VolleyBots: A Testbed for Multi-Drone Volleyball Game Combining Motion Control and Strategic Play

Zelai Xu[1*], Ruize Zhang[1*], Chao Yu[12+], Huining Yuan[1], Xiangmin Yi[1], Shilong Ji[1], Chuqi Wang[1], Wenhao Tang[1], Feng Gao[1], Wenbo Ding[1], Xinlei Chen[1], Yu Wang[1]

[1] Tsinghua University, [2] Beijing Zhongguancun Academy,

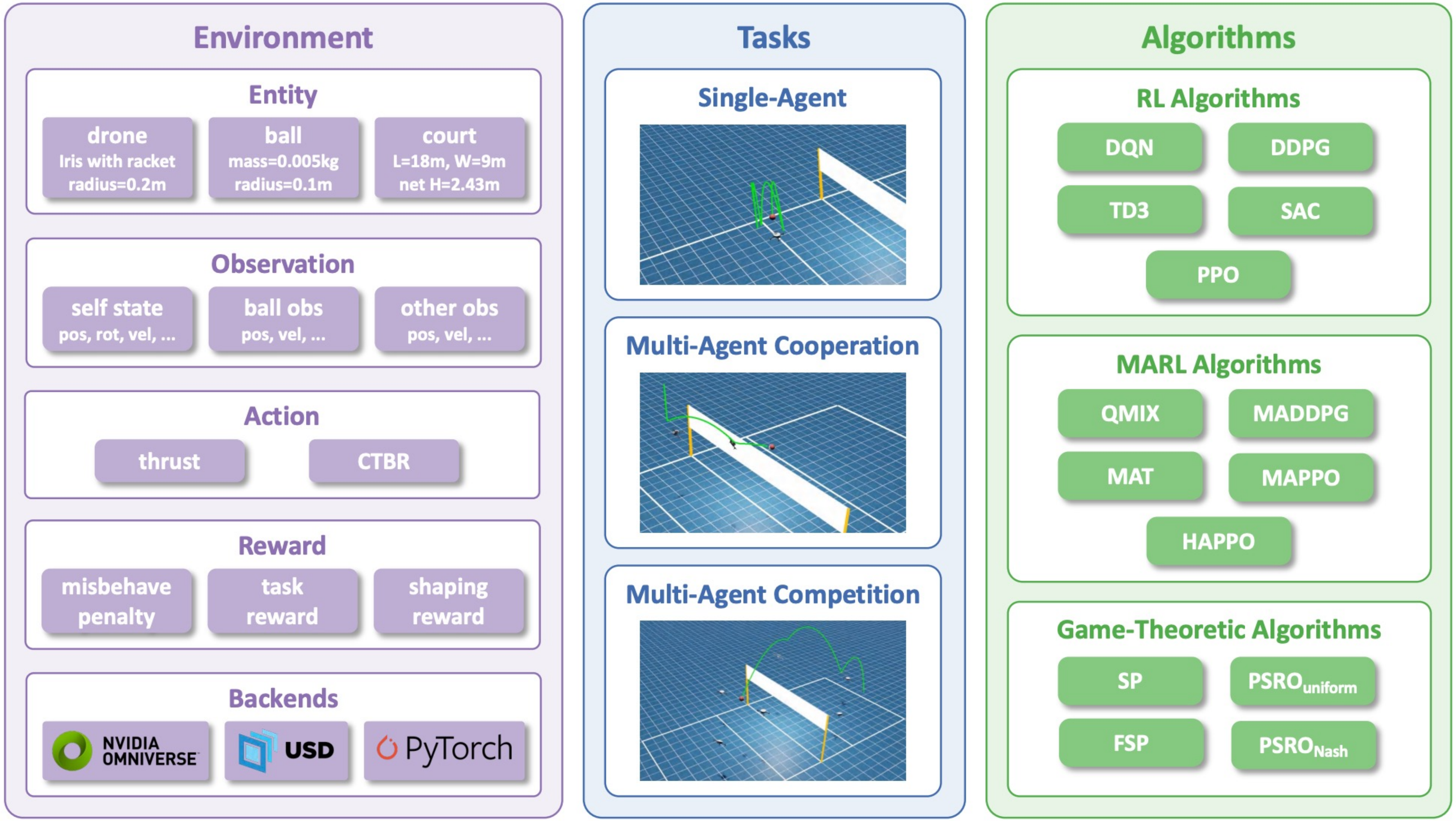
* Equal Contribution, + Corresponding Authors

Introduction

- In this paper, we present ***VolleyBots***, a novel robot sports testbed where multiple drones cooperate and compete in the sport of volleyball under physical dynamics.
- VolleyBots integrates *three features* within a unified platform:
  - Competitive and cooperative gameplay;
  - Turn-based interaction structure;
  - Agile 3D maneuvering.
- Comparison between VolleyBots and other robot sports platforms:

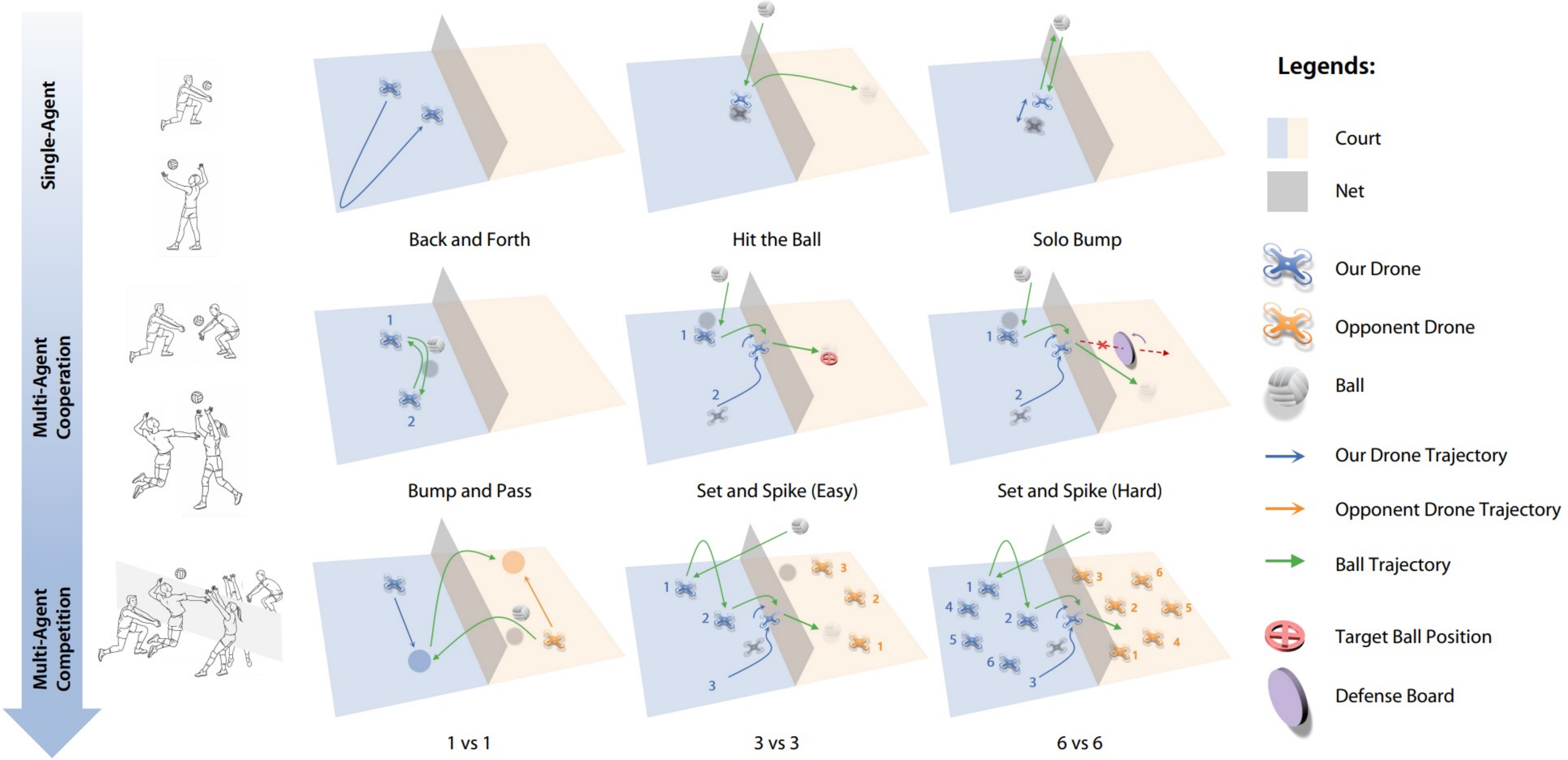
	Multi-Agent Task			Game Type	Entity	Hierarchical Policy	Open Source	Baseline Provided
	coop.	comp.	mixed					
Robot Table Tennis [5]	✗	✓	✗	turn-based	robotic arm	✓	✗	✗
Badminton Robot [13]	✗	✗	✗	turn-based	robotic arm	✗	✗	✗
Quadruped Soccer [2]	✗	✗	✗	simultaneous	quadruped	✓	✗	✗
MQE [4]	✓	✓	✓	simultaneous	quadruped	✓	✓	✓
Humanoid Football [1]	✗	✓	✓	simultaneous	humanoid	✓	✓	✗
SMPLOlympics [14]	✗	✓	✓	simu. & turn-based	humanoid	✗	✓	✓
Pursuit-Evasion [7]	✓	✗	✗	simultaneous	drone	✗	✓	✓
Drone-Racing [15]	✗	✗	✗	simultaneous	drone	✗	✗	✗
VolleyBots (Ours)	✓	✓	✓	turn-based	drone	✓	✓	✓

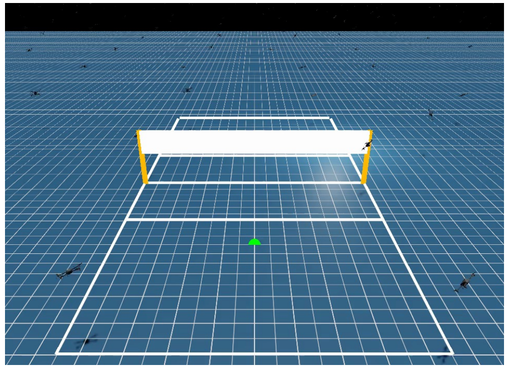
Overview

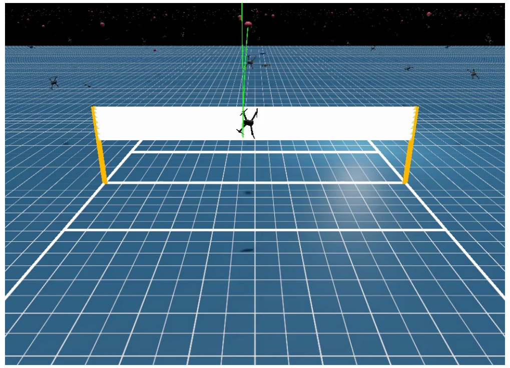


