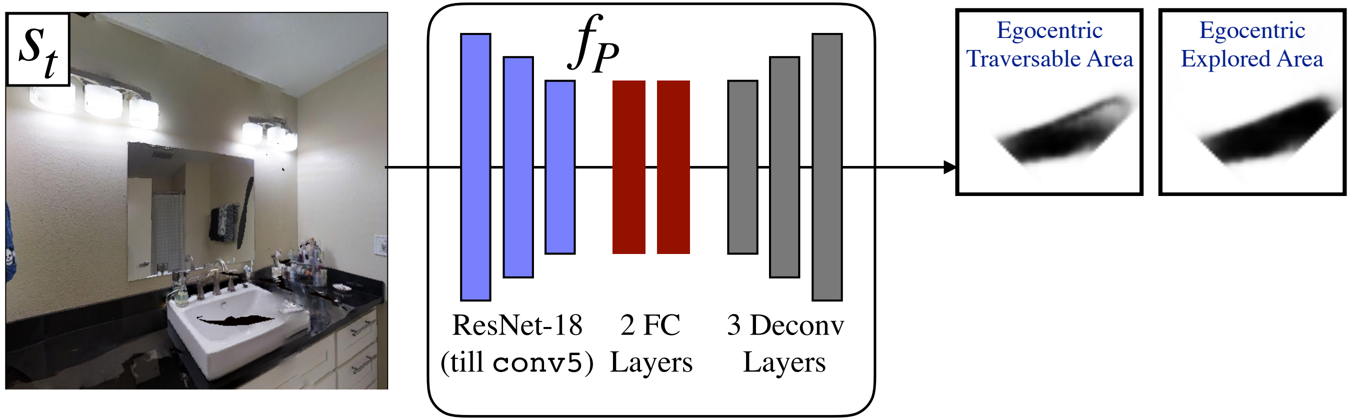
Active Neural Mapping



Samples a navigation action to reach the short-term goal

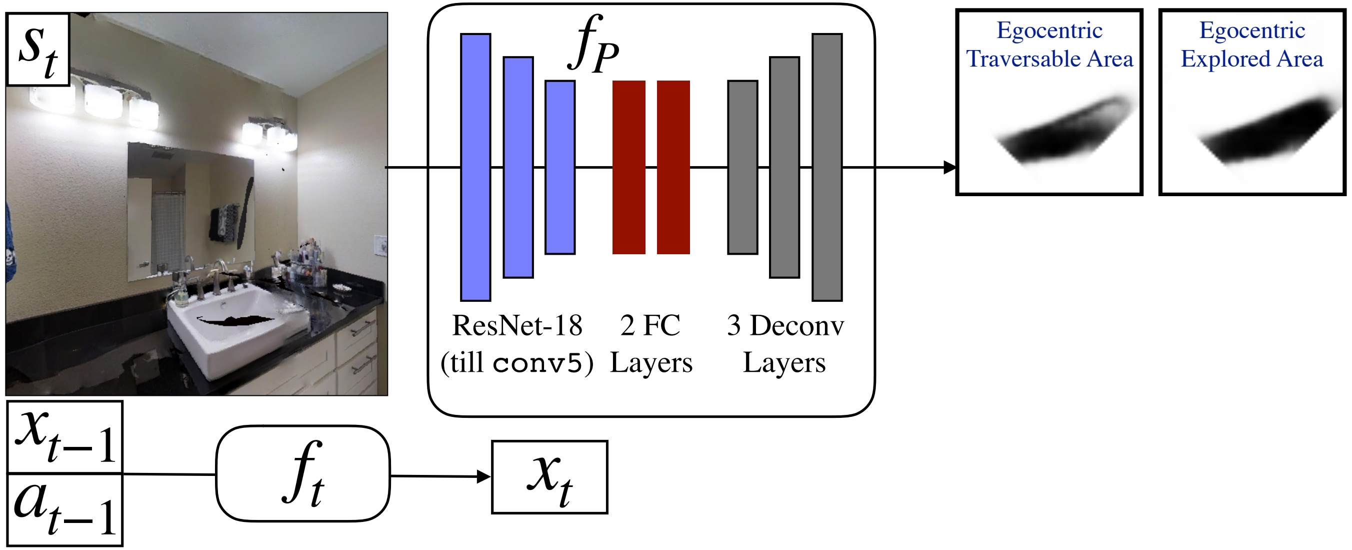
Mapper

Mapper



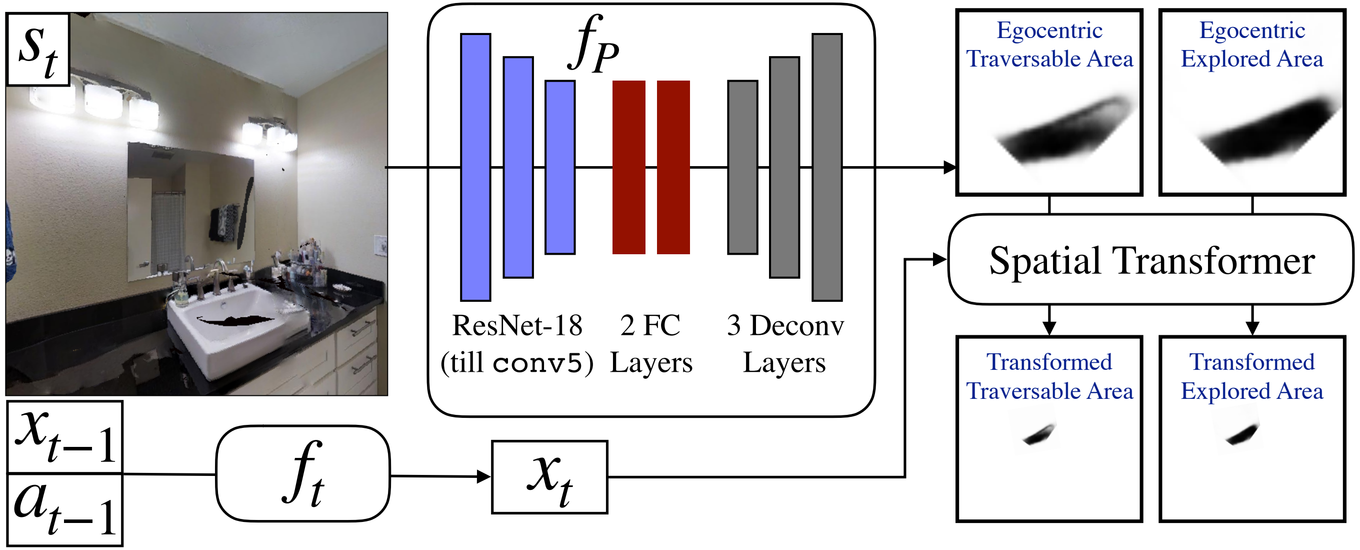
Projection Unit f_P: takes an RGB frame and outputs an egocentric top-down 2D spatial map, predicting obstacles and explored area in the current observation

Mapper



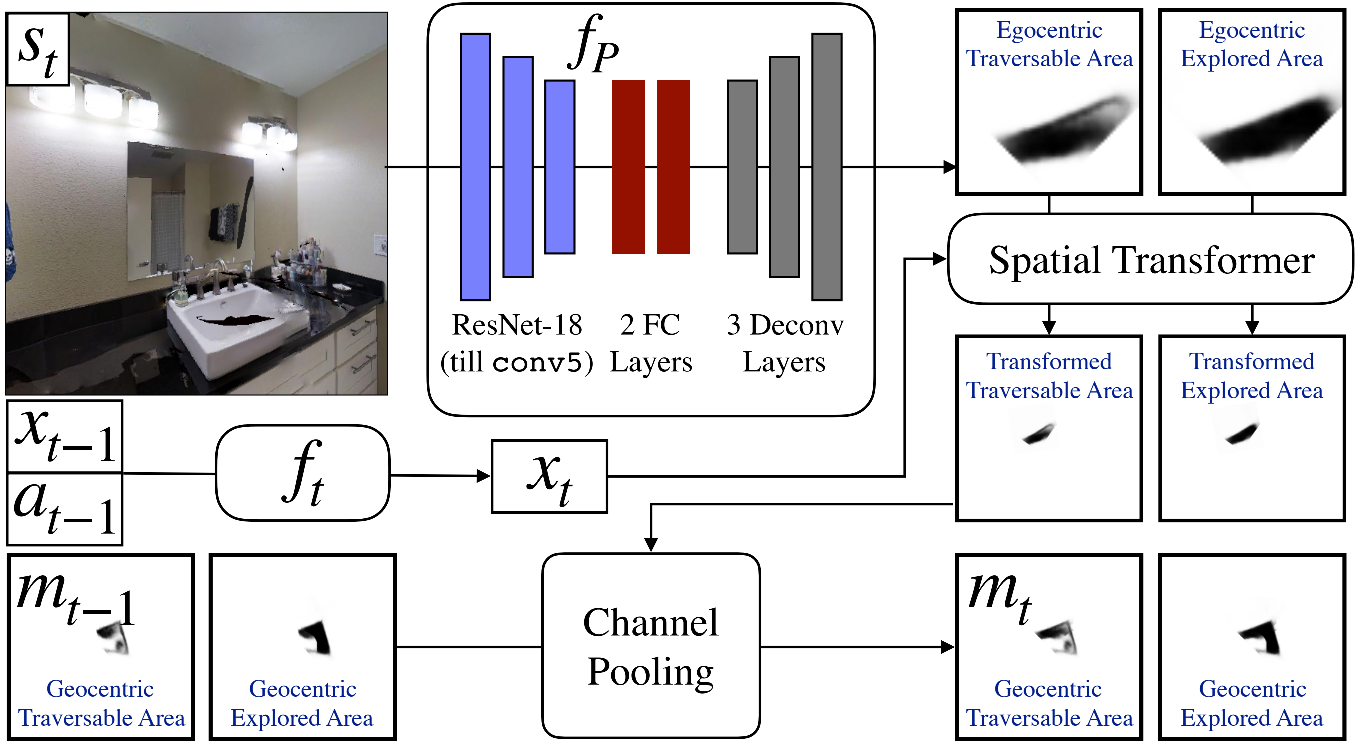
The pose of the agent is computed based on the previous action using transition function f_t

Mapper



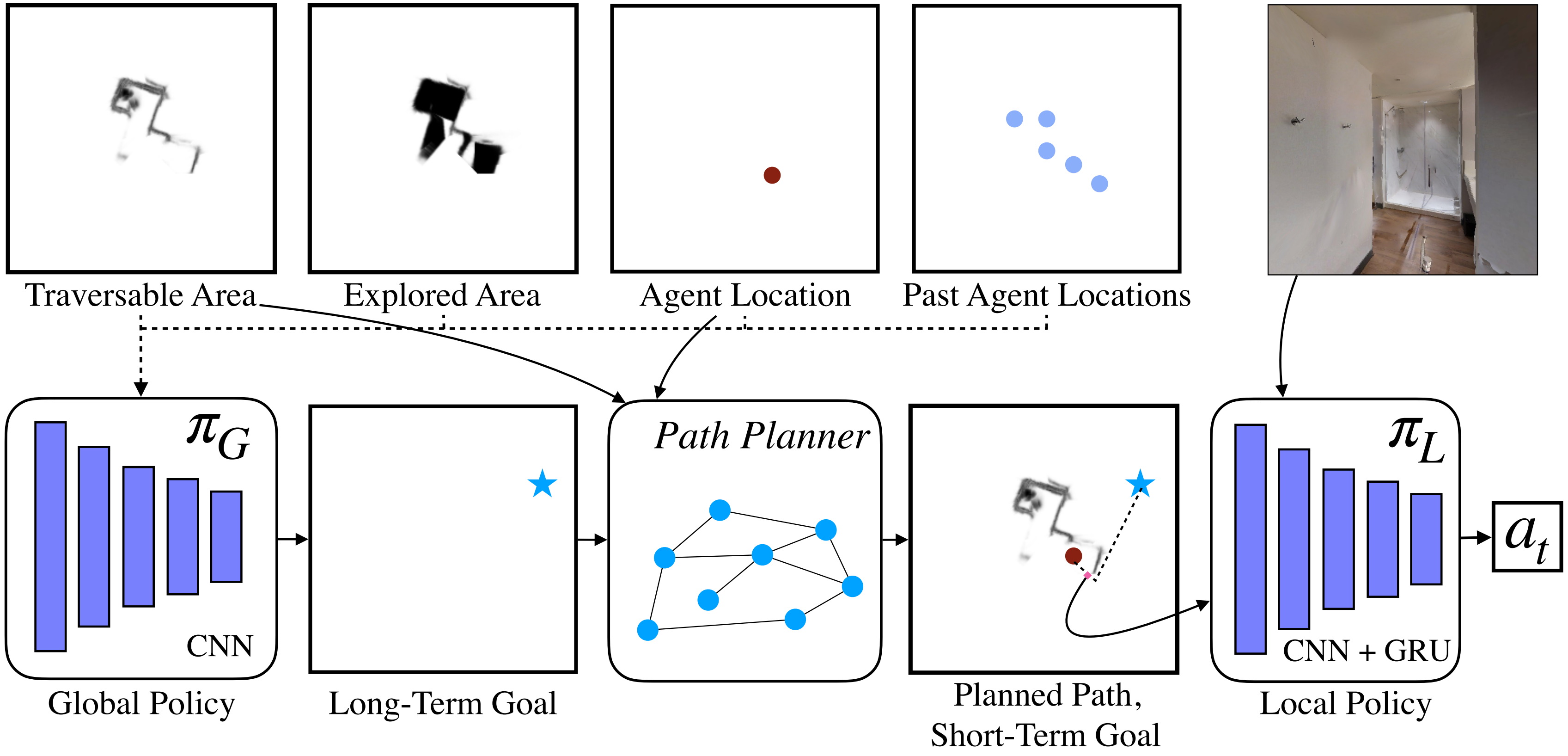
Egocentric spatial map is transformed into geocentric frame using the current pose of the agent X_t using Spatial Transformation

Mapper

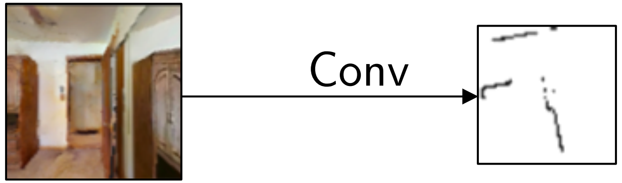


Geocentric map prediction of the current frame is aggregated with the previous spatial map M_{t-1} using Channel-wise Pooling

