The Empirical Impact of Forgetting and Transfer in Continual Visual Odometry

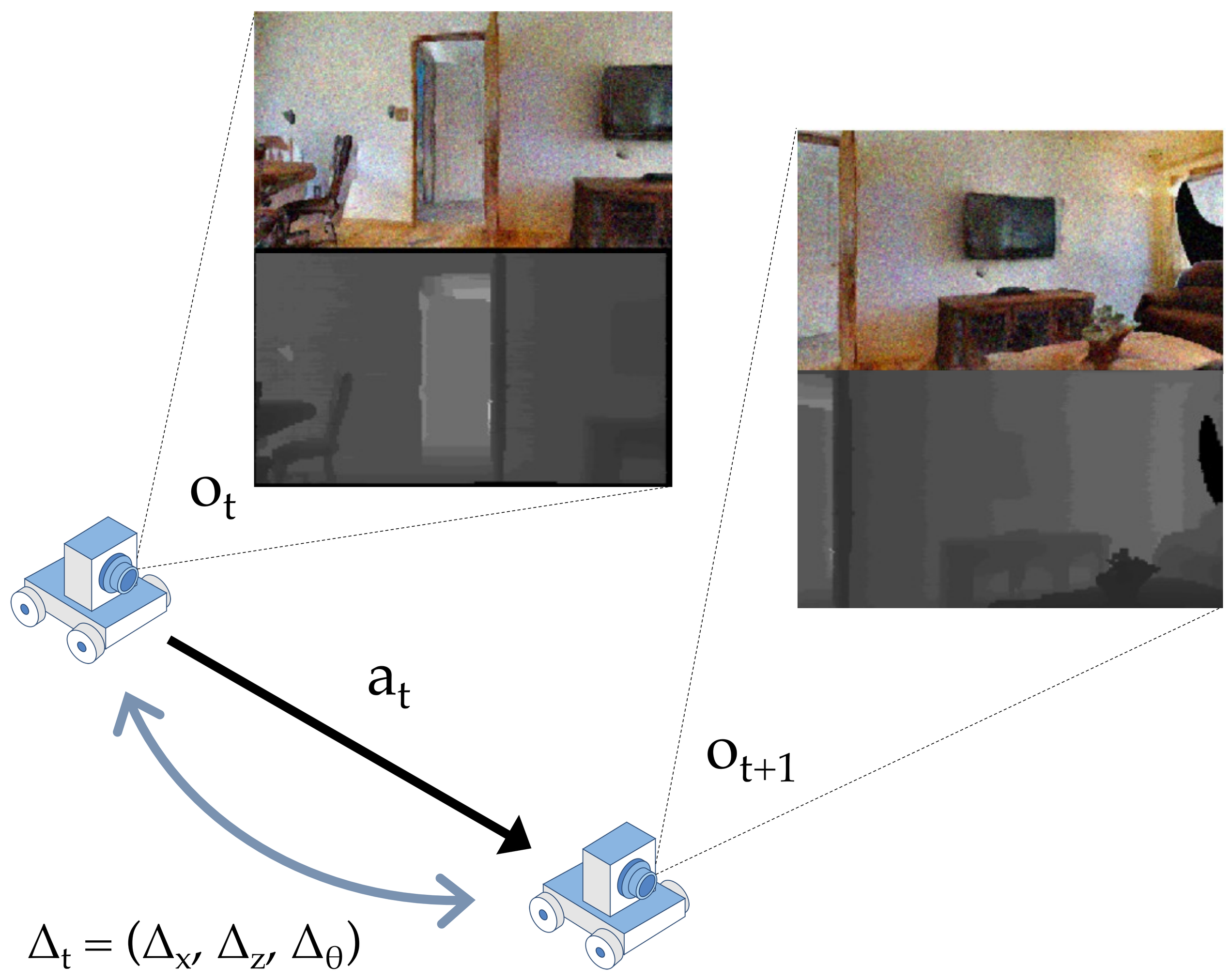
Paolo Cudrano* Xiaoyu Luo* Matteo Matteucci

