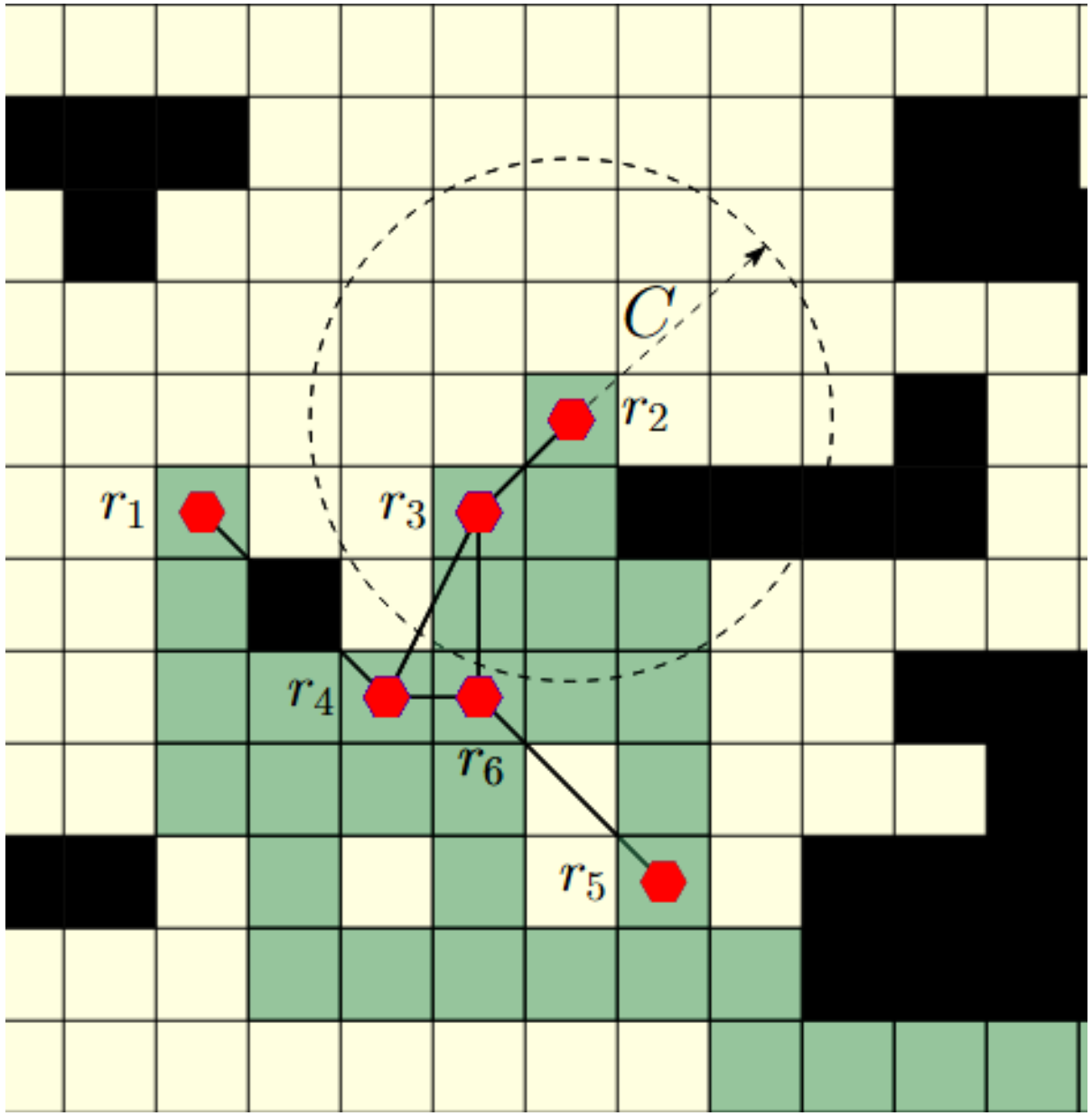
MARCC: Scalable Multi-robot Coverage with Connectivity Maintenance Via Deep Reinforcement Learning

ICTAI 2026

Jacomo Corrieri (UF), Iliya Kulbaka, Anirban Ghosh, Ayan Dutta (UNF)

Connected Multi-Robot Exploration

- **Multi-robot applications:** Environmental monitoring, precision agriculture, etc.
- **Computationally Difficult**
- **Core Trade-off:** Maximizing area coverage while preserving connectivity.