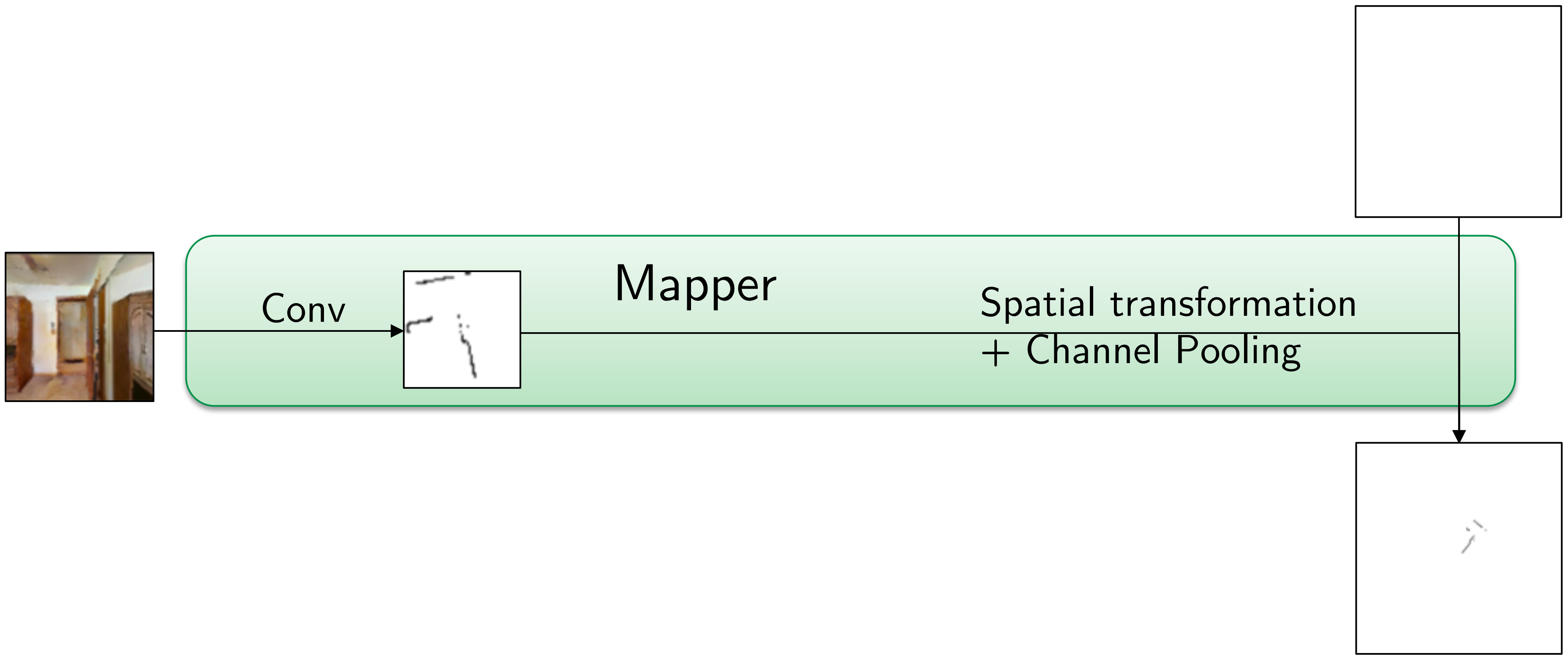
Global and Local Policies



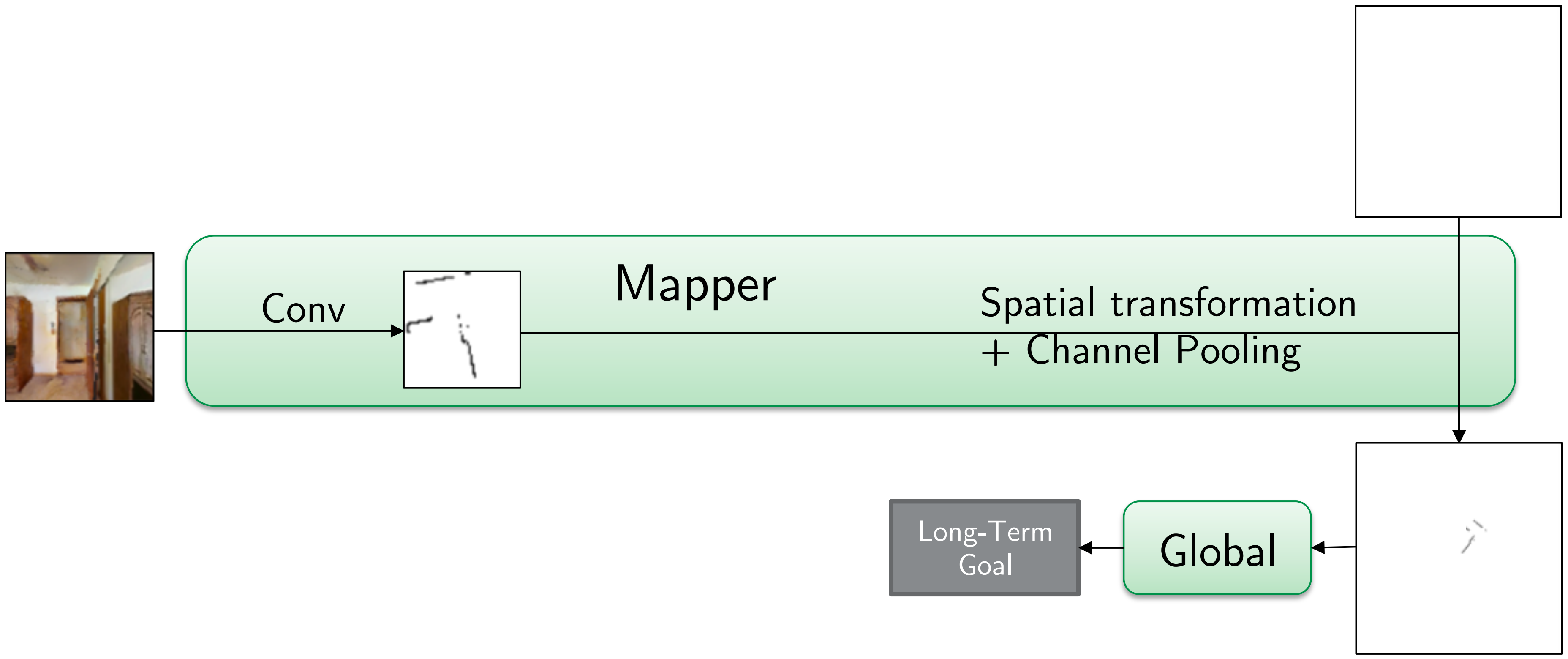
Active Neural Mapping



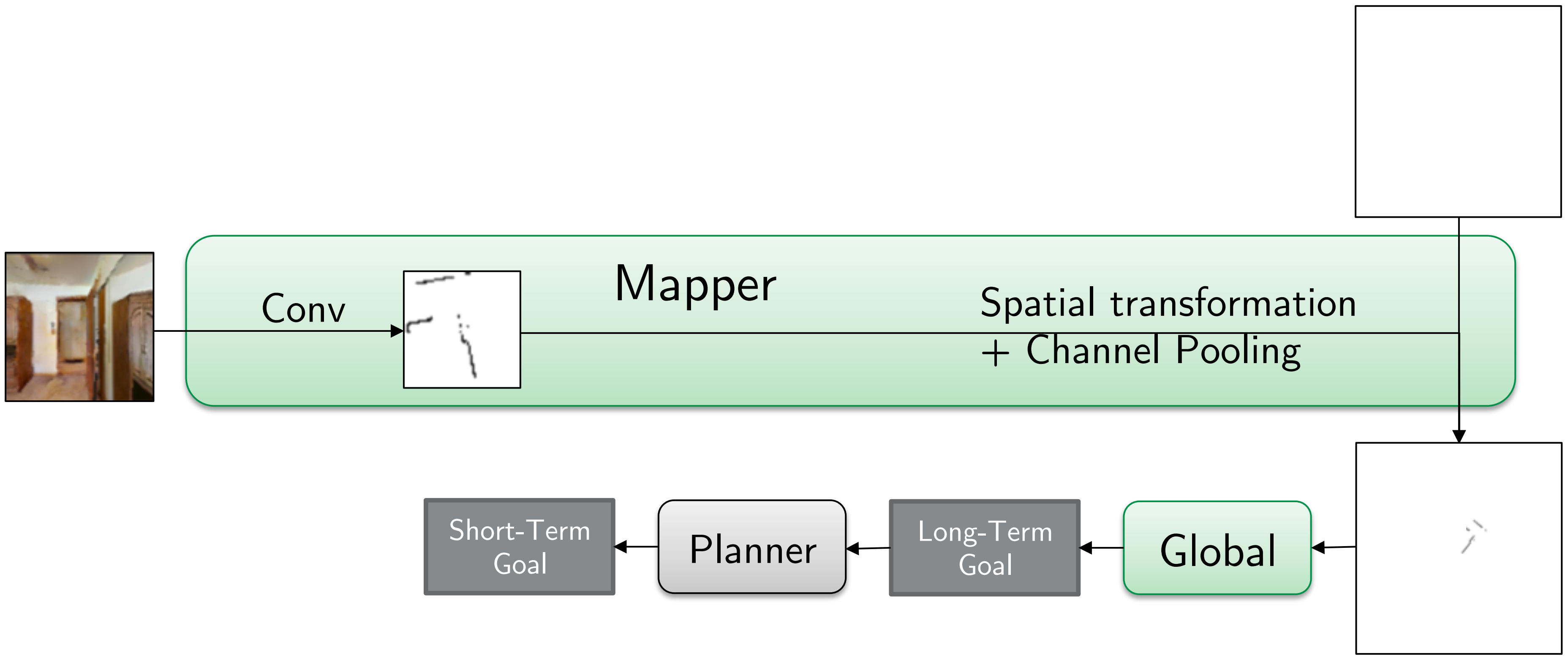
Active Neural Mapping



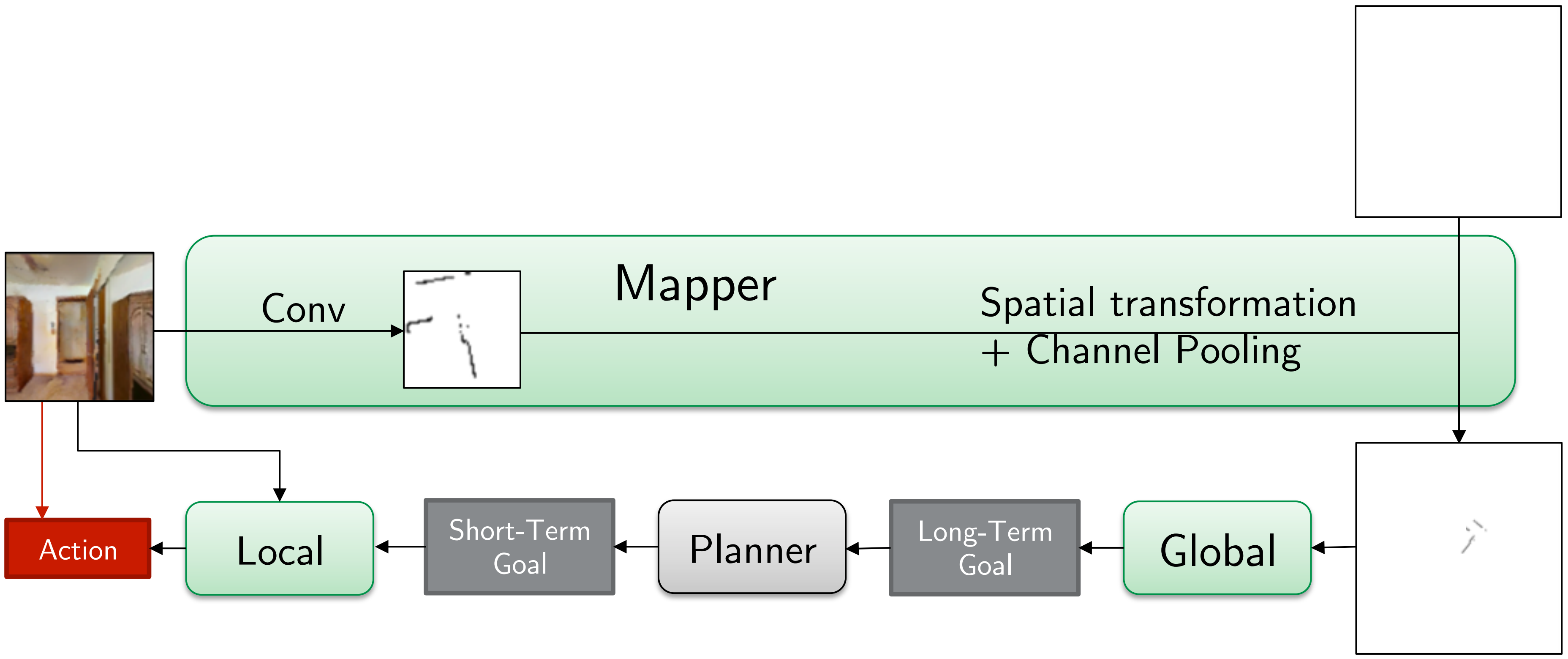
Active Neural Mapping



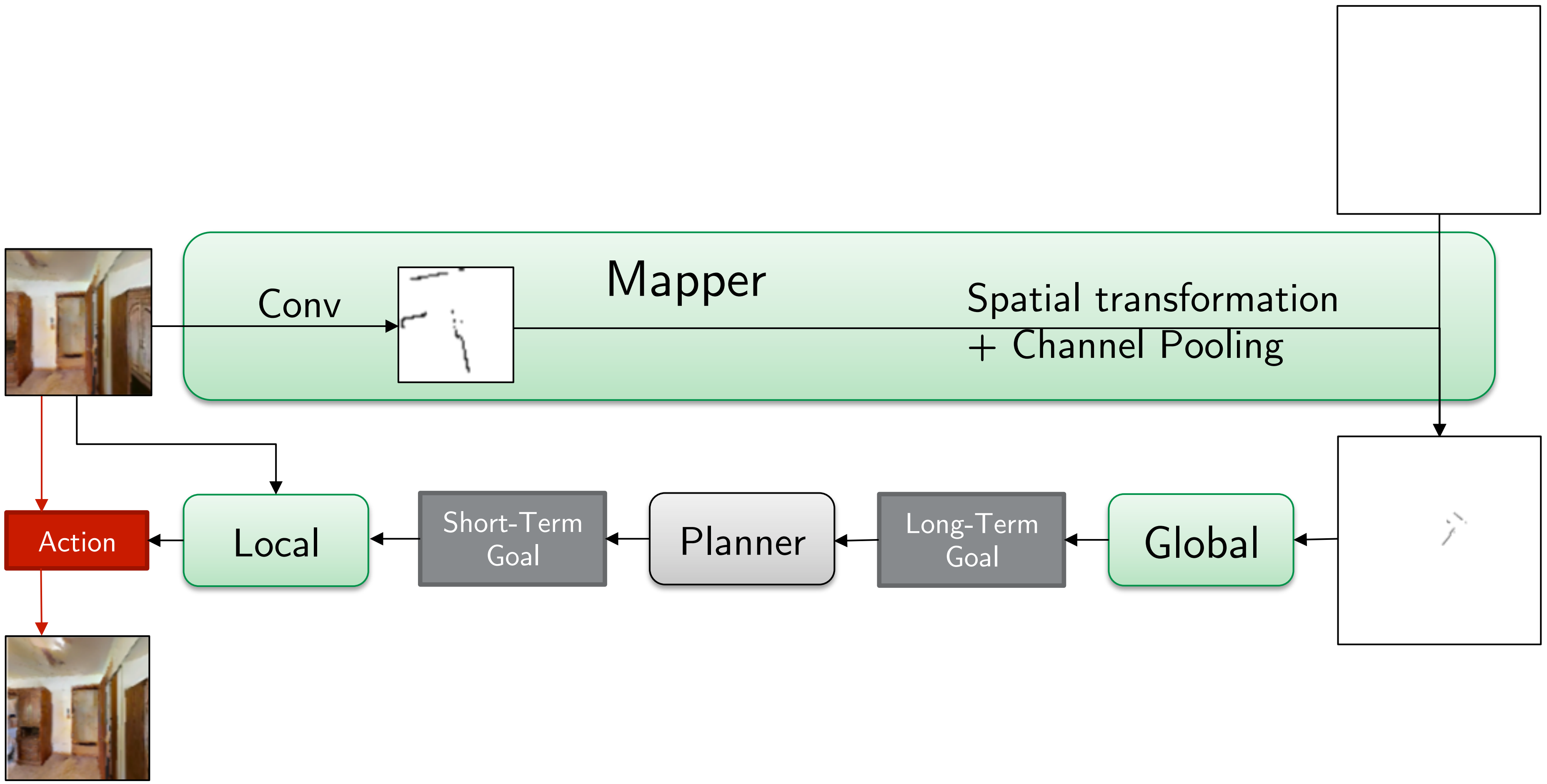
Active Neural Mapping



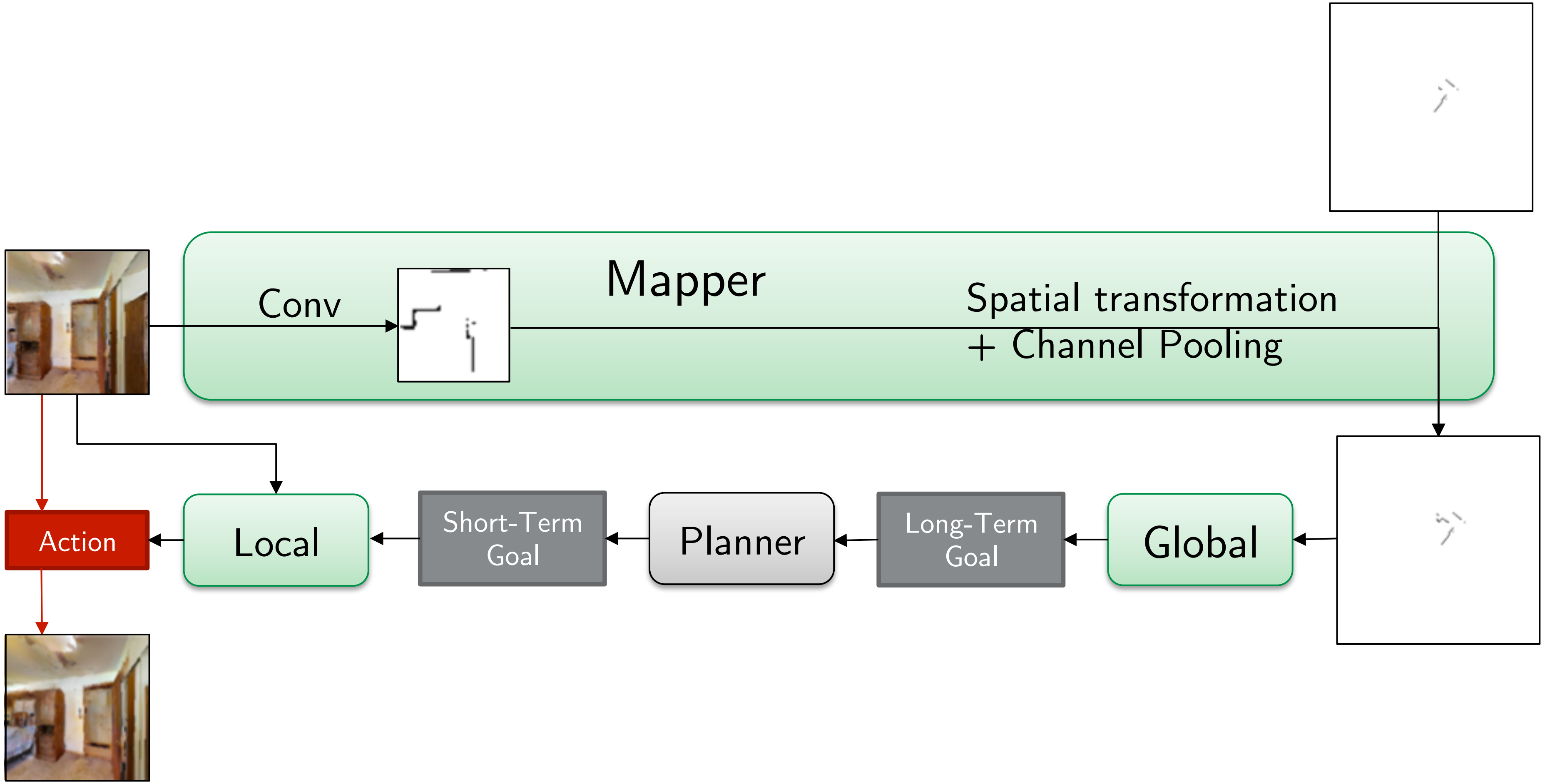
Active Neural Mapping



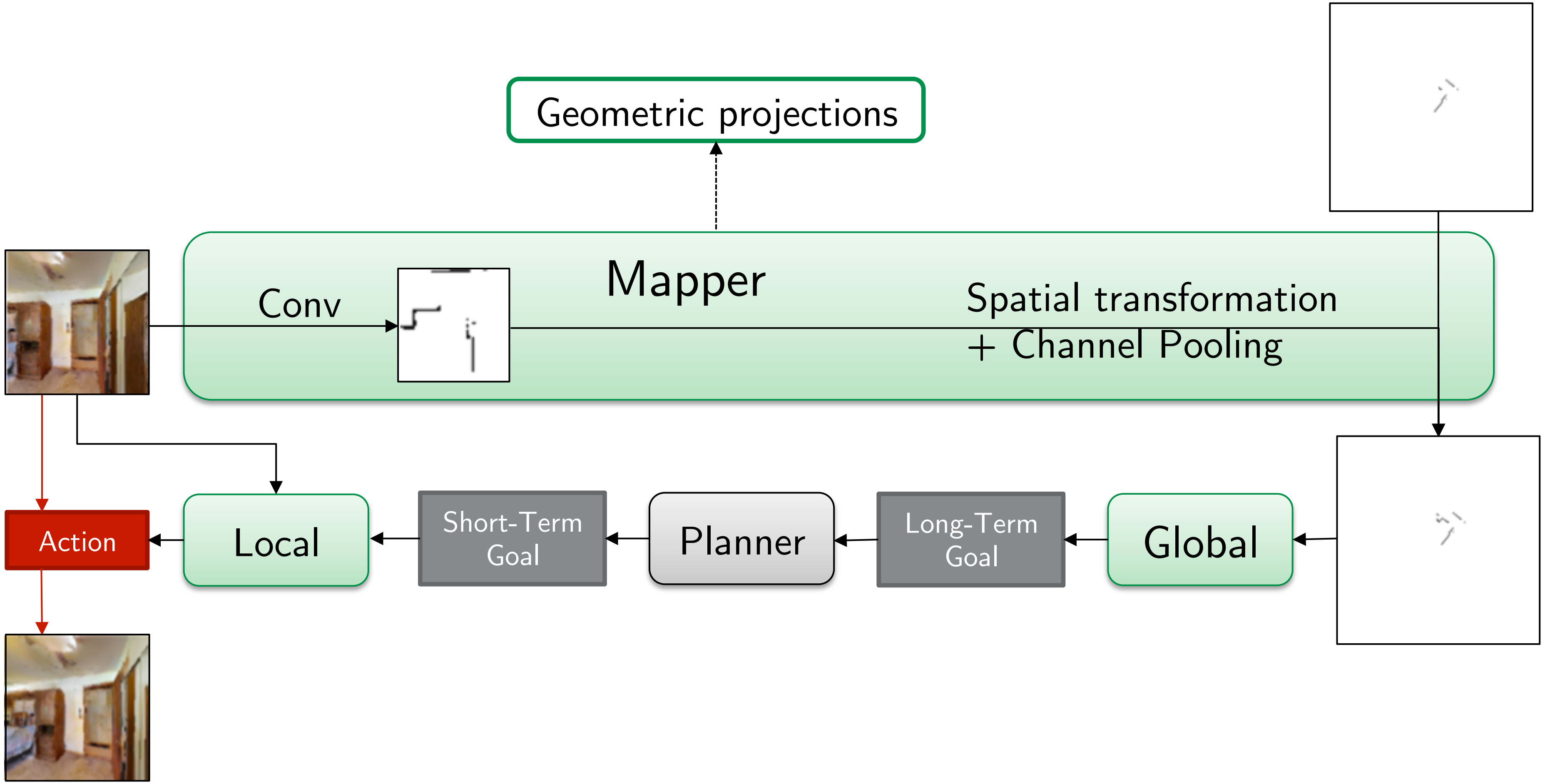
Active Neural Mapping



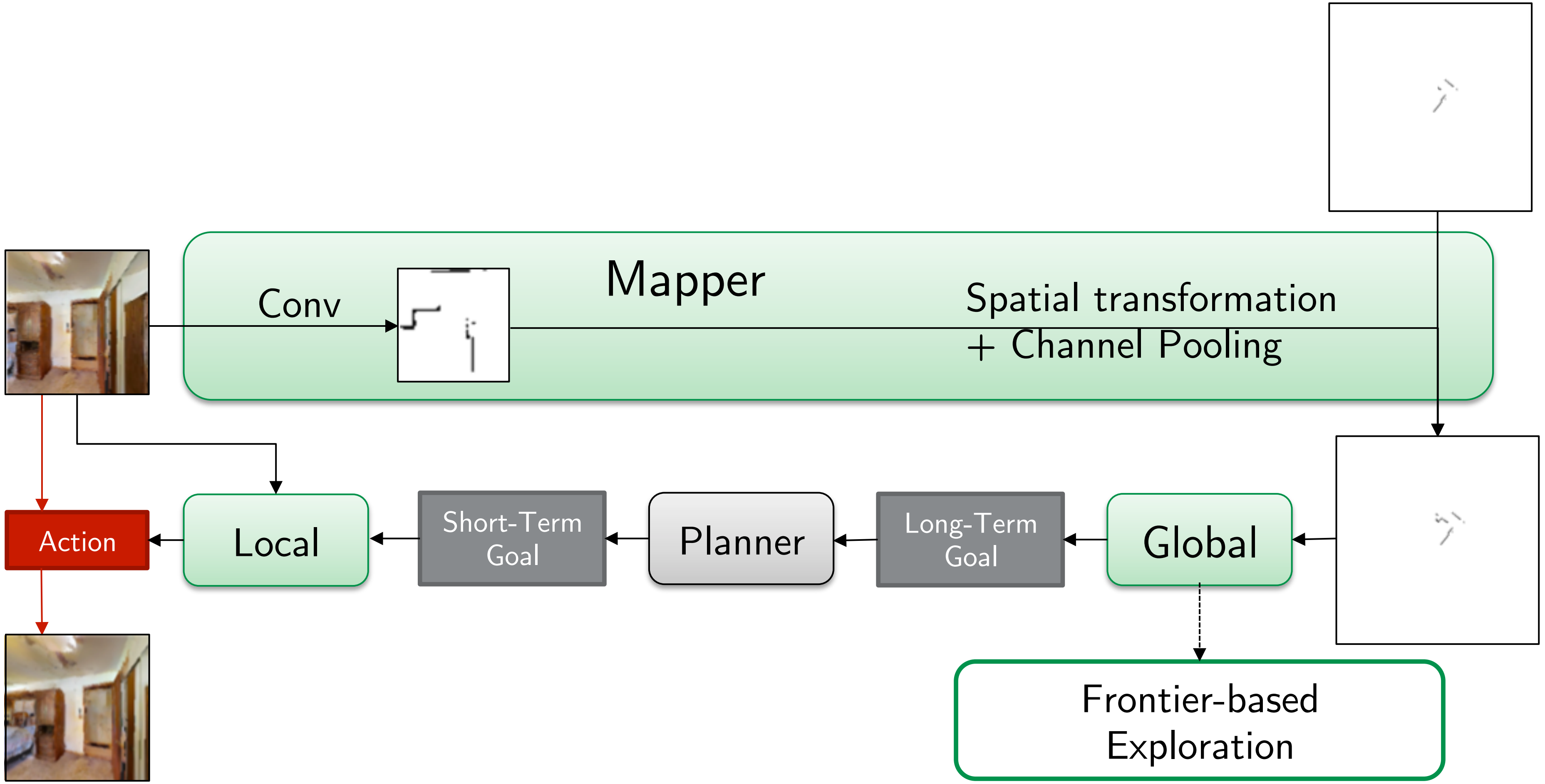
Active Neural Mapping



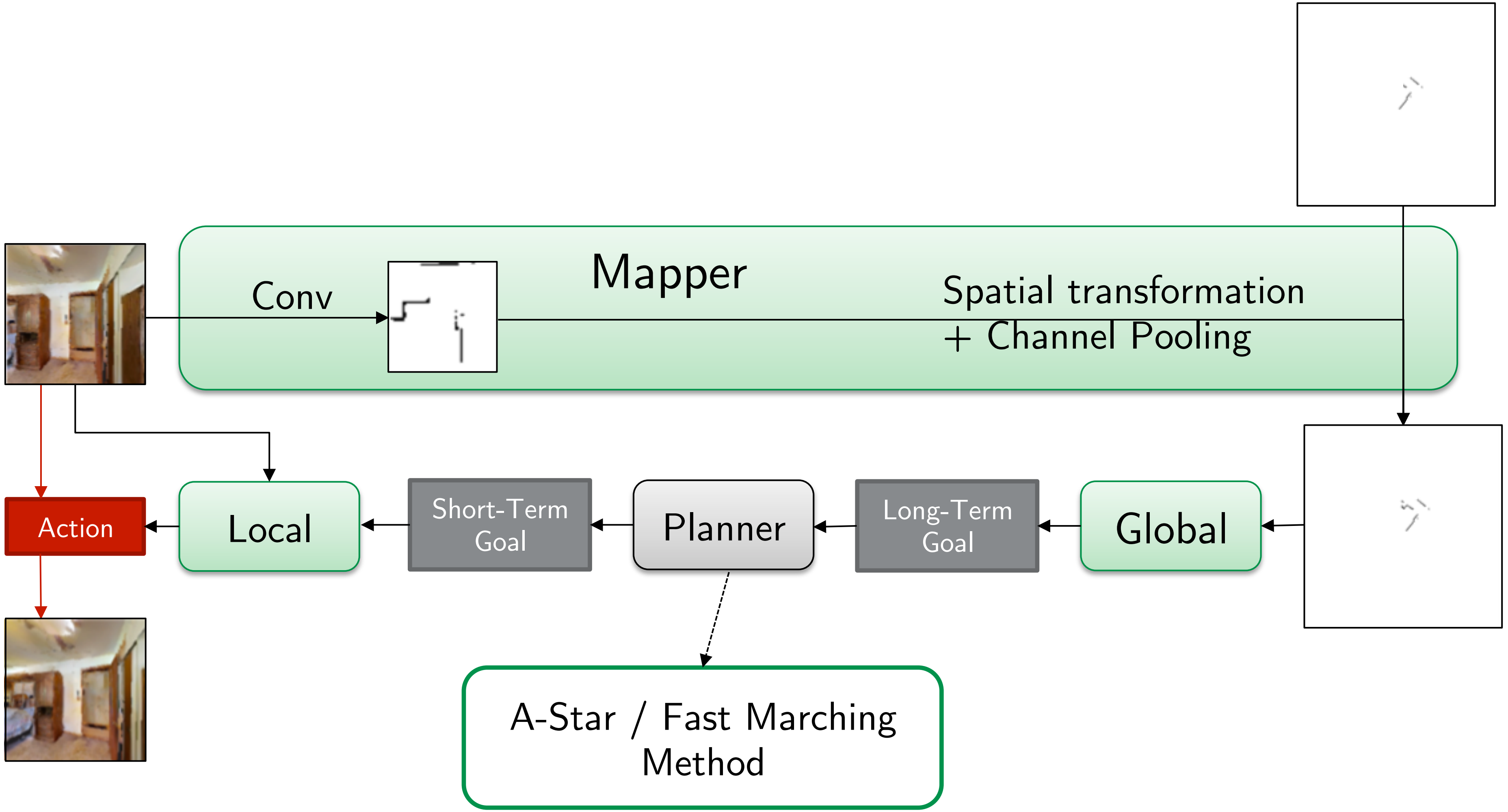
Active Neural Mapping



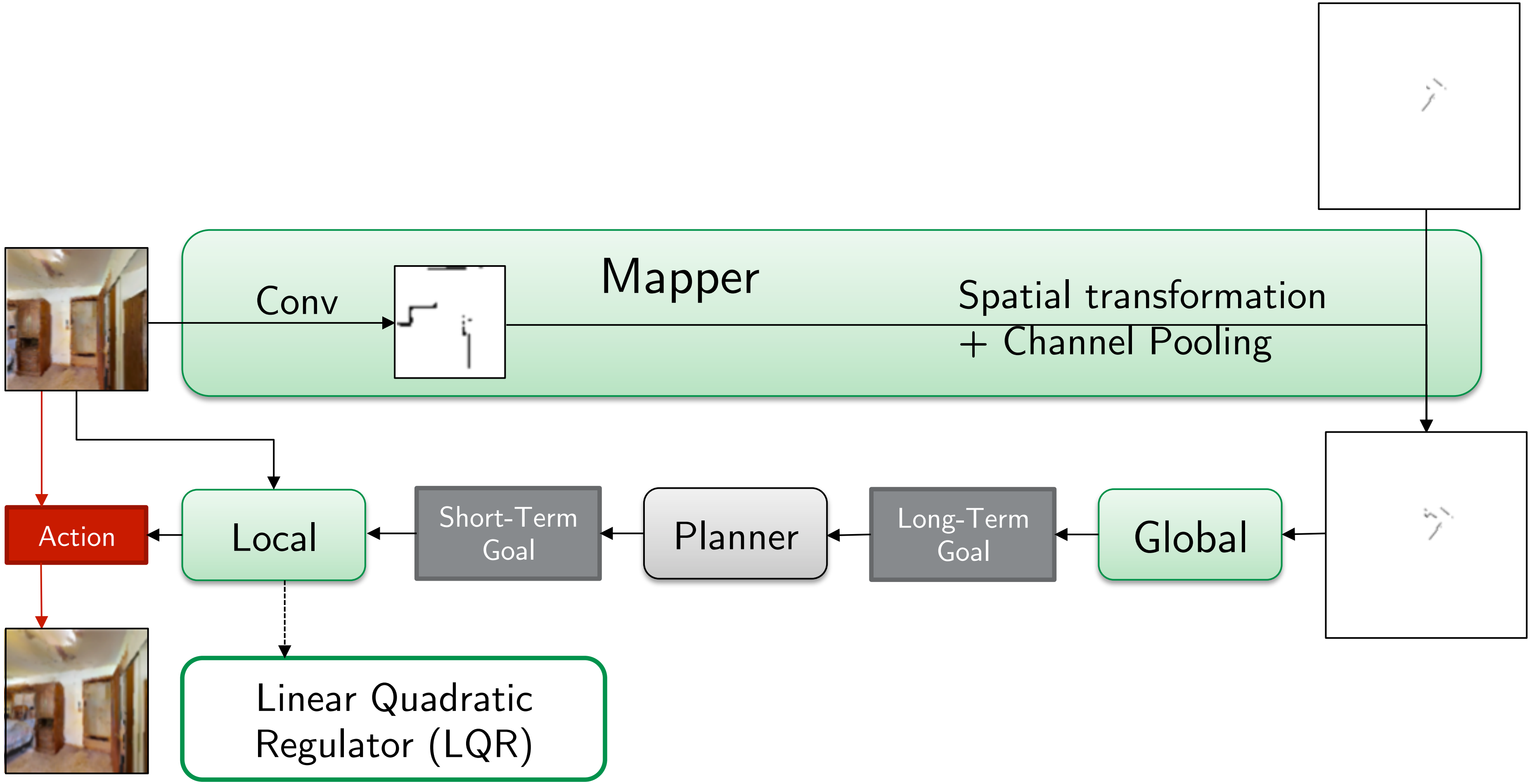
Active Neural Mapping



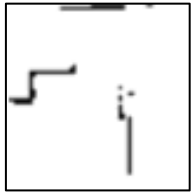
Active Neural Mapping



Active Neural Mapping



Training



- Mapper – Supervised on Projection