AIRLAB – Artificial Intelligence and Robotics Lab · Politecnico di Milano 1863 · CoLLAs

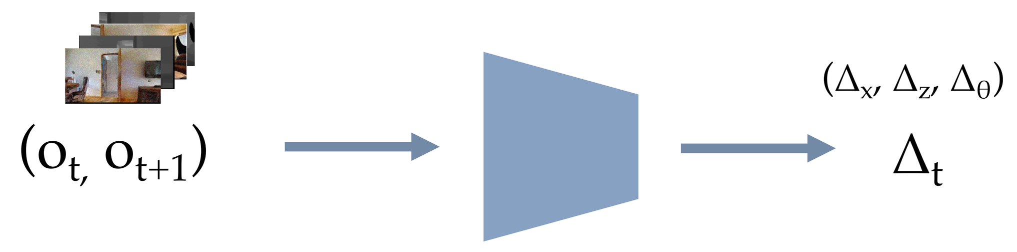
Supervised continual learning has been studied extensively on internet-based datasets, but **embodied scenarios** are one of the most promising directions.

Does traditional continual learning extend to embodied scenarios?

Visual Odometry



Estimate the robot displacement Δ_t, given consecutive RGB-D observations (o_t, o_{t+1}).

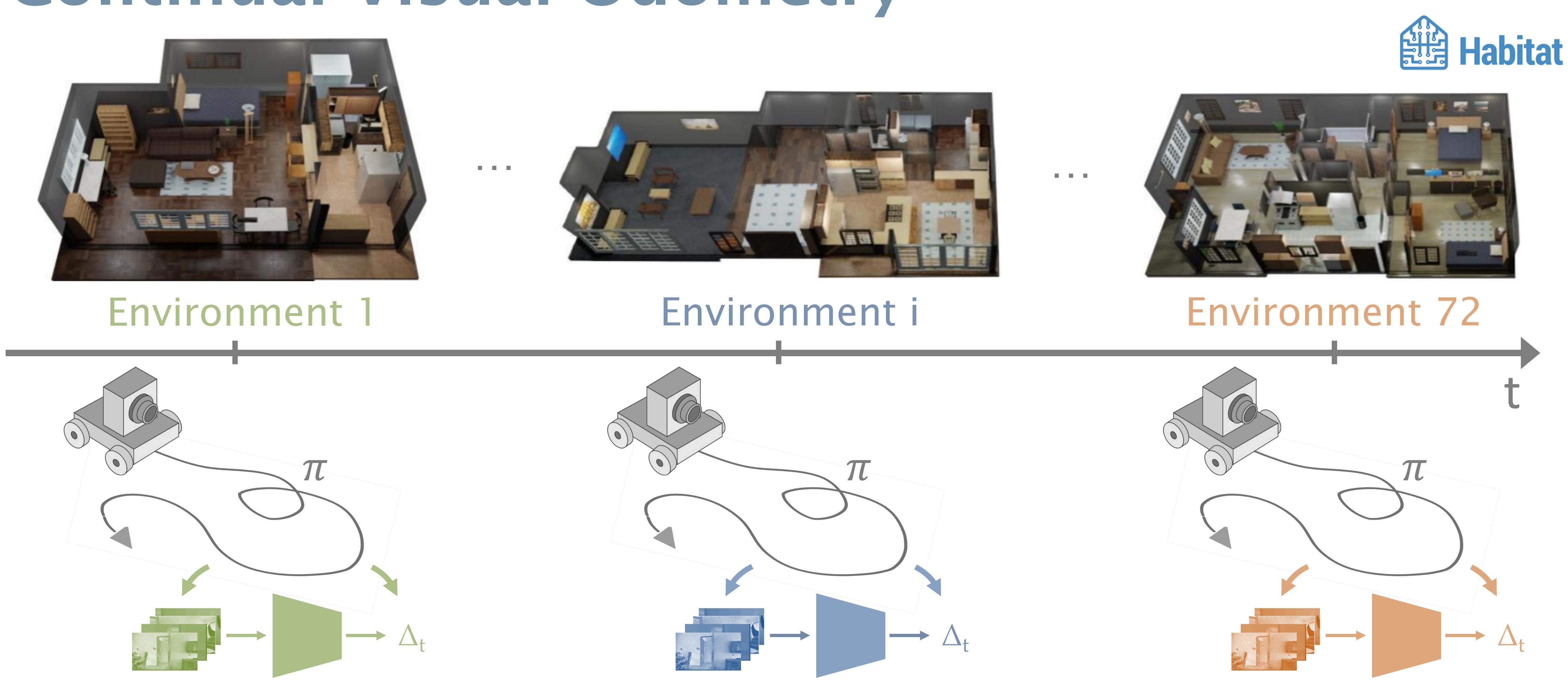


Continual Learning in Embodied Scenarios

CL Scenario Granularity			Level of interaction
environment-level ← our work			
trajectory-level	our work →	passive interaction	
observation-level		active interaction	

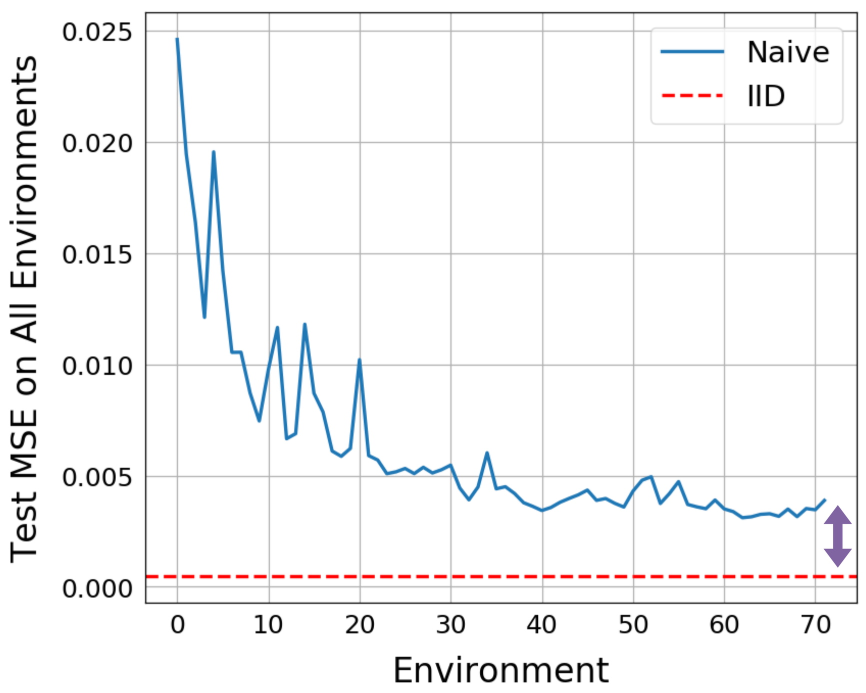
CL Scenario Granularity: environment-level, trajectory-level, observation-level. Level of interaction: no interaction, passive interaction, active interaction. Our work: environment-level, passive interaction.

