

Grid World Transfer Learning - from Small to Large Environments



Leibniz
Universität
Hannover



Florian Fingscheidt, Merle Krauss, Mathis Pöhlse
Poster Presentations in context of Reinforcement Learning Lecture

1

TL;DR

- We compare transfer learning and curriculum learning to baseline training for DQN agents in grid environments with different sizes.
- **Transfer learning:** Faster convergence, fewer training steps than baseline.
- **Curriculum learning:** Outperforms both baseline and direct transfer in most cases.

2

Motivation & Problem Setting

Motivation

- DRL agents (e.g., DQNs) need long training times for good performance.
- Grid worlds are simple but ideal for controlled DRL experiments.
- Transfer learning has been proven to speed up learning for other agents such as PPO [Chevalier-Boisvert et al. 2023].
- Goal: Reduce training time without sacrificing performance.

3

Approach

- We use an empty grid environment to pretrain our agent to navigate to a defined goal.
- For the grid environment we use Gym's Minigrid [Chevalier-Boisvert et al. 2023].
- For the agent we use the DQN implementation from Stable Baseline3 [Raffin et al. 2021].
- We used the built-in CnnPolicy but replace its default feature extractor with a **custom designed neural network**:

CNN

$h \times w \times 1$

Conv(3 × 3, 32)

$h \times w \times 32$

Conv(2 × 2, 64)

$h \times w \times 64$

Dropout(0.2)

$h \times w \times 64$

AdaptiveAvgPool(3 × 3)

$3 \times 3 \times 64$

Flatten

576

Linear Head

Linear(576, 128)

128

Dropout(0.2)

128

Linear(128, 128)

128

Experiment setup:

- Same hyperparameter setup across all experiments.
- 100,000 training steps per experiment.
- 20 random seeds for statistical robustness.

4

Key Insights

Performance metric:

- Log mean episode reward per step.
- Performance curves show **running maximum rewards** with 95% confidence bands across 20 seeds.

Transfer learning (Small → Medium):

- Baseline: Train DQN from scratch on 7x7 grid.
- Transfer: Train on 5x5 → finetune on 7x7 grid.

Timesteps	7x7 baseline (n=20)	Transfer 5x5 to 7x7 (n=20)
0	0.0	0.0
20000	0.15	0.45
40000	0.35	0.65
60000	0.55	0.75
80000	0.70	0.80
100000	0.80	0.85

Transfer and curriculum learning (Small → Large):

- Baseline: Train DQN from scratch on 9x9 grid.
- Transfer: Train on 5x5 → finetune on 9x9 grid.
- Curriculum: Train on 5x5 → transfer to 7x7 → transfer to 9x9 grid.

Timesteps	9x9 Baseline (n=20)	Transfer 5x5 to 9x9 (n=20)	Curriculum 5x5 to 9x9 (n=20)
0	0.0	0.0	0.0
20000	0.10	0.30	0.40
40000	0.25	0.50	0.70
60000	0.40	0.60	0.80
80000	0.55	0.65	0.85
100000	0.65	0.70	0.85

5

Future Works

- Apply to more complex environments with obstacles like **walls** or **lava**.
- Explore reverse transfer (large → small) for robustness.