- Global Policy (PPO) – Explored area as reward
- Local Policy (PPO) – Distance reduced to short-term goal as reward
- Planner – Fast Marching Method

Exploration Task

- Use Habitat simulator with Gibson and Matterport3D datasets
- Objective: Maximize the explored area
- Metric: Explored area or coverage (m^2)
- A cell is explored if it is either
  - Known to be traversable or
  - Known to be an obstacle
- All methods trained for 10 million frames
- Fixed episode length of 500 steps (about 3 mins)

Exploration results

	Gibson Val
Model	Coverage (m^2)
Random	11.52
RL + 3LConv + GRU [1]	21.60
RL + Res18 + GRU	24.48
RL + Res18 + GRU + AuxDepth [2]	28.80
RL + Res18 + GRU + ProjDepth [3]	30.24
Active Neural Mapping (ANM)	**43.20**

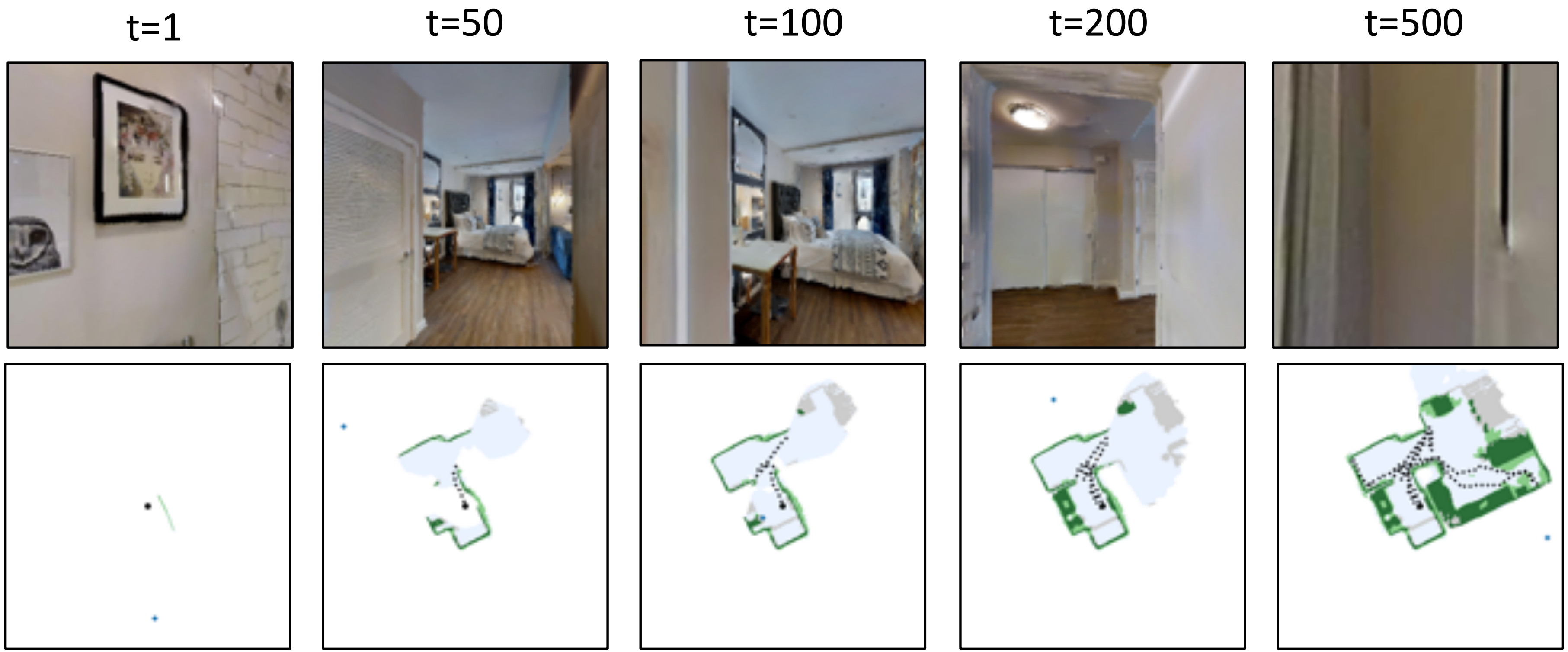