Continual Visual Odometry



Domain-incremental, single regression task

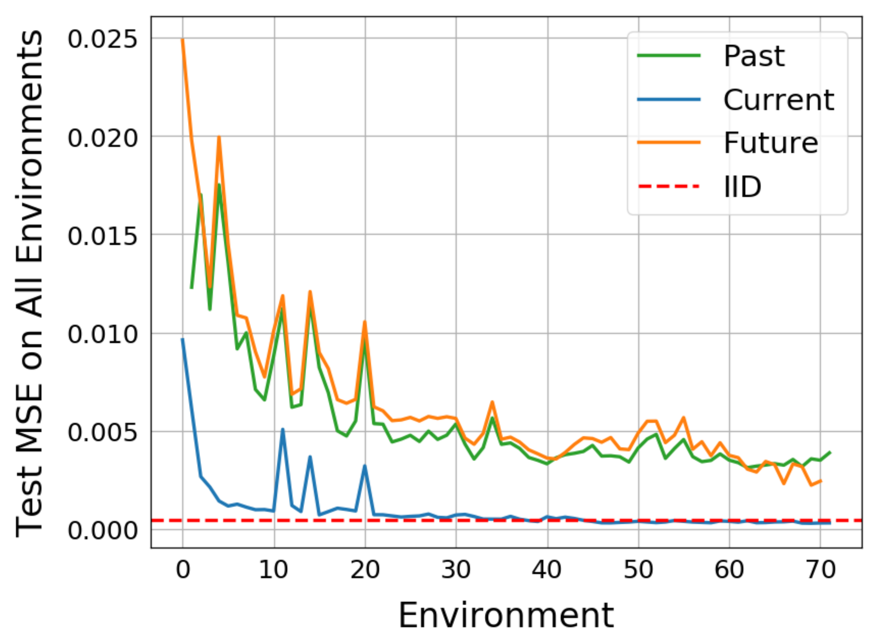
Baselines



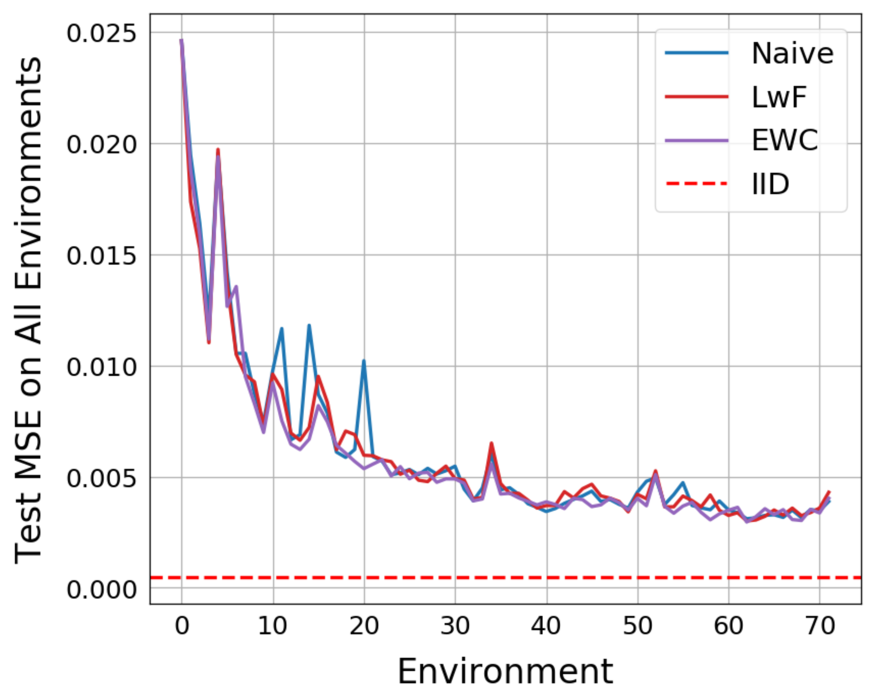
Performance is given by the MSE over test samples taken from all 72 environments. We compute this after training on each environment.