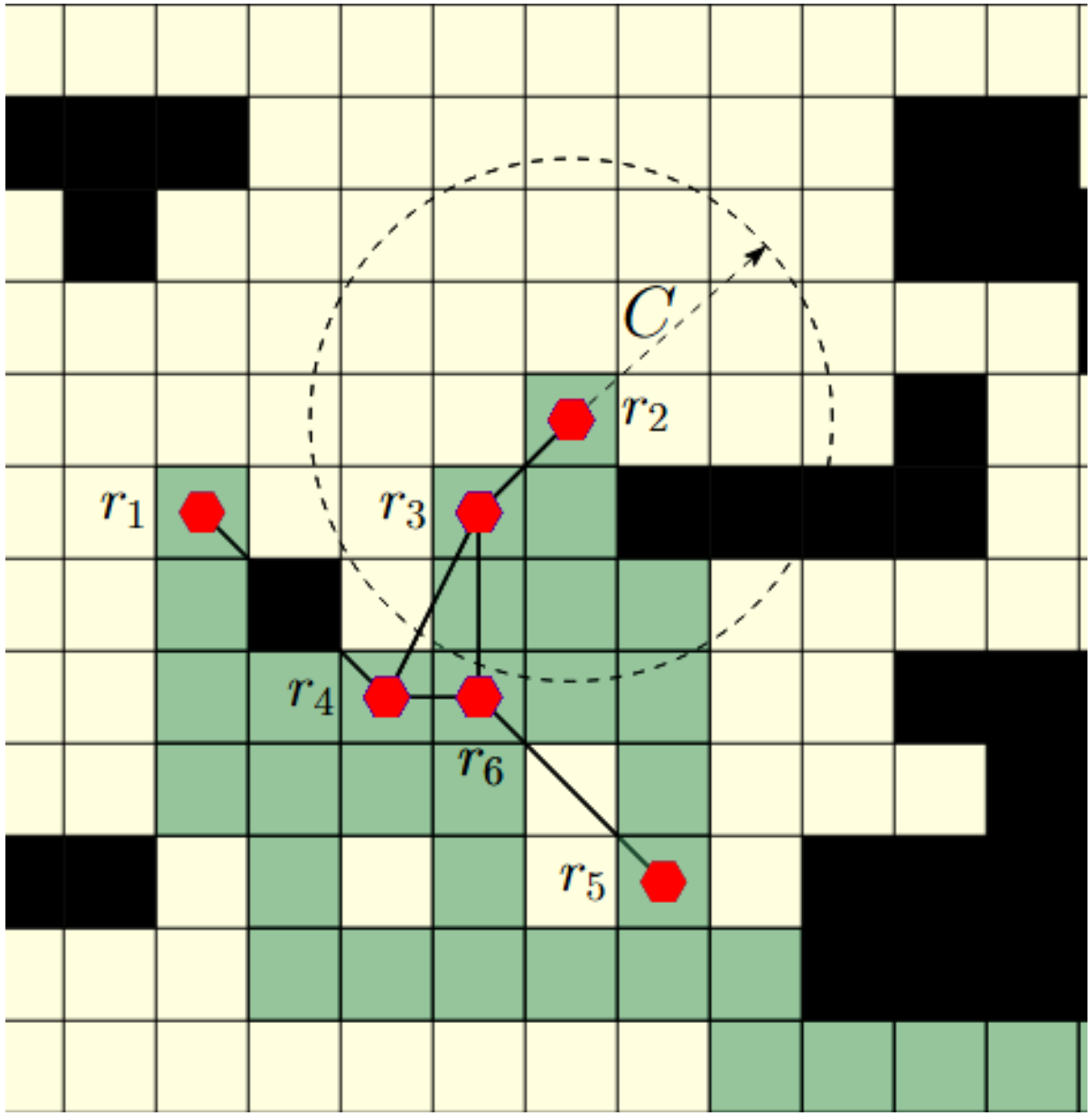


Original Work and Motivation

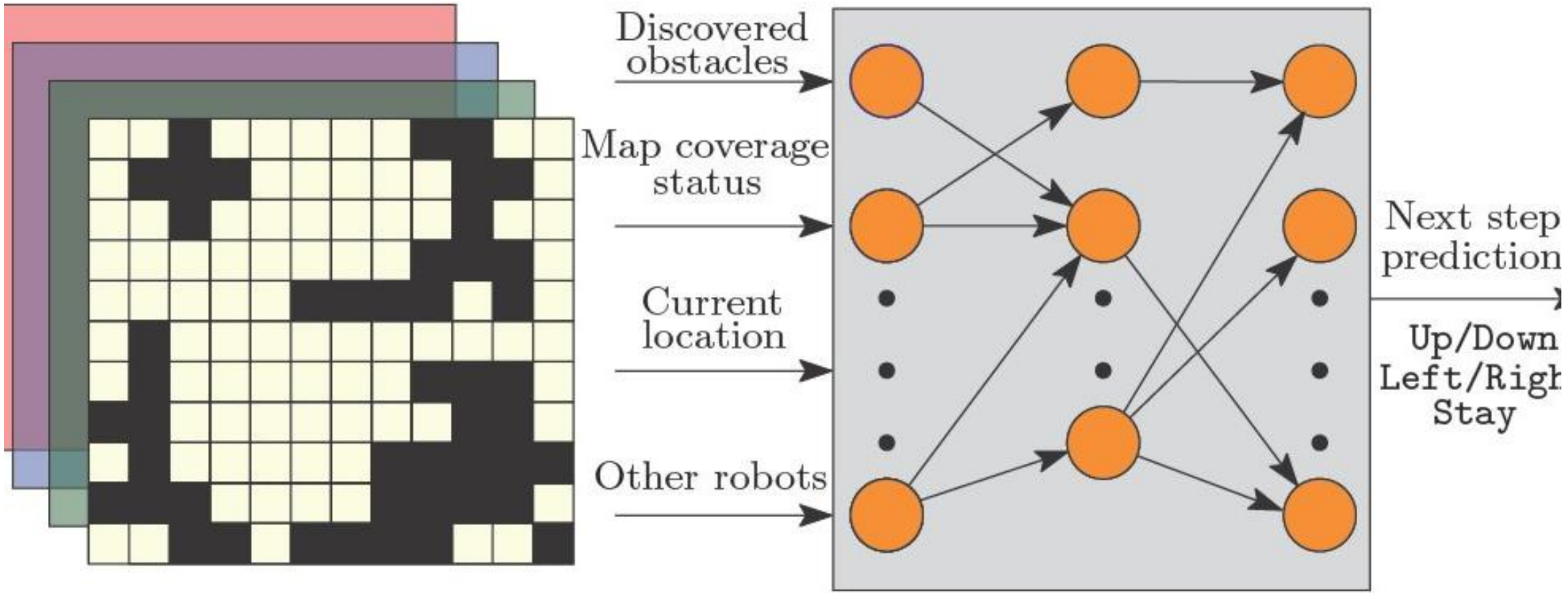
- Prior work [4] presents a path planning algorithm for multi-robot information collection, maintaining continuous connectivity for up to 10 robots.
- However, with the increase in n, the algorithm slows down considerably.