*Adapted from [1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17, [3] Chen el al. ICLR-19

Exploration results: Domain Generalization

	Gibson Val	MP3D Test
Model	Coverage (m^2)	Coverage (m^2)
Random	11.52	25.92
RL + 3LConv + GRU [1]	21.60	33.55
RL + Res18 + GRU	24.48	33.12
RL + Res18 + GRU + AuxDepth [2]	28.80	45.36
RL + Res18 + GRU + ProjDepth [3]	30.24	41.04
Active Neural Mapping (ANM)	**43.20**	**63.07**

*Adapted from [1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17, [3] Chen el al. ICLR-19

Examples



Pointgoal Evaluation

- Objective: Navigate to goal coordinates
- Global Policy: Always gives pointgoal as a long-term goal
- All methods trained for 10 million frames
- Metric: Success weighted by inverse Path Length (SPL)

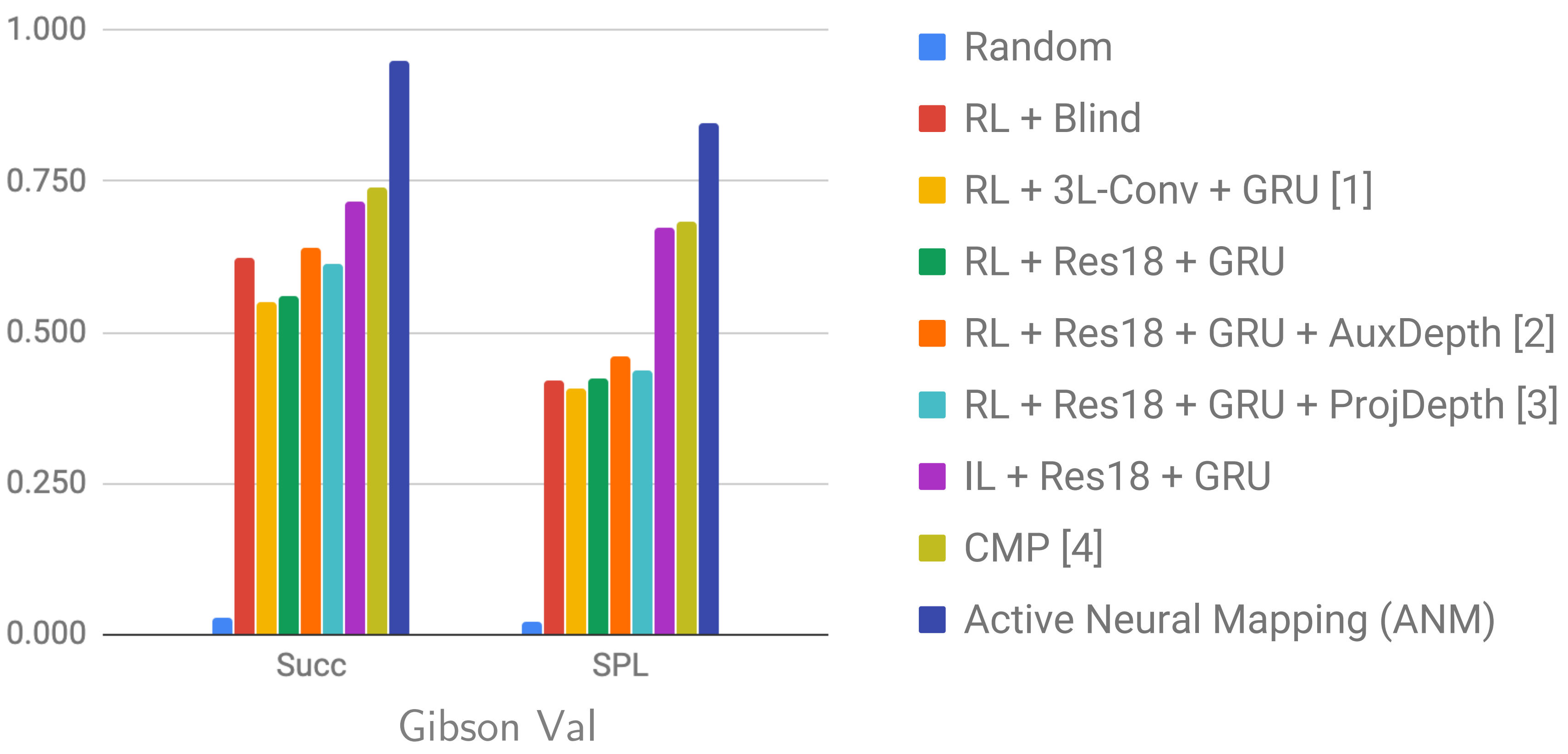
$$\frac{1}{N}\sum_{i=1}^{N} S_i \frac{\ell_i}{\max(p_i, \ell_i)}.$$

