Master's Thesis Proposal: Comparing Visual and Graph-Based Representations for Learning General Policies in Classical Planning Domains

November 23, 2025

Abstract

This proposal describes research to compare visual representation learning (using CNNs on image observations) with graph-based representations (using GNNs on relational state encodings) for learning generalized policies in classical planning domains. While recent work has demonstrated that GNN-based approaches can learn policies that generalize across instance sizes, it remains unclear whether visual encoders can achieve similar generalization when given appropriate image-based state representations. This project is about doing a systematic comparison using, for example, the PDDLGym environments, which provide both symbolic PDDL states and visual renderings. The study will evaluate various CNN-based RL algorithms (e.g., DQN, PPO, etc, with visual encoders) against GNN baselines, analyzing generalization performance, sample efficiency, and the types of structural invariances captured by each representation. This work aims to provide insights into the necessity of explicit relational inductive biases for generalized planning.

1 Introduction

Learning generalized policies for classical planning domains has traditionally relied on symbolic or graph-based representations that explicitly encode the relational structure of PDDL states [12, 13]. Graph neural networks (GNNs) have shown particular promise due to their ability to capture permutation invariance and relational reasoning, leading to policies that generalize across instances with varying object counts and configurations.

Concurrently, deep reinforcement learning with visual observations has achieved remarkable success in complex domains like Atari games [1], robotics [2], and 3D navigation [3]. Convolutional neural networks (CNNs) can learn powerful spatial features from raw pixels, and modern RL algorithms combined with visual encoders have solved tasks previously thought to require explicit symbolic reasoning.

This raises a natural question: *Can visual RL methods, using CNN encoders on image-based state representations, learn generalized policies that transfer across classical planning instances as effectively as GNN-based methods?* Understanding the answer has both theoretical implications (what inductive biases are necessary for generalized planning?) and practical implications (can we leverage the extensive visual RL tooling and pre-trained models for planning tasks?).

2 Background and Related Work

GNN-based generalized planning. Recent work has demonstrated that graph neural networks can learn generalized policies for classical planning [12]. By encoding PDDL states as graphs (objects as nodes, predicates as edges), GNNs naturally capture permutation invariance and relational structure, enabling policies to generalize across instances of different sizes. Follow-up work has explored policy gradient methods [13] and goal-conditioned RL with first-order representations [14].

Visual RL and representation learning. Deep Q-Networks (DQN) [1] pioneered the use of CNNs for RL with visual observations, learning directly from pixels in Atari games. Subsequent work has improved sample efficiency through techniques like auxiliary tasks [5], world models [6], and contrastive learning [7]. Modern visual RL algorithms include Rainbow DQN, PPO with visual encoders, SAC for continuous control, and model-based approaches like Dreamer [8].

PDDLGym and planning benchmarks. PDDLGym [4] provides a standardized interface for classical planning domains with OpenAI Gym compatibility. Importantly, it offers visual rendering capabilities for many domains, making it possible to train RL agents on pixel observations. This enables direct comparison between symbolic/graph-based and visual approaches on the same underlying planning problems.

Visual reasoning and structured representations. Recent work has explored whether neural networks can learn structured reasoning from visual observations. Some studies suggest that CNNs can implicitly learn object-centric representations [10], while others argue that explicit structural inductive biases are necessary for systematic generalization [11]. This work will contribute to this debate in the context of classical planning.

Bridging visual and symbolic AI. This work bridges visual RL and symbolic planning by asking whether visual representations can serve as an alternative to explicit graph structures for learning generalized policies. The comparison provides insights into what types of inductive biases are truly necessary for generalization in classical planning domains.

3 Research Questions

1. Can CNN-based visual RL methods learn generalized policies that transfer to larger unseen instances in classical planning domains?
2. How does the generalization performance of visual RL compare to GNN-based methods across different planning domains?
3. What types of visual representations (raw pixels, spatial maps, object-centric renderings) work best for generalized planning?
4. Which visual RL algorithms (DQN variants, PPO, SAC, model-based) are most effective for this task?