- **Main Motivation:** If robots can learn a policy offline, and those learned policies are executed on them, it might alleviate the complexity challenge and allow greater scalability beyond the original 10 robots.

MARCC

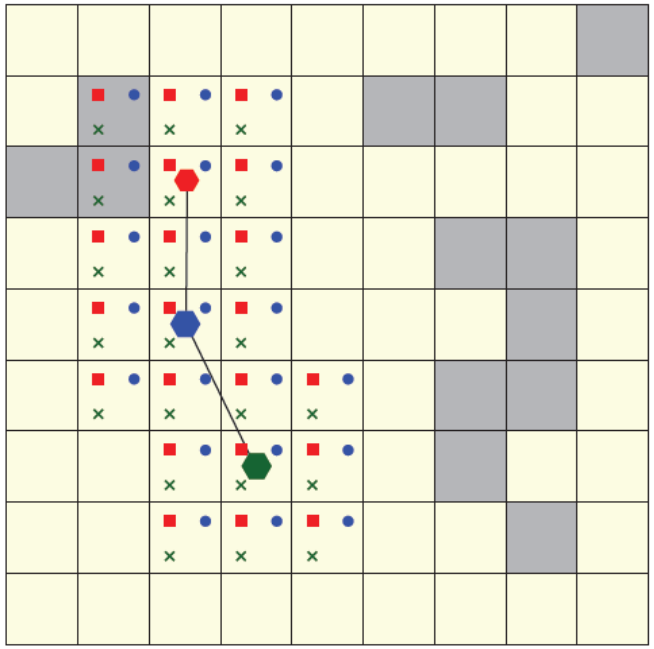
- Multi Agent Reinforcement Learning
- Trained and tested up to 50 robots
- Consistently saw 99% coverage and connectivity
- Generalized to unseen maps and team configurations