Pointgoal Results

Test Setting ->	Gibson Val	
Method	Succ	SPL
Random	0.027	0.021
RL + Blind	0.625	0.421
RL + 3L-Conv + GRU [1]	0.550	0.406
RL + Res18 + GRU	0.561	0.422
RL + Res18 + GRU + AuxDepth [2]	0.640	0.461
RL + Res18 + GRU + ProjDepth [3]	0.614	0.436
IL + Res18 + GRU	0.716	0.673
IL + CMP [4]	0.738	0.683
Active Neural Mapping (ANM)	**0.951**	**0.848**

*Adapted from [1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17, [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Pointgoal Results



*Adapted from
[1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17,
[3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Harder Goals

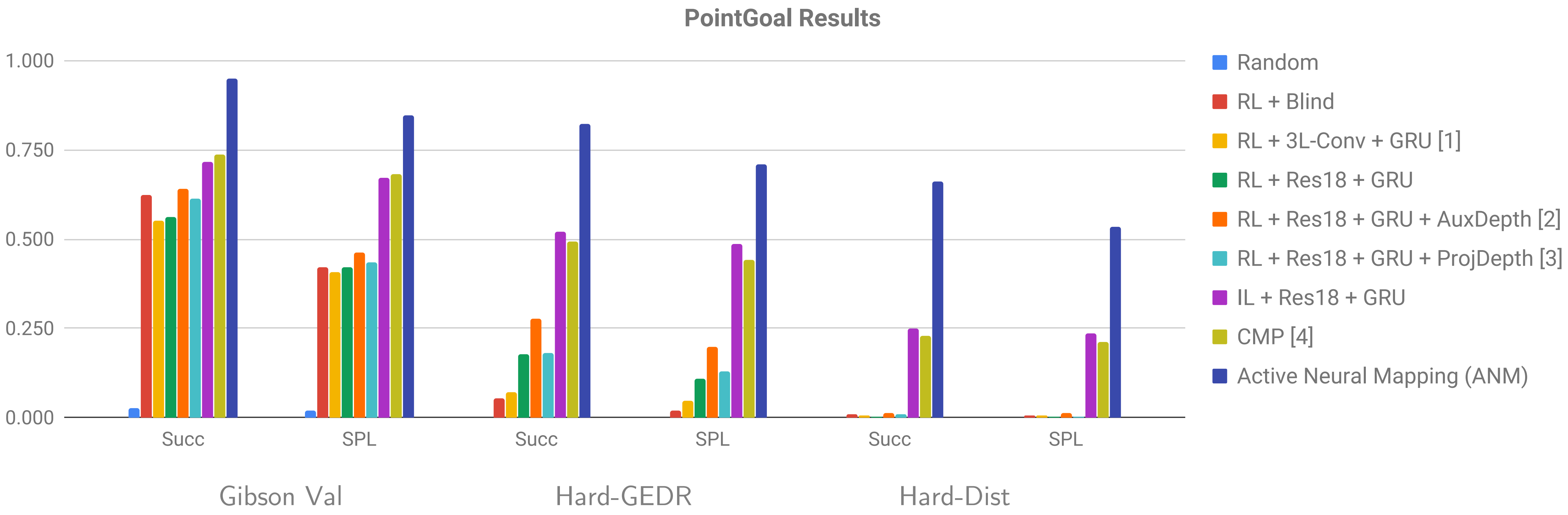
- Higher Geodesic to Euclidean Distance Ratio (Hard-GEDR)
- Higher Geodesic Distance (Hard-Dist)

Pointgoal Results

			Goal Generalization			
Test Setting ->	Gibson Val		Hard-GEDR		Hard-Dist	
Method	Succ	SPL	Succ	SPL	Succ	SPL
Random	0.027	0.021	0.000	0.000	0.000	0.000
RL + Blind	0.625	0.421	0.052	0.020	0.008	0.006
RL + 3L-Conv + GRU [1]	0.550	0.406	0.072	0.046	0.006	0.006
RL + Res18 + GRU	0.561	0.422	0.176	0.109	0.004	0.003
RL + Res18 + GRU + AuxDepth [2]	0.640	0.461	0.277	0.197	0.013	0.011
RL + Res18 + GRU + ProjDepth [3]	0.614	0.436	0.180	0.129	0.008	0.004
IL + Res18 + GRU	0.716	0.673	0.521	0.486	0.248	0.234
IL + CMP [4]	0.738	0.683	0.492	0.443	0.228	0.212
Active Neural Mapping (ANM)	**0.951**	**0.848**	**0.824**	**0.710**	**0.662**	**0.534**

*Adapted from [1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17, [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Pointgoal Results



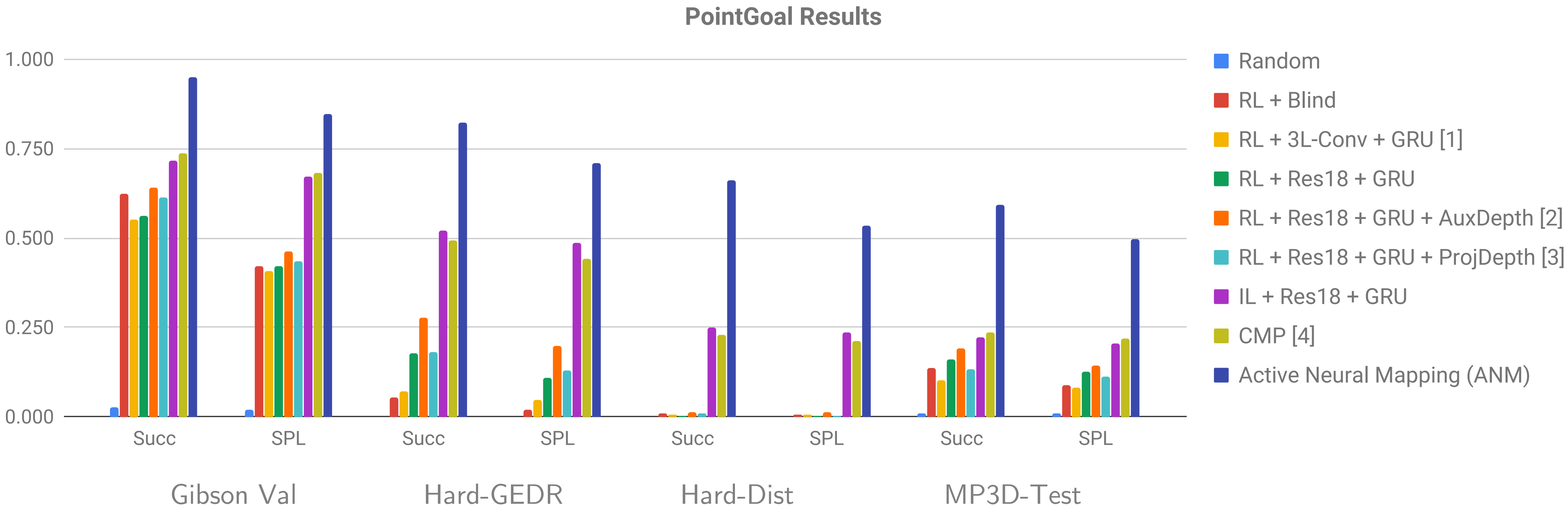
*Adapted from [1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17, [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Pointgoal Results

	Goal Generalization						Domain Generalization	
Test Setting ->	Gibson Val		Hard-GEDR		Hard-Dist		MP3D Test	
Method	Succ	SPL	Succ	SPL	Succ	SPL	Succ	SPL
Random	0.027	0.021	0.000	0.000	0.000	0.000	0.010	0.010
RL + Blind	0.625	0.421	0.052	0.020	0.008	0.006	0.136	0.087
RL + 3L-Conv + GRU [1]	0.550	0.406	0.072	0.046	0.006	0.006	0.102	0.080
RL + Res18 + GRU	0.561	0.422	0.176	0.109	0.004	0.003	0.160	0.125
RL + Res18 + GRU + AuxDepth [2]	0.640	0.461	0.277	0.197	0.013	0.011	0.189	0.143
RL + Res18 + GRU + ProjDepth [3]	0.614	0.436	0.180	0.129	0.008	0.004	0.134	0.111
IL + Res18 + GRU	0.716	0.673	0.521	0.486	0.248	0.234	0.221	0.205
IL + CMP [4]	0.738	0.683	0.492	0.443	0.228	0.212	0.237	0.219
Active Neural Mapping (ANM)	**0.951**	**0.848**	**0.824**	**0.710**	**0.662**	**0.534**	**0.593**	**0.496**

*Adapted from [1] Lample & Chaplot. AAAI-17, [2] Mirowski et al. ICLR-17, [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Pointgoal Results



Pointgoal: Task Transfer

- Pointgoal task: Navigation to goal coordinates
- Global Policy: always gives pointgoal as long-term goal
- Task Transfer: Use Local Policy and Mapper trained for exploration

Pointgoal Results

				Goal Generalization				Domain Generalization	
	Test Setting ->	Gibson Val		Hard-GEDR		Hard-Dist		MP3D Test	
Train Task	Method	Succ	SPL	Succ	SPL	Succ	SPL	Succ	SPL
Pointgoal	Random	0.027	0.021	0.000	0.000	0.000	0.000	0.010	0.010
	RL + Blind	0.625	0.421	0.052	0.020	0.008	0.006	0.136	0.087
	RL + 3L-Conv + GRU [1]	0.550	0.406	0.072	0.046	0.006	0.006	0.102	0.080
	RL + Res18 + GRU	0.561	0.422	0.176	0.109	0.004	0.003	0.160	0.125
	RL + Res18 + GRU + AuxDepth [2]	0.640	0.461	0.277	0.197	0.013	0.011	0.189	0.143
	RL + Res18 + GRU + ProjDepth [3]	0.614	0.436	0.180	0.129	0.008	0.004	0.134	0.111
	IL + Res18 + GRU	0.716	0.673	0.521	0.486	0.248	0.234	0.221	0.205
	IL + CMP [4]	0.738	0.683	0.492	0.443	0.228	0.212	0.237	0.219
	Active Neural Mapping (ANM)	**0.951**	**0.848**	**0.824**	**0.710**	0.662	**0.534**	**0.593**	**0.496**
Exploration	ANM - Task Transfer	0.950	0.846	0.821	0.703	**0.665**	0.532	0.588	0.490

Sample Efficiency

	Training Frames	Succ	SPL
RL + Res18 + GRU + ProjDepth [3]	10 million	0.640	0.461
RL + Res18 + GRU + ProjDepth [3]	75 million	0.678	0.486
IL + Cognitive Mapping & Planning [4]	10 million	0.738	0.683
Active Neural Mapping	1 million	0.789	0.703
Active Neural Mapping	10 million	0.951	0.848

*Adapted from [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Sample Efficiency

	Training Frames	Succ	SPL
RL + Res18 + GRU + ProjDepth [3]	10 million	0.640	0.461
RL + Res18 + GRU + ProjDepth [3]	75 million	0.678	0.486
IL + Cognitive Mapping & Planning [4]	10 million	0.738	0.683
Active Neural Mapping	1 million	0.789	0.703
Active Neural Mapping	10 million	0.951	0.848

> 75x speedup as compared to best RL baseline

*Adapted from [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

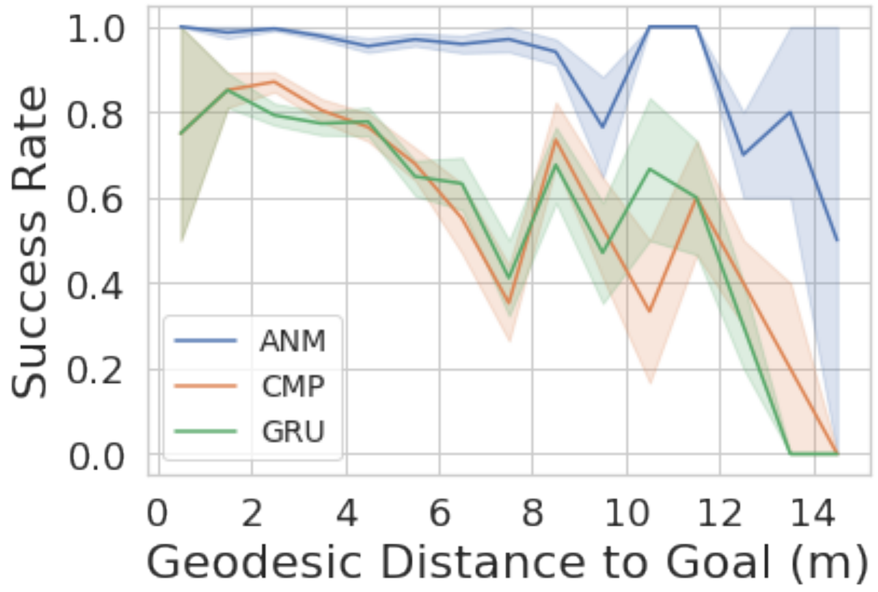
Sample Efficiency

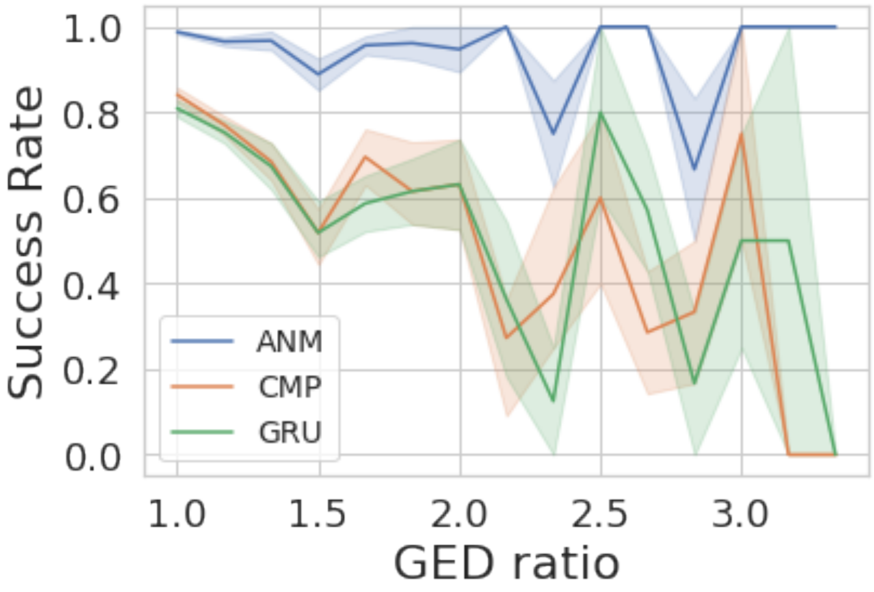
	Training Frames	Succ	SPL
RL + Res18 + GRU + ProjDepth [3]	10 million	0.640	0.461
RL + Res18 + GRU + ProjDepth [3]	75 million	0.678	0.486
IL + Cognitive Mapping & Planning [4]	10 million	0.738	0.683
Active Neural Mapping	1 million	0.789	0.703
Active Neural Mapping	10 million	0.951	0.848