5. Do visual methods capture similar structural invariances (permutation, translation) as GNN methods?

6. What is the sample efficiency trade-off between visual and graph-based approaches?

7. In which domains does explicit relational structure provide clear advantages, and where might visual approaches be competitive?

4 Methodology

4.1 Overview

The proposed study consists of the following components:

1. **Environment setup.** Use PDDLGym [4] to access classical planning domains with both symbolic PDDL states and visual renderings. Focus on domains with clear visual structure: Blocksworld, Sokoban, Hanoi, Gripper, and others. Many domains have existing visual renderings, so start with those.

2. **Visual encoders.** Design and implement various CNN architectures to encode visual observations:
   - **Nature CNN:** Standard architecture from DQN paper [1]
   - **IMPALA CNN:** Residual architecture from IMPALA [9]
   - Do some literature research to identify other promising CNN architectures for visual RL.

3. **Visual RL baselines.** Implement and train multiple visual RL algorithms:
   - **DQN and variants:** Standard DQN, Double DQN, Dueling DQN, Rainbow
   - **Policy gradient methods:** PPO with CNN encoder
   - If you find some interesting architecture that you would like to try out.

4. **GNN baselines.** Use existing GNN-based generalized planning methods [12, 13] as comparison baselines. I can help you with setting these up.

5. (Additional) **Hybrid approaches.** Explore combinations:
   - CNN+GNN: Extract visual features with CNN, then process with GNN
   - Attention mechanisms: Use visual attention to build object representations
   - Object-centric visual encoders: Explicitly segment objects in images

4.2 Training Paradigm

- **Training set:** Small to medium-sized instances (e.g., 4-8 objects for Blocksworld, 5×5 to 7×7 for Sokoban)

- **Test set:** Larger instances (12-20 objects, 10×10 grids) to evaluate generalization

- **Reward structure:** Sparse rewards (goal achievement) with optional step penalties

4.3 Evaluation Dimensions

- **Coverage:** Number of test instances solved
- **Trajectory lengths:** Average length of solution trajectories
- **Sample efficiency:** Number of environment interactions to reach top performance
- **Generalization gap:** Performance drop from training to test instance sizes

5 Datasets and Benchmarks

The PDDLGym environment will be used initially due to its dual symbolic and visual interface. But some research could be done into other environments if time permits.

6 Baselines

- **GNN-based methods:** Policies from [12, 13]
- **Symbolic planners:** Fast Downward, LAMA for reference solution quality
- Other ideas?

7 Expected Contributions

- **Systematic comparison:** First comprehensive study comparing visual and graph-based RL for generalized planning
- **Empirical insights:** Evidence on whether explicit relational inductive biases are necessary for generalization in classical planning
- **Best practices:** Guidelines for visual representation design in planning domains
- **Algorithm recommendations:** Identification of visual RL algorithms best suited for generalized planning

8 Recommended Reading

- [1] for understanding DQN and visual RL fundamentals.
- [4] for understanding the PDDLGym environment.
- [12] for understanding Relational-GNNs and general policies in classical planning.

9 Library Suggestions

- PDDLGym (`https://github.com/tomsilver/pddlgym`) for accessing classical planning domains with visual renderings.
- For encoding PDDL states as GNNs *mimir-rgnn* (`https://github.com/simon-stahlberg/mimir-rgnn`) can be used, which is based on the architectures in [12].
- For running baseline RL algorithms and to use as base for the new implementation, *mimir-rl* (`https://github.com/simon-stahlberg/mimir-rl`) can be used
- *mimir* (`https://github.com/simon-stahlberg/mimir`) is a library for generalized planning that includes PDDL parsing, state space expansion and is a good library for functionalities around classical planning.