Local Observation Encoding

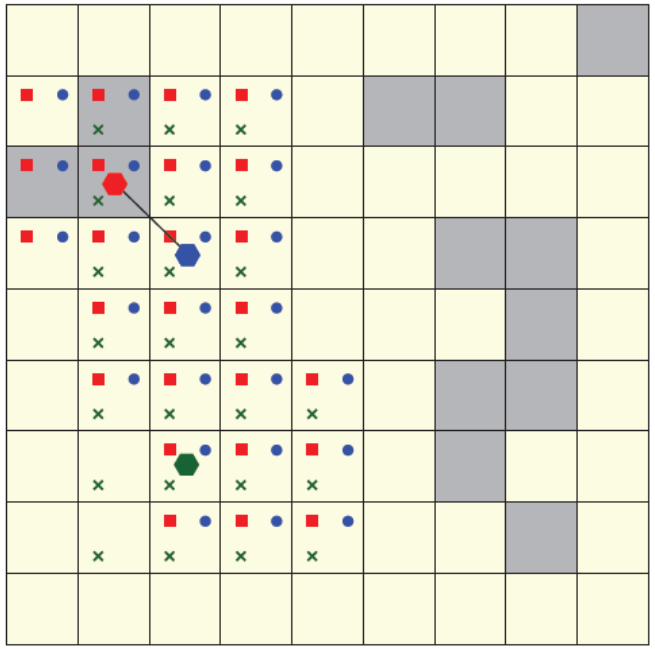


Observation Merging



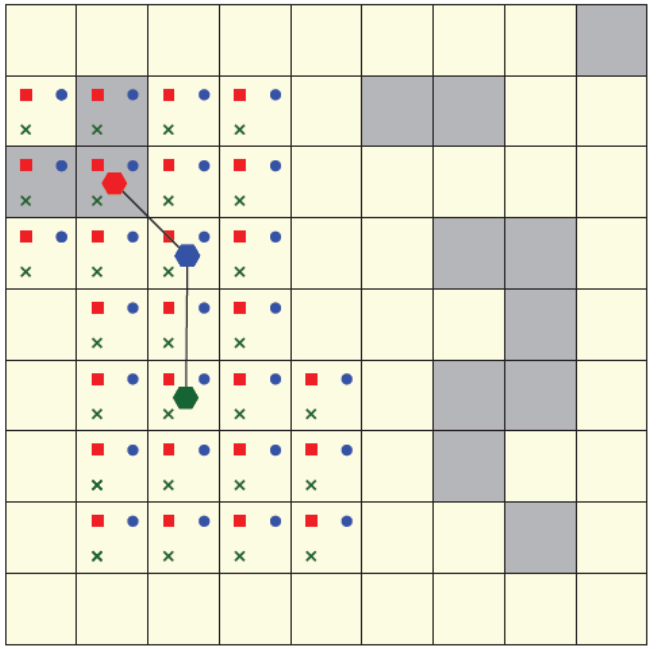
(a)

Agents are initially connected



(b)