> 10x speedup as compared to best IL baseline

*Adapted from [3] Chen el al. ICLR-19, [4] Gupta et al. CVPR-17

Analysis


1.0
0.8
0.6
0.4
0.2
0.0
Success Rate
0
2
4
6
8
10
12
14
Geodesic Distance to Goal (m)
ANM
CMP
GRU


1.0
0.8
0.6
0.4
0.2
0.0
Success Rate
1.0
1.5
2.0
2.5
3.0
GED ratio
ANM
CMP
GRU

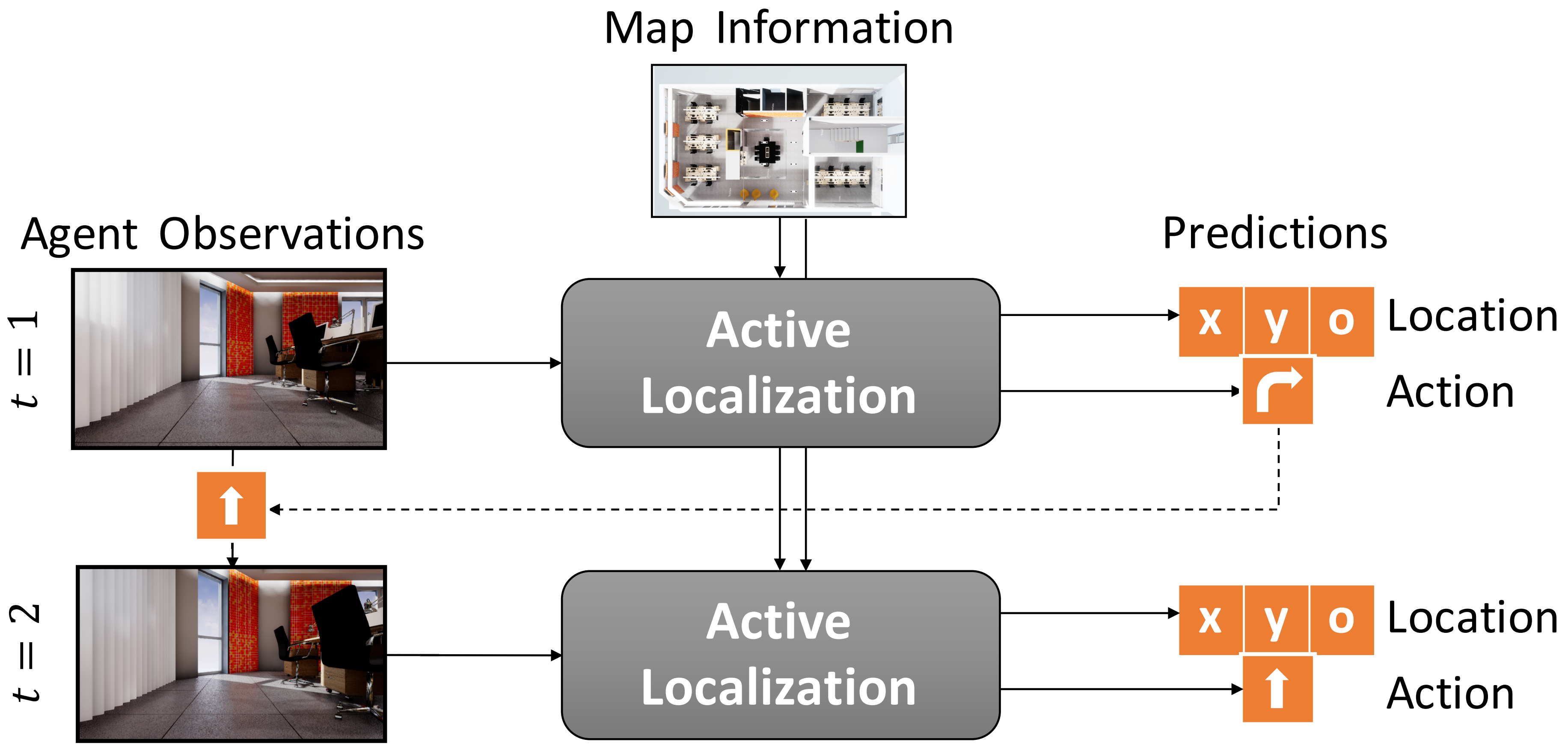
Summary

- Modular navigation model, effective at both
  - Exploration and
  - Pointgoal navigation
- Generalization across goals, tasks, domains
- Effective at long-term planning
- Extensions:
  - Pose Estimation / Odometry
  - Relocalization
  - Semantics

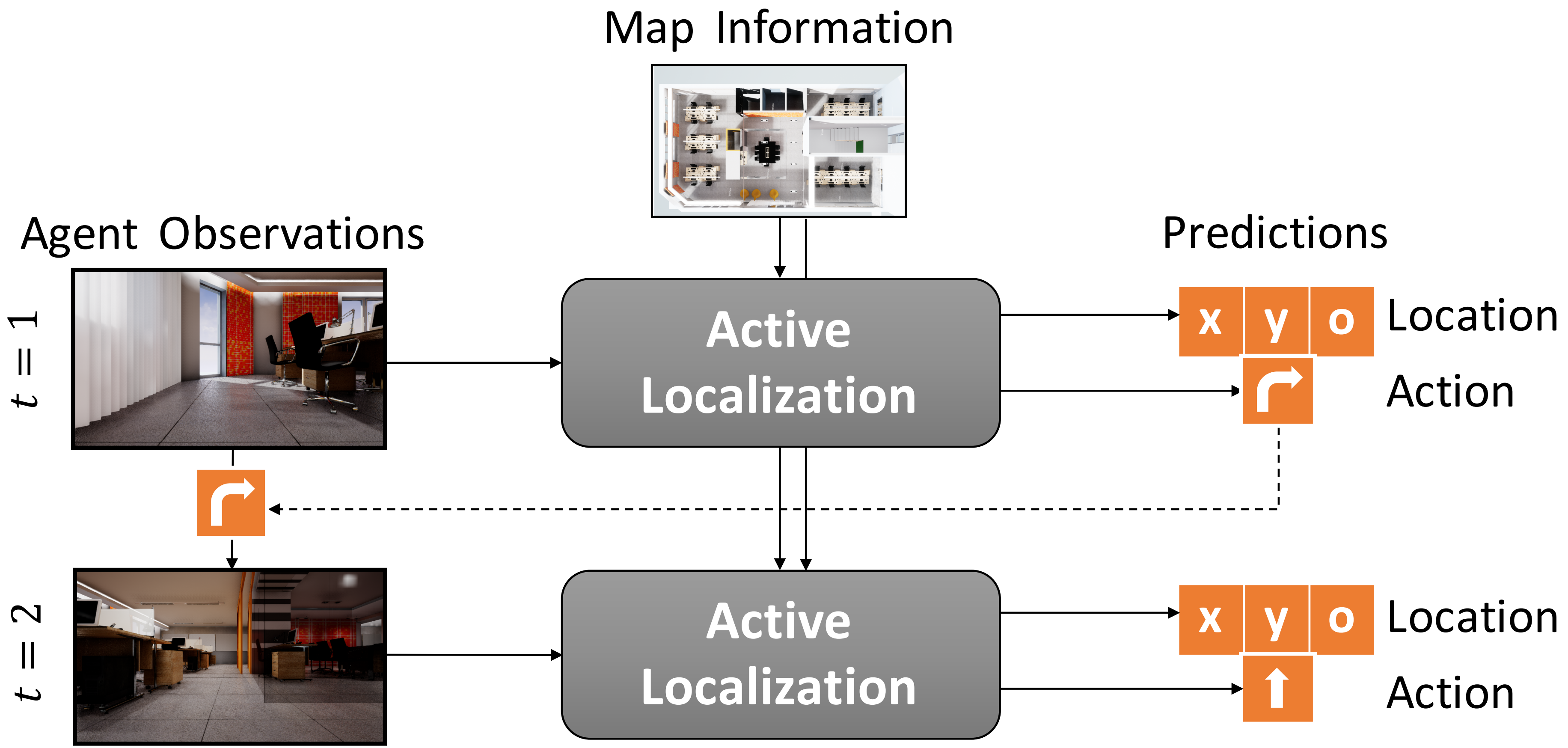
Talk Outline

- Modular Visual Navigation using Active Neural Mapping
- Active Neural Localization: Towards Deep SLAM
- MineRL NeurIPS Competition

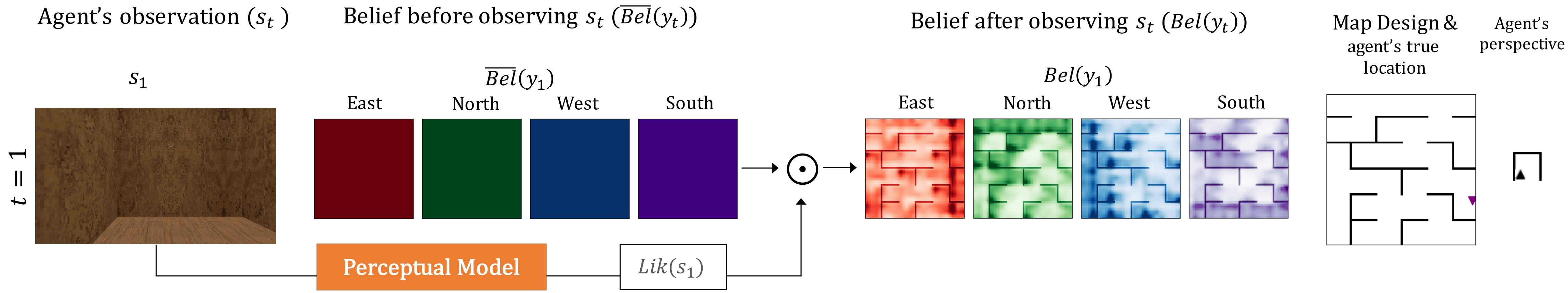
Active Localization



Active Localization

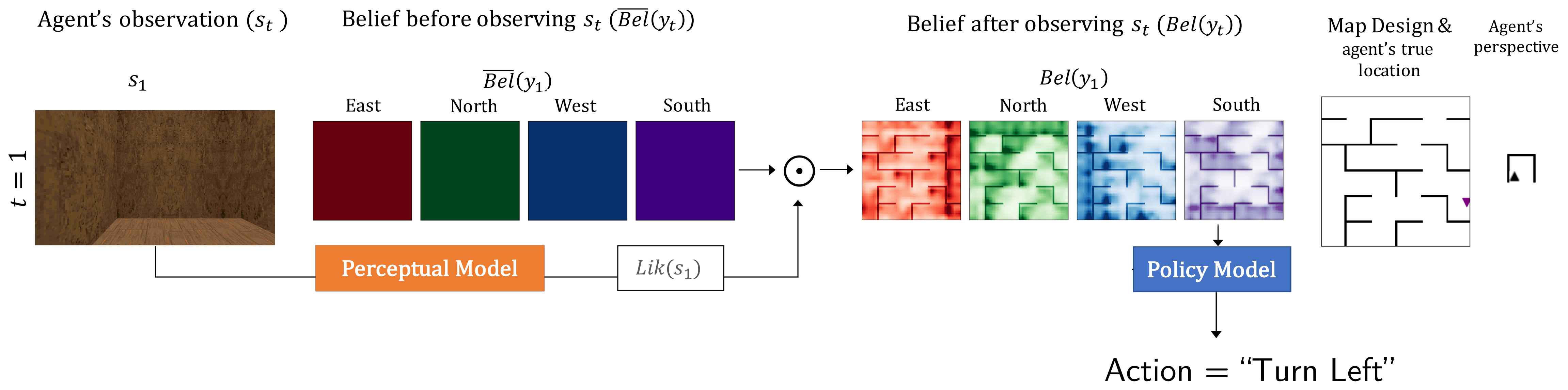


Active Neural Localization



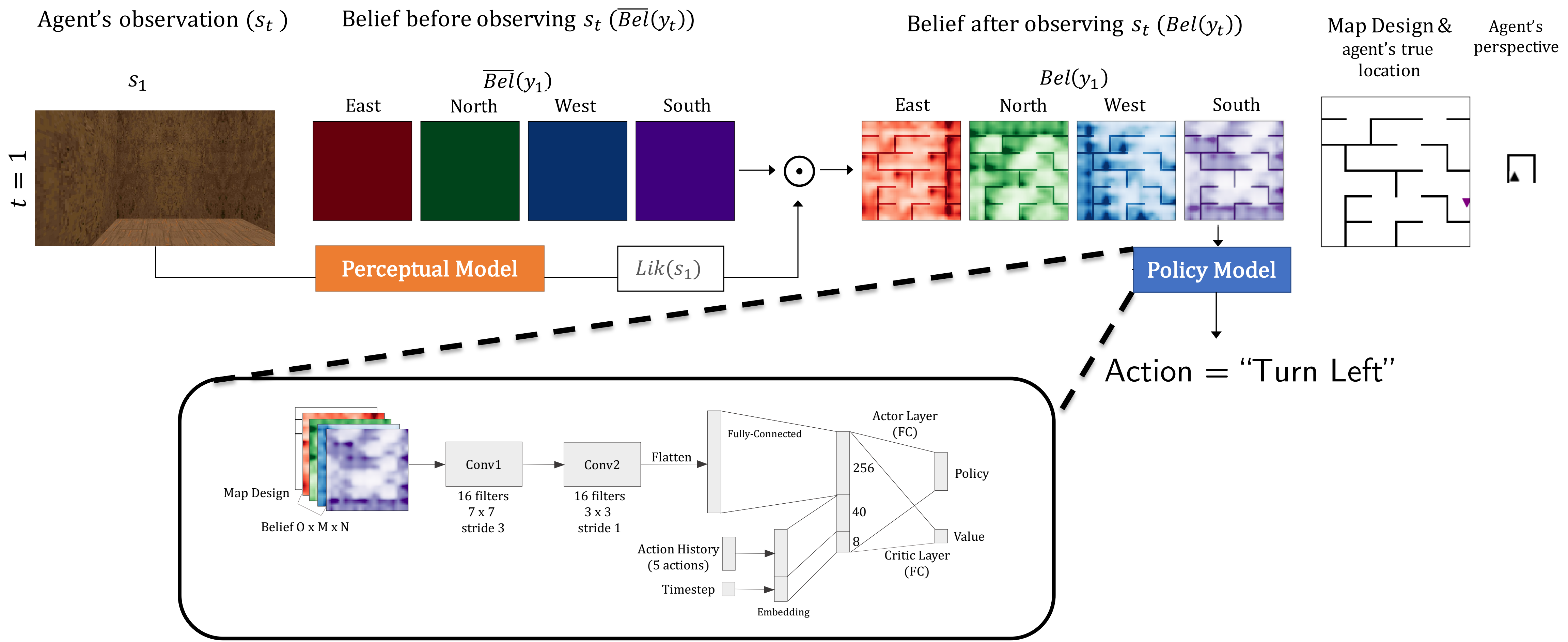
Chaplot, Parisotto, Salakhutdinov, ICLR 2018