References

[1] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015. URL: `https://www.nature.com/articles/nature14236`.

[2] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-End Training of Deep Visuomotor Policies. *Journal of Machine Learning Research*, 17(39):1–40, 2016. URL: `https://arxiv.org/abs/1504.00702`.

[3] Piotr Mirowski, Razvan Pascanu, Fabio Viola, Hubert Soyer, Andrew J. Ballard, Andrea Banino, Misha Denil, Ross Goroshin, Laurent Sifre, Koray Kavukcuoglu, Dharshan Kumaran, and Raia Hadsell. Learning to Navigate in Complex Environments. In *ICLR*, 2017. URL: `https://arxiv.org/abs/1611.03673`.

[4] Tom Silver, Rohan Chitnis, Joshua Tenenbaum, Leslie Pack Kaelbling, and Tomás Lozano-Pérez. Learning Symbolic Operators for Task and Motion Planning. In *IROS*, 2021. (PDDLGym environment introduced). URL: `https://github.com/tomsilver/pddlgym`.

[5] Max Jaderberg, Volodymyr Mnih, Wojciech Marian Czarnecki, Tom Schaul, Joel Z. Leibo, David Silver, and Koray Kavukcuoglu. Reinforcement Learning with Unsupervised Auxiliary Tasks. In *ICLR*, 2017. URL: `https://arxiv.org/abs/1611.05397`.

[6] David Ha and Jürgen Schmidhuber. World Models. In *NeurIPS*, 2018. URL: `https://arxiv.org/abs/1803.10122`.

[7] Michael Laskin, Aravind Srinivas, and Pieter Abbeel. CURL: Contrastive Unsupervised Representations for Reinforcement Learning. In *ICML Workshops / PMLR*, 2020. URL: `https://arxiv.org/abs/2004.04136`.

[8] Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to Control: Learning Behaviors by Latent Imagination. In *ICLR*, 2020. URL: `https://arxiv.org/abs/1912.01603`.

[9] Lasse Espeholt, Hubert Soyer, Remi Munos, Karen Simonyan, Volodymir Mnih, Tom Ward, Yotam Doron, Vlad Firoiu, Tim Harley, Iain Dunning, Shane Legg, and Koray Kavukcuoglu. IMPALA: Scalable Distributed Deep-RL with Importance Weighted Actor-Learner Architectures. In *ICML*, 2018. URL: `https://arxiv.org/abs/1802.01561`.

[10] Klaus Greff, Raphaël Lopez Kaufman, Rishabh Kabra, Nick Watters, Chris Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, and Alexander Lerchner. Multi-Object Representation Learning with Iterative Variational Inference. In *ICML*, 2019. URL: `https://arxiv.org/abs/1903.00450`.

[11] Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, Caglar Gulcehre, Francis Song, Andrew Ballard, Justin Gilmer, George Dahl, Ashish Vaswani, Kelsey Allen, Charles Nash, Victoria Langston, Chris Dyer, Nicolas Heess, Daan Wierstra, Pushmeet Kohli, Matt Botvinick, Oriol Vinyals, Yujia Li, and Razvan Pascanu. Relational inductive biases, deep learning, and graph networks. *arXiv:1806.01261*, 2018. URL: `https://arxiv.org/abs/1806.01261`.

[12] Simon Ståhlberg, Blai Bonet, and Hector Geffner. Learning General Optimal Policies with Graph Neural Networks: Expressive Power, Transparency, and Limits. In *ICAPS*, 2022.

[13] Simon Ståhlberg, Blai Bonet, and Hector Geffner. Learning General Policies with Policy Gradient Methods. 2023. (Conference / proceedings).

[14] Simon Ståhlberg, Blai Bonet, and Hector Geffner. First-Order Representation Languages for Goal-Conditioned RL. In *PRL*, 2025. URL: `https://prl-theworkshop.github.io/prl2025-icaps/papers/2.pdf`.