Agents disconnect

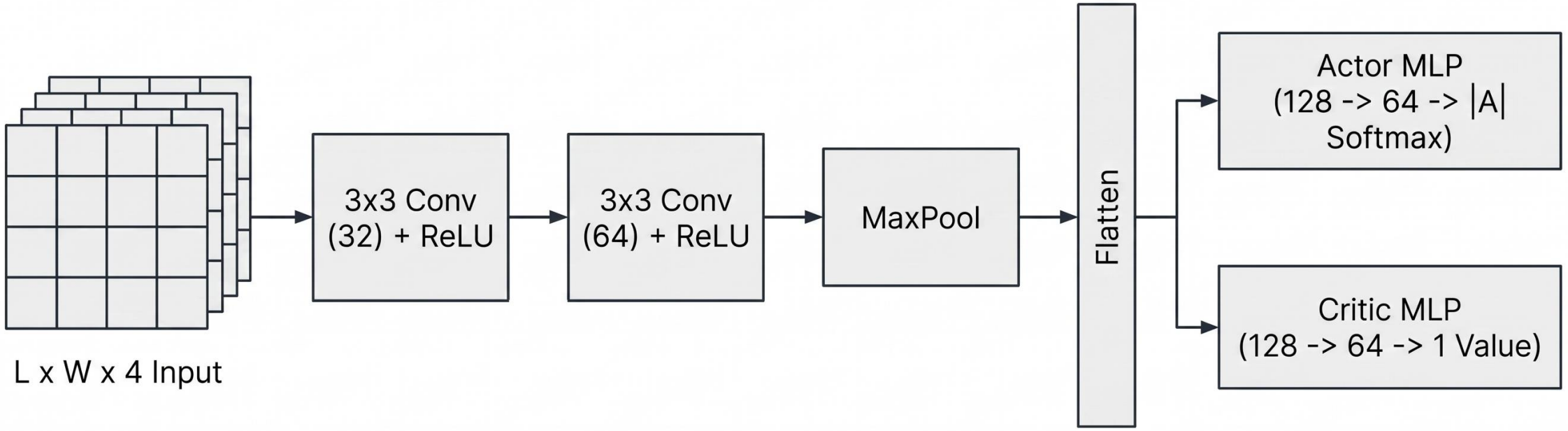


(c)

Agents reconnect, updating shared observations

Neural Network Architecture

- CNN-Backbone with shared parameters.
- Actor-Critic Proximal Policy Optimization
- ~1.2M actor-network parameters.



Reward Shaping

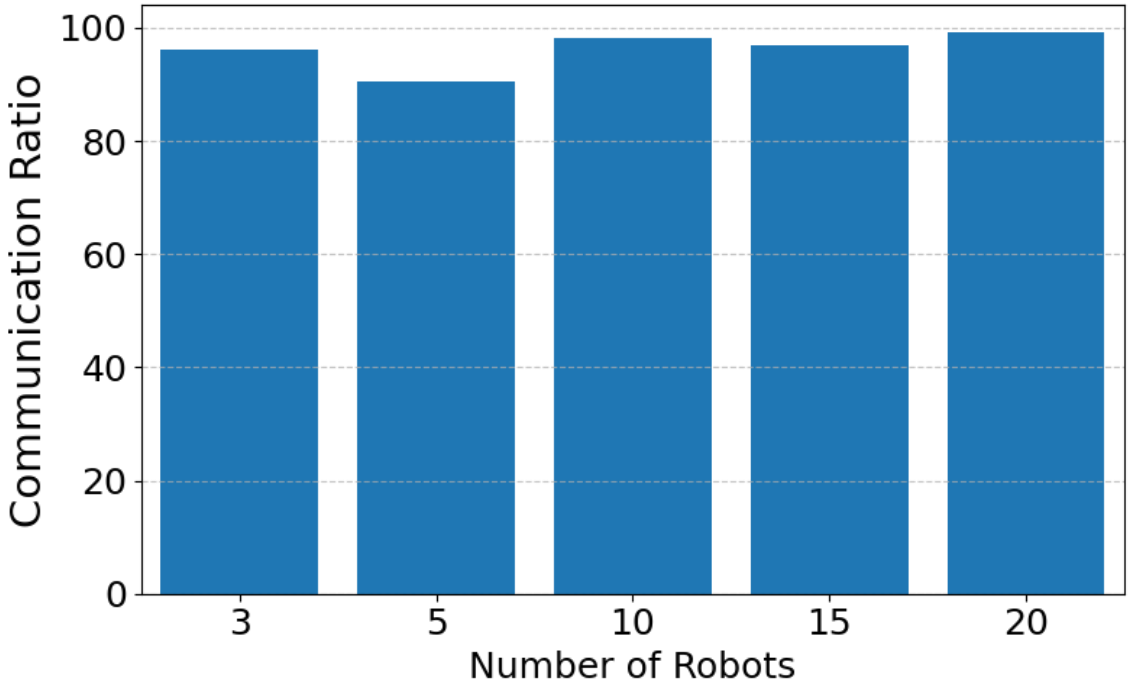
$$R^i(t) = R_{\text{cover}}(t) - R^i_{\text{obs}}(t) - R_{\text{disconn}}(t) - R_{\text{step}}$$