Active Neural Localization



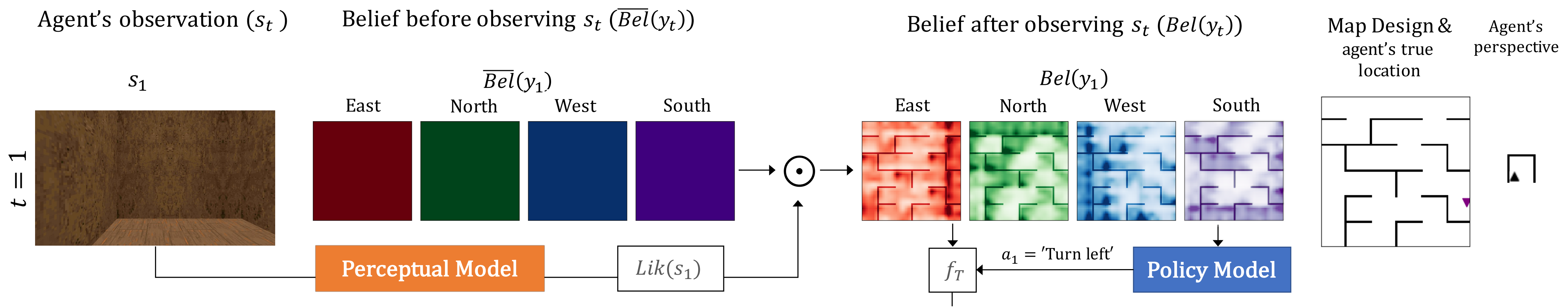
Chaplot, Parisotto, Salakhutdinov, ICLR 2018

Active Neural Localization



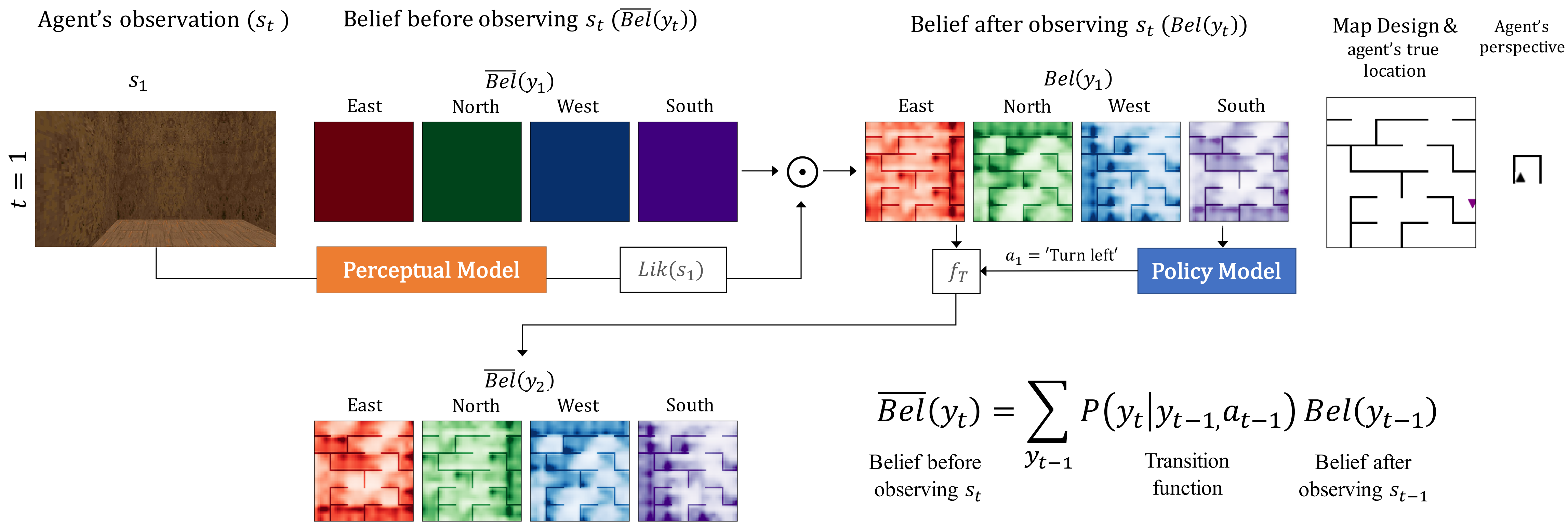
Chaplot, Parisotto, Salakhutdinov, ICLR 2018

Active Neural Localization



Chaplot, Parisotto, Salakhutdinov, ICLR 2018

Active Neural Localization

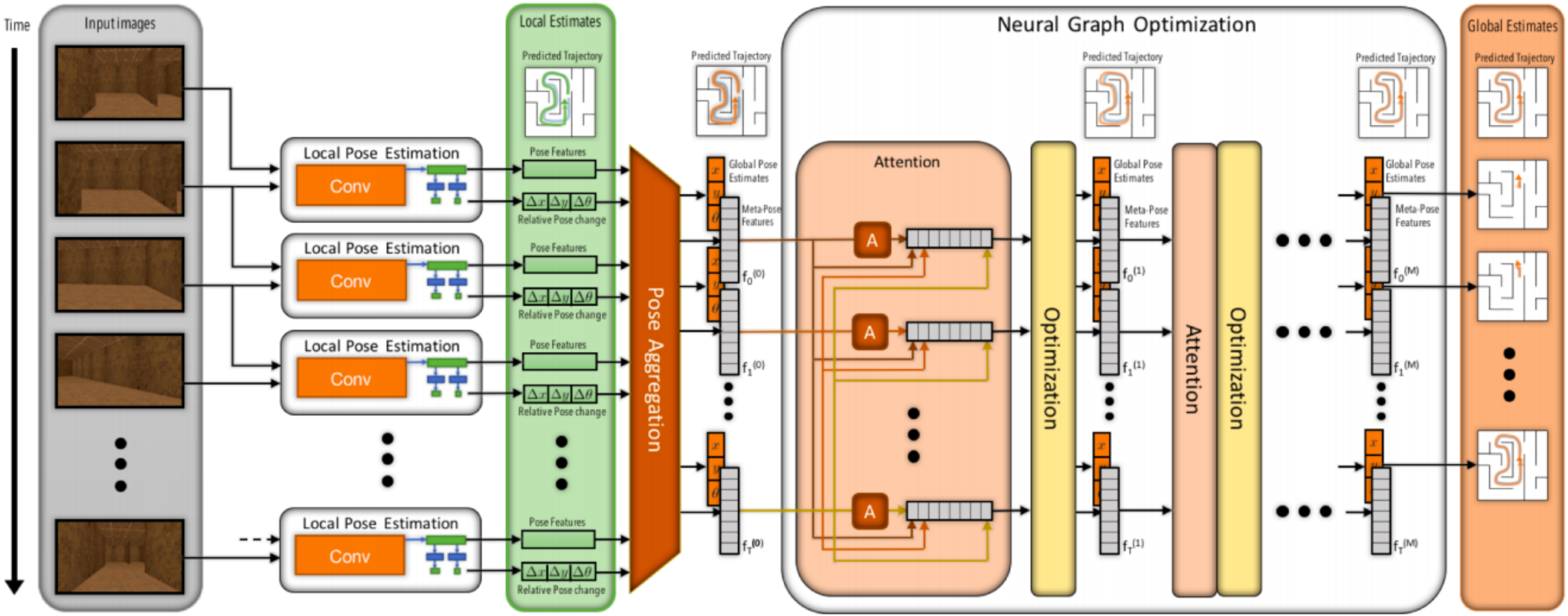


$$\overline{Bel}(y_t) = \sum_{y_{t-1}} P(y_t | y_{t-1}, a_{t-1})\, Bel(y_{t-1})$$

Belief before observing s_t | Transition function | Belief after observing s_{t-1}

Chaplot, Parisotto, Salakhutdinov, ICLR 2018

Pose Estimation: Towards Deep SLAM



Parisotto et al., CVPR Workshop on Visual SLAM 2018

Talk Outline

- Modular Visual Navigation using Active Neural Mapping
- Active Neural Localization: Towards Deep SLAM
- MineRL NeurIPS Competition

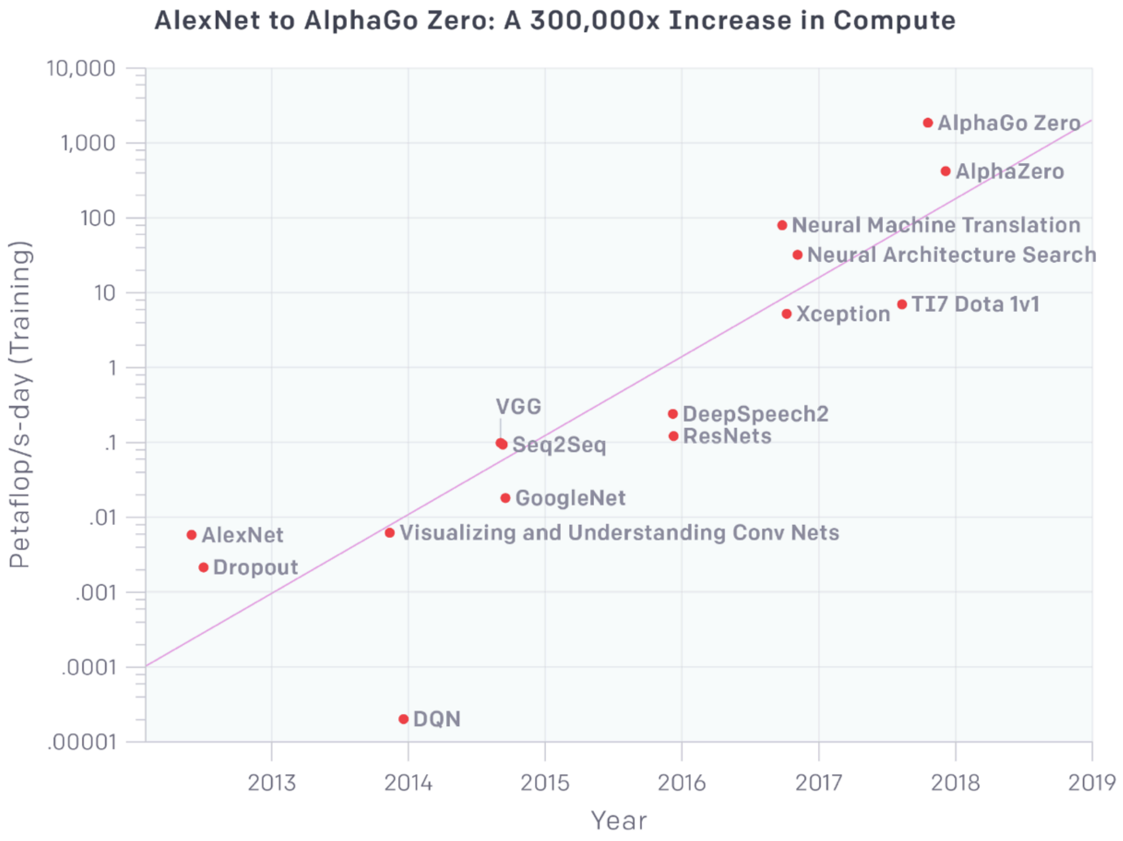
MineRL

Towards Sample Efficient Reinforcement Learning

William H. Guss$*$, Brandon Houghton$*$, Nicholay Topin , Phillip Wang , Cayden Codel , Manuela Veloso and **Ruslan Salakhutdinov**

The growing problem of sample inefficiency in RL

- The **number of environment samples** to train policies on domains of increasing complexity is **growing exponentially**

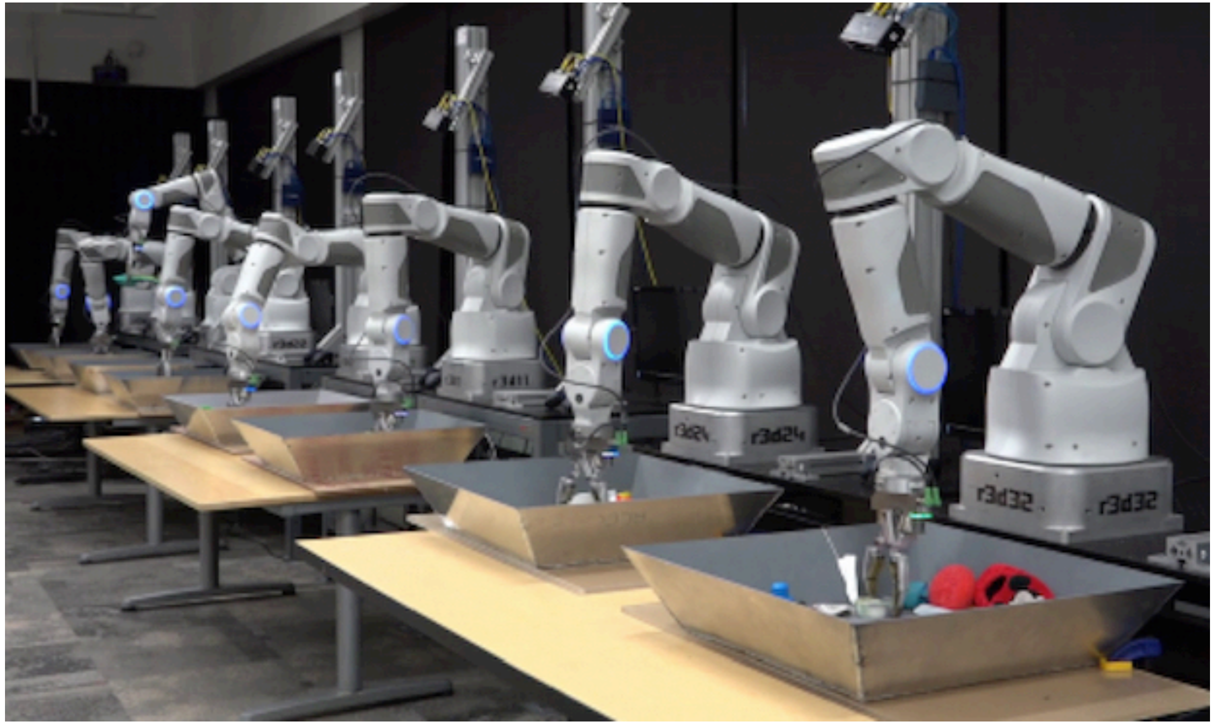


Dario Amodei & Danny Hernandez Open AI 2019.