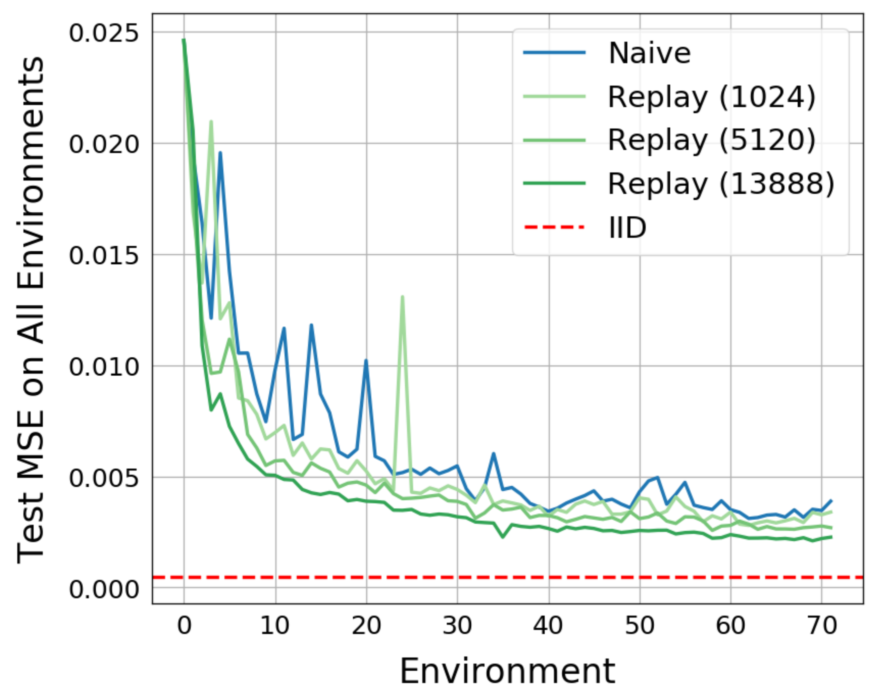
Naive incremental training improves performance over time, but results in a large performance gap w.r.t. IID.

A general improvement on future environments suggests overall forward transfer, yet Naive overly focuses on the current scene, matching IID performance at the expense of forgetting anything specifically learned on past scenes.



Regularization and Replay

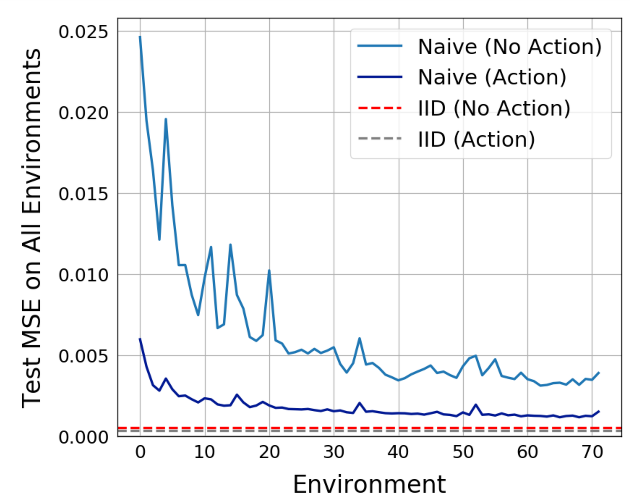




Regularization strategies show no improvement from the Naive baseline.
Replay partially closes the gap with IID when using large buffers (longer training times).

Additional Information

In real-world scenarios, the agent has access to additional information, such as which motion action was just performed. Providing this information to the model almost closes the gap between Naive and IID, as it completely changes the learning problem.



Not really.

Embodied scenarios introduce:

- **Additional challenges** (longer sequences, sensor and control noise, less common tasks, complex real-world dynamics).
- **New problem dimensions** (continual scenario granularity, levels of interaction).

Even in a **simple embodied setup** (domain incremental, single task, coarse granularity, passive-only environmental interaction), naive incremental learning presents a **large performance gap wrt IID**, and known CL techniques are not sufficient to close it.

However, the added complexities associated with an embodied system **can actually become advantageous**: multiple sensors and actuators mean that additional information is often available, and can significantly **alter the learning problem**.