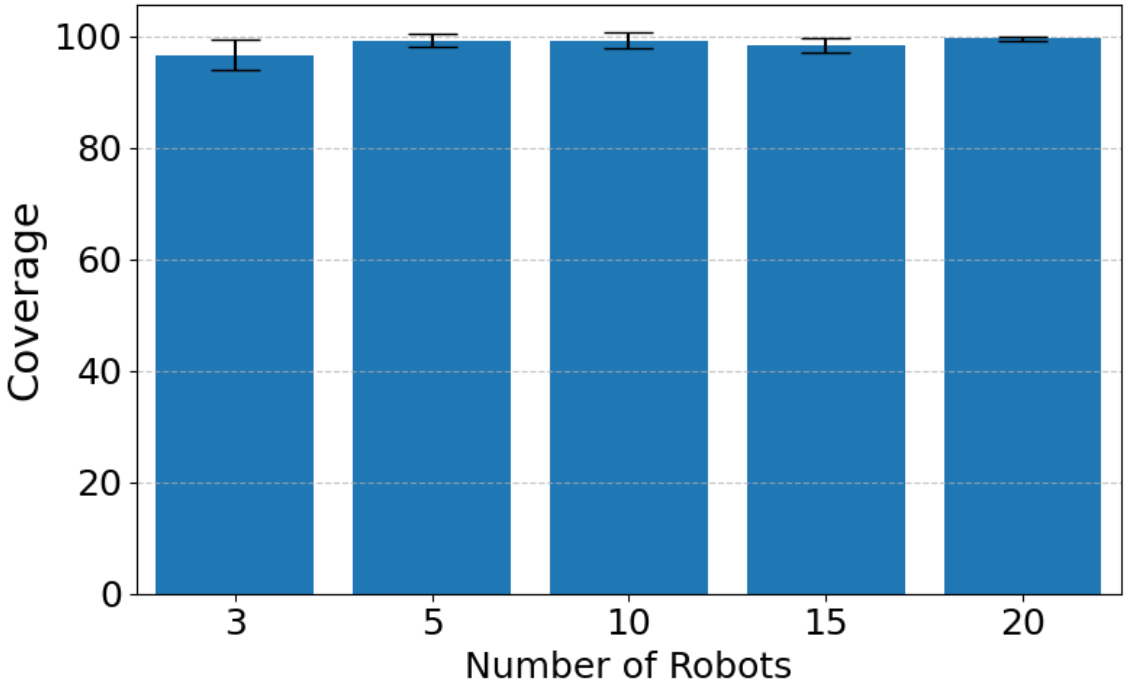
1. **Coverage Reward**: +100 * fractional change in coverage
2. **Obstacle penalty**: -0.1 for obstacle, robot, or boundary collision
3. **Disconnected graph**: -0.5 to all agents on disconnection
4. **Timestep penalty**: -0.01 for all agents each timestep
5. **Termination Bonus**: +50 on map completion

Evaluation Methodology

- **Training maps**: 50 randomized environments, 10,000 episodes, centralized critic.
- **Testing maps**: 50 randomized environments, unseen during training with 10 test runs per map (total of 500 episodes). No centralized critic.
- **Robot counts**: 3–20 robots during training, tested with the same number of robots, as well as different numbers than what the model was trained on.
- **Communication range**: 10 tiles
- **Sensing (FOV) range**: 8 tiles
- **Max episode length (steps)**: 1000

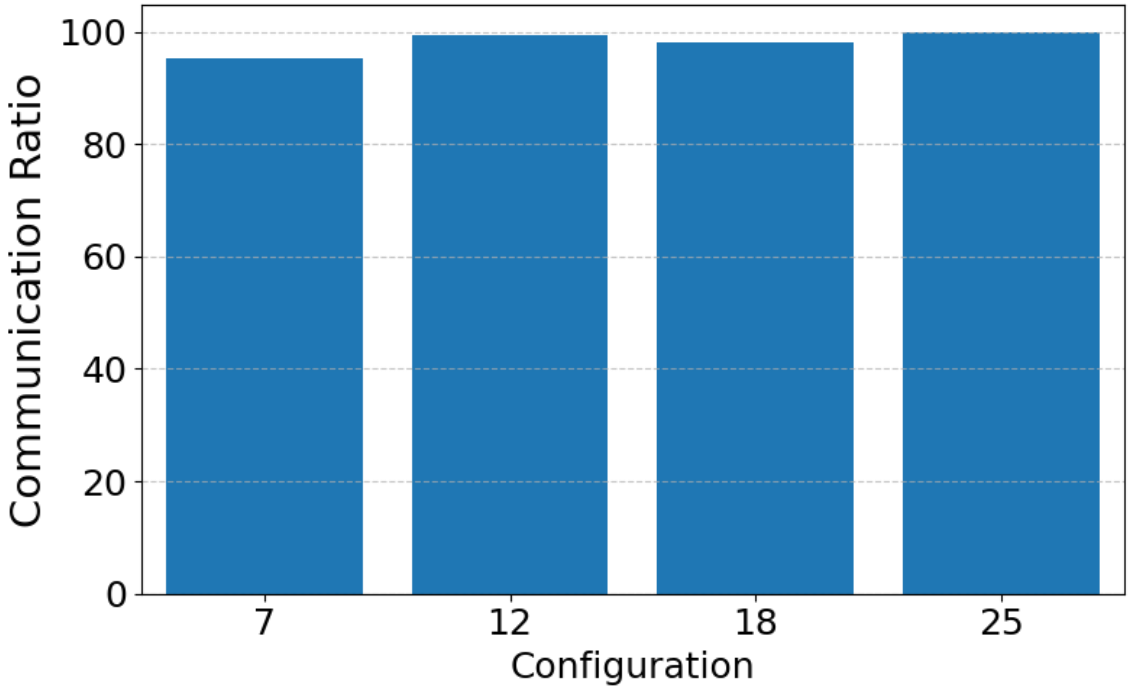
Same N in Testing: Results

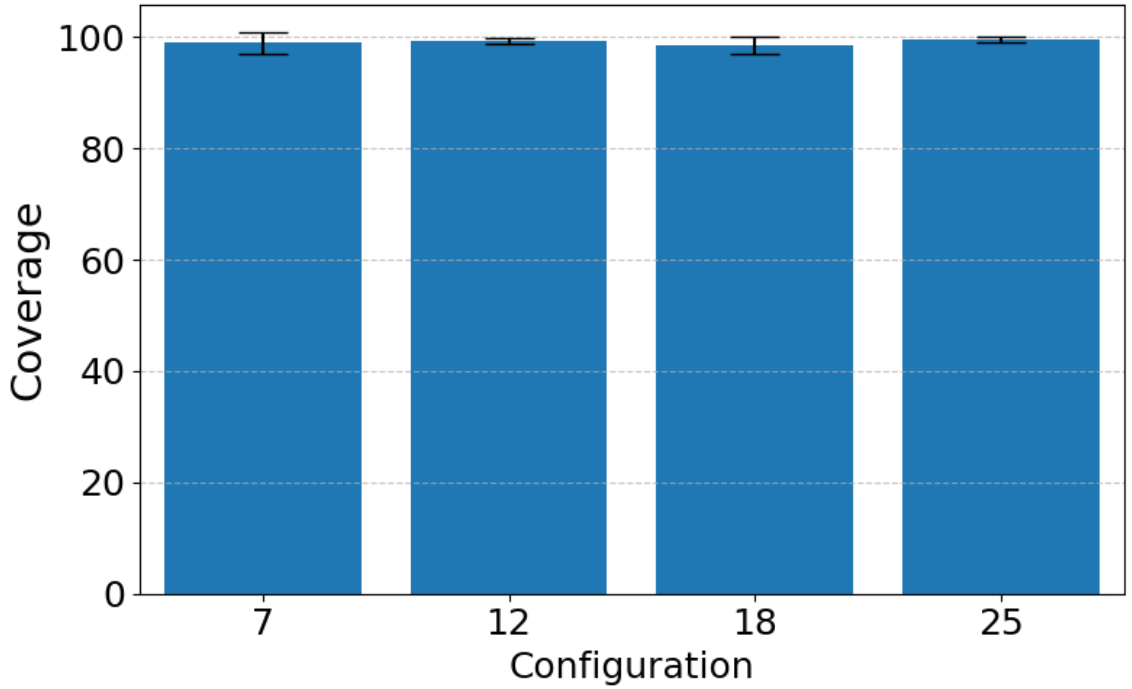




Different N in Testing: Results

**For trained n = {5, 10, 15, 20}*





Comparison Against Classical Baseline

- **Closest comparison: Rooker & Birk [5]**
- Across tested team sizes, MARCC reports ~15% higher average coverage while maintaining similar connectivity.
- At 20 robots: time to 90% coverage
- MARCC: 6.11 s and 67.32 steps | Rooker & Birk: 18.68 s and 678.68 steps
- **Approximately 3× faster to 90% coverage; roughly 10× fewer steps.**

Comparison Against Coverage-Only DQN

- **Coverage-only DDQN baseline [14]**
- **Purpose**: Validate that our coverage remains comparatively high despite communication reward shaping and quantify the decrease in connectivity without the penalty term in the reward.
- Compared against original team sizes (n = 3 through n = 20)
- Coverage: 95.83%, 98.74%, 99.44%, 99.26%, and 97.68%, respectively.
- Connectivity: 26.59%, 20.29%, 25.22%, 46.44%, and 59.22%, respectively.

Graph Connectivity Check

- We use Breadth-First-Search (BFS) vs Algebraic connectivity.
- Benchmarked performance on complete graphs up to size n = 3847 nodes ($m \approx 7.40 \times 10^6$ edges).
- BFS remains <0.1s, while computing $\lambda 2$ reaches 32.46s, an approximately 392x slowdown.

Ablations and Generalization

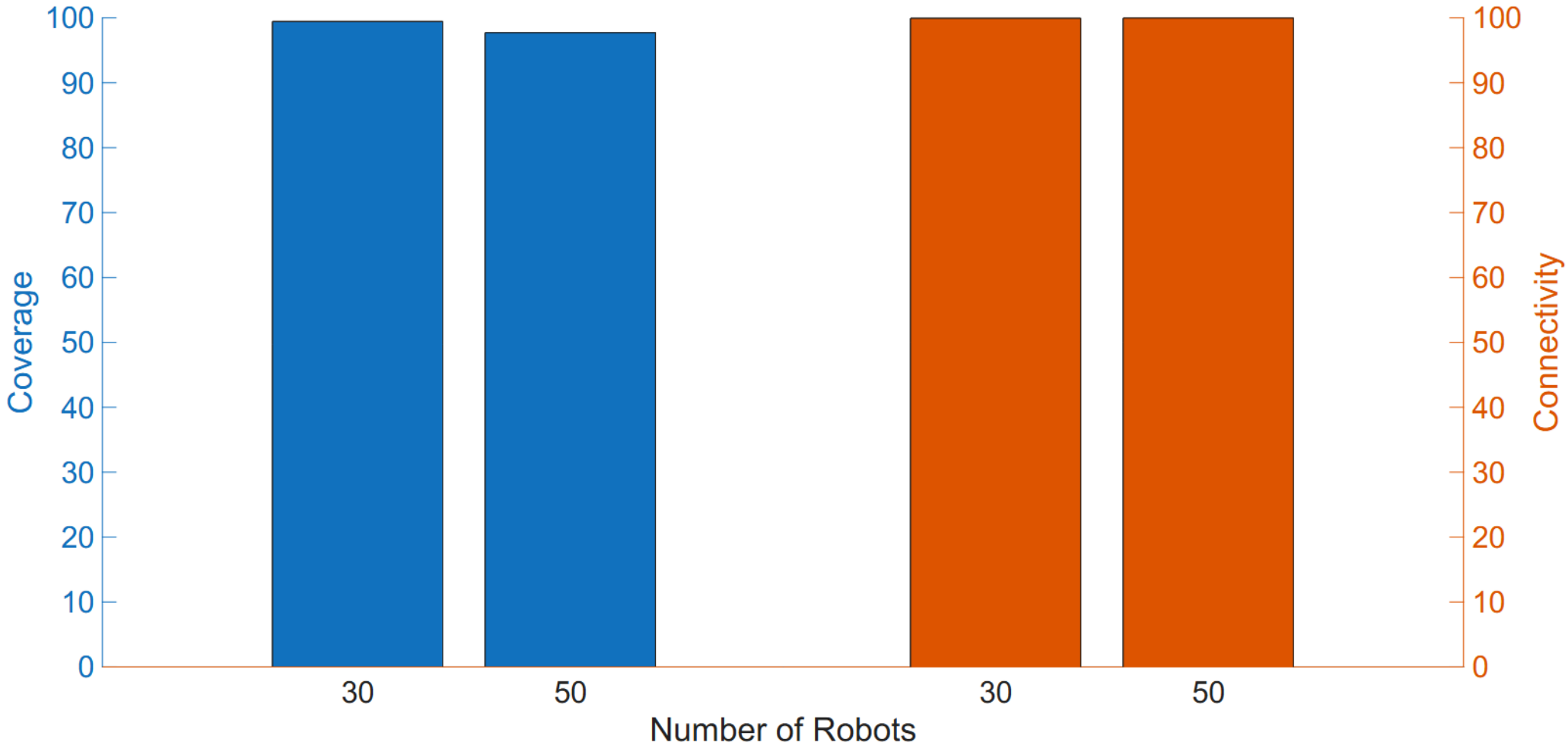
Remove Conv64 layer:

- Coverage: 96.67% → 92.39%
- Connectivity: 96.24% → 92.44%

Different FOV range in test:

Train $\mathcal{F}$	Test $\mathcal{F}$	Coverage Ratio	Comm. ratio	Makespan
8	10	99.62%	99.0%	874.31
8	8	99.63%	99.1%	873.51
8	5	99.61%	99.27%	885.28

From 20 Robots to 50





Thank you!

Contact:

jacomocorrieri@ufl.edu