The growing problem of sample inefficiency in RL

- The **number of environment samples** to train policies on domains of increasing complexity is **growing exponentially**

- Training **complex policies** in real-world environments is quickly becoming **intractable**, without significant infrastructure



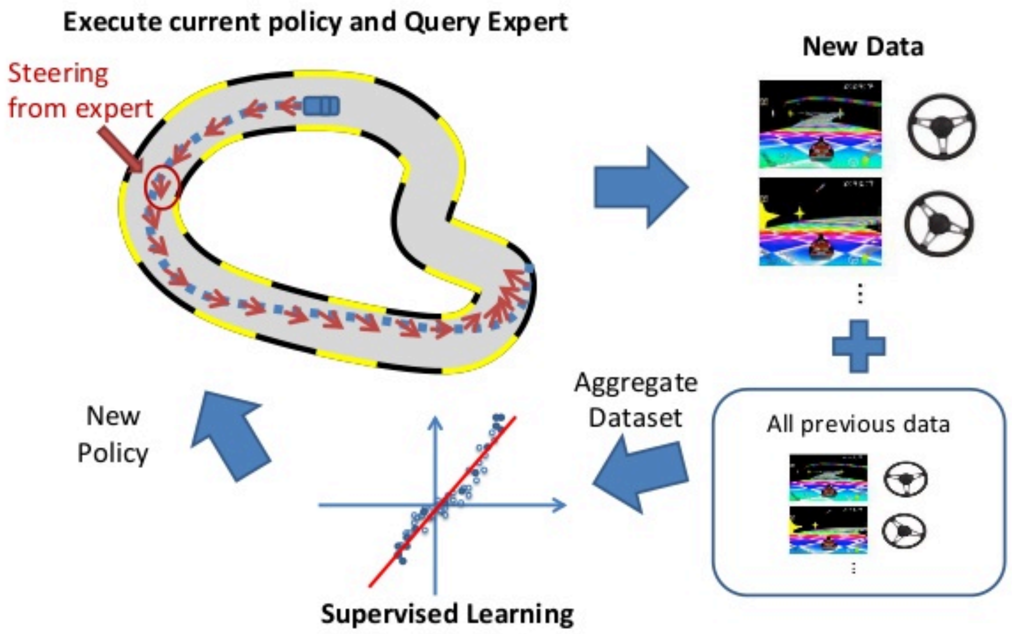
Levine et. al. 2016

Demonstration as an Answer to Sample Inefficiency

- The number of samples required can be drastically reduced using expert demonstrations.

- No open, large-scale dataset of demonstrations across a variety of open/closed world tasks exists

Stéphane. Ross, Geoffrey J. Gordon, and J. Andrew. Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In , 2011.



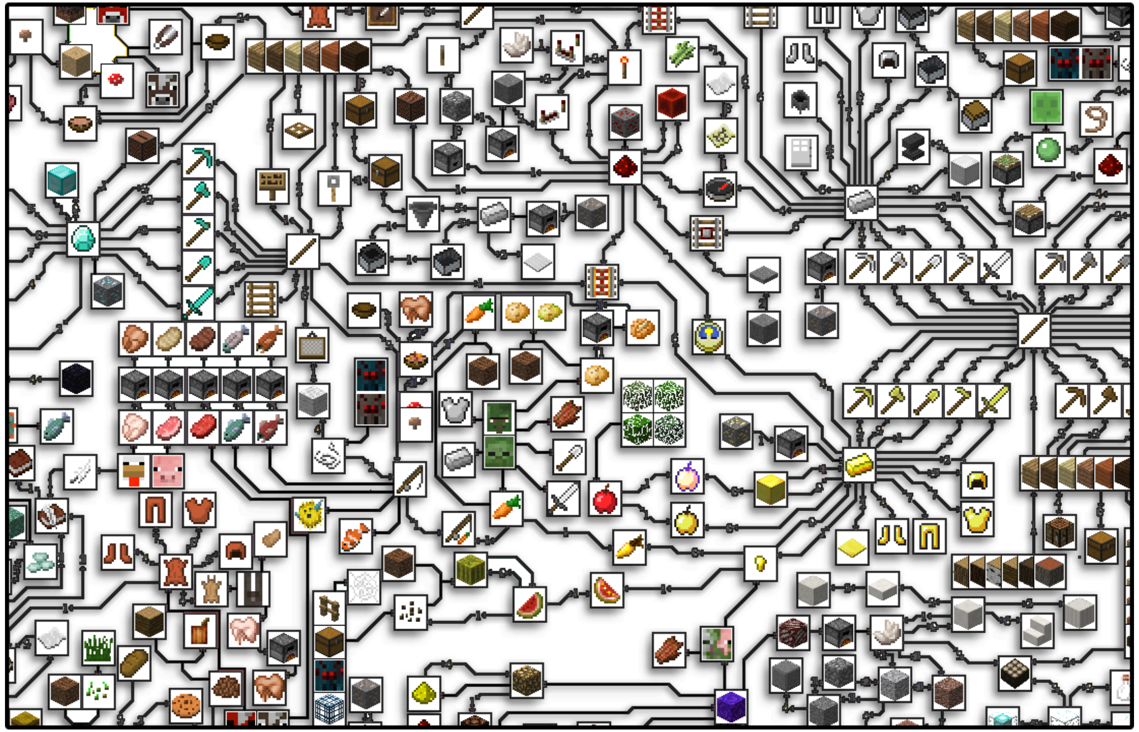
MineRL: A Large-Scale Dataset of Minecraft Demonstrations

William H. Guss∗ , Brandon Houghton∗ , Nicholay Topin , Phillip Wang , Cayden Codel , Manuela Veloso and **Ruslan Salakhutdinov.** IJCAI 2019.

- We have created one of the largest imitation learning datasets with over **60 million frames of recorded** human player data across **6**+ complex tasks in Minecraft.

MineRL: Why Minecraft?

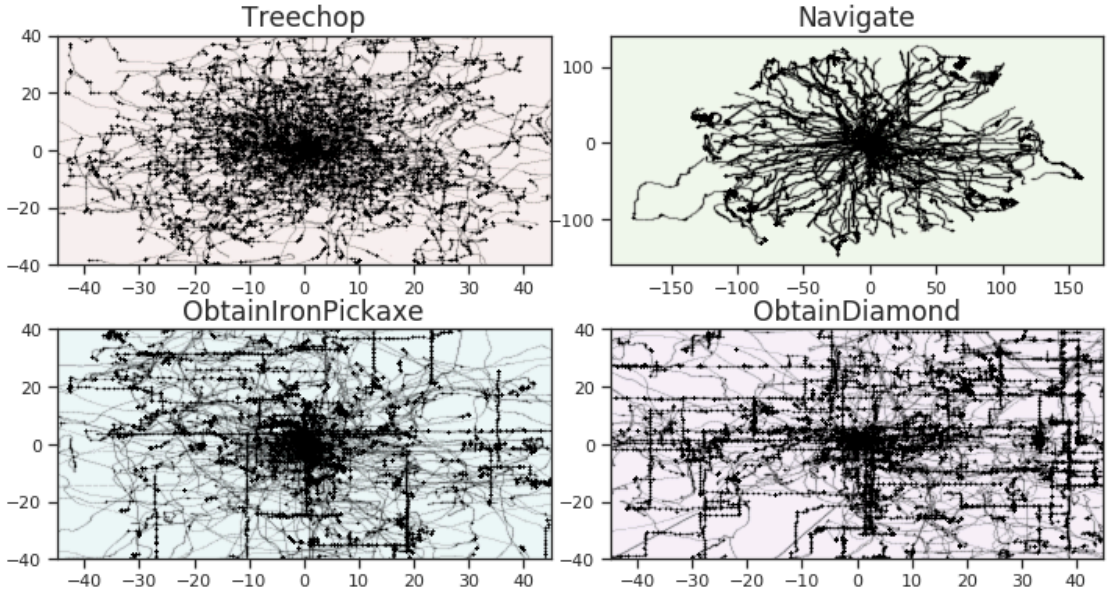
- Open-world, infinite/procedurally generated
- Sparse & dense rewards
- Many innate task hierarchies and subgoals
- Encompasses many of problems we must solve as we approach the problem of general AI.



A glimpse into the Minecraft item hierarchy

MineRL: Dataset Details

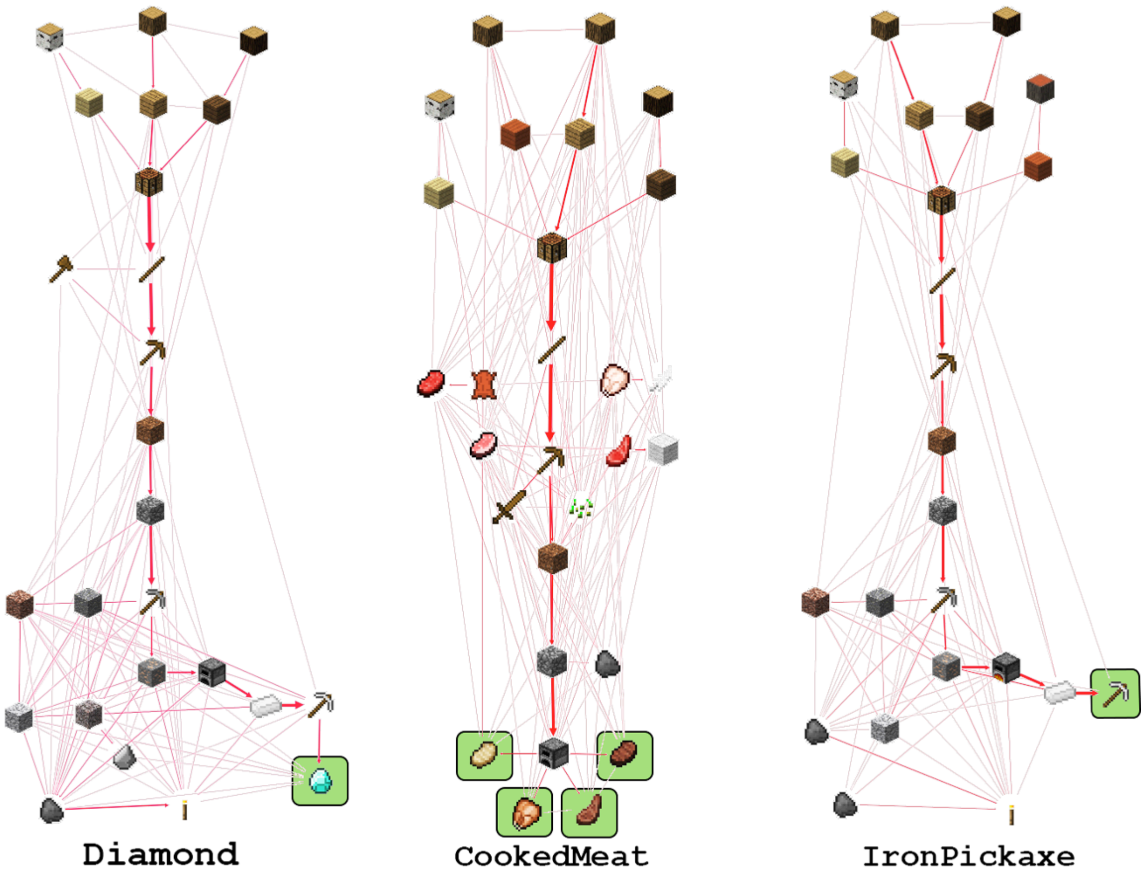
- Consists of over 500+ hours of human demonstrations over 1000+ unique player sessions.
- Rich set of annotations including: **subtask completion, rewards, player meta-data, gamestate.**
- **Rerenderable!** We record game-state not just player-pixels



Plots of XY positions of players in several tasks (diversity & rich annotations)

MineRL: Hierarchality of Data

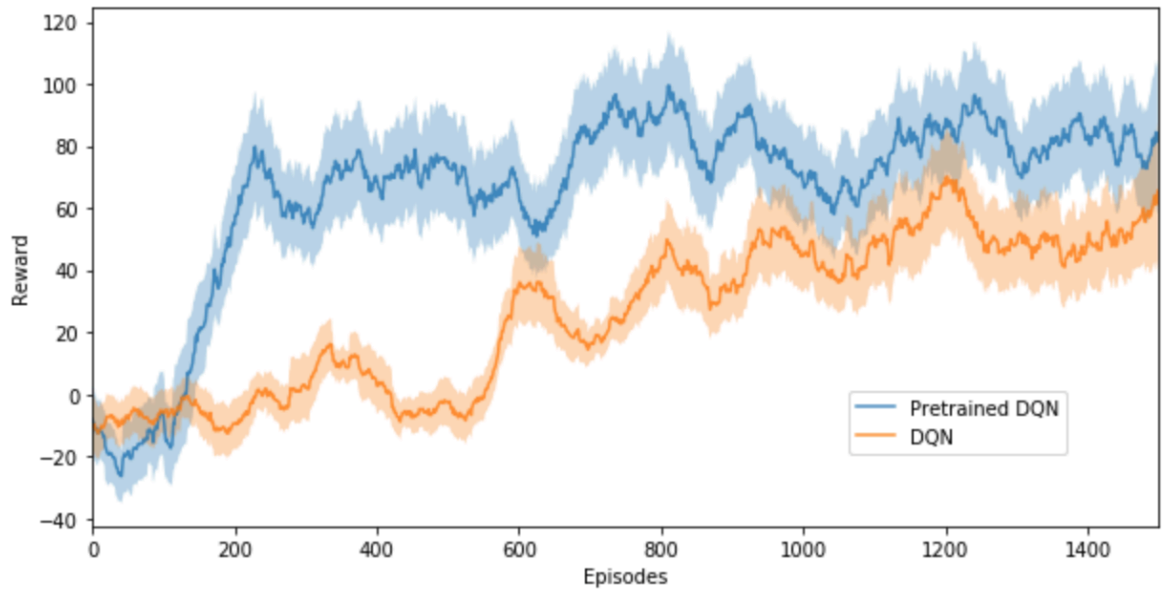
- Players complete sparsely rewarded tasks **following a specific task hierarchy/dependence graph.**
- Many ways to obtain an item, but data exhibits the existence of **canonical pathways.**



MineRL: Expert demonstrations help

- On the **Navigate** task, using the MineRL-v0 dataset helps **drastically reduce** the number of samples for standard algorithms.

- However, **better algorithms still need to be developed**, especially for the long-term, hierarchical tasks exhibited in Minecraft.



MineRL: Get started now!

http://minerl.io/

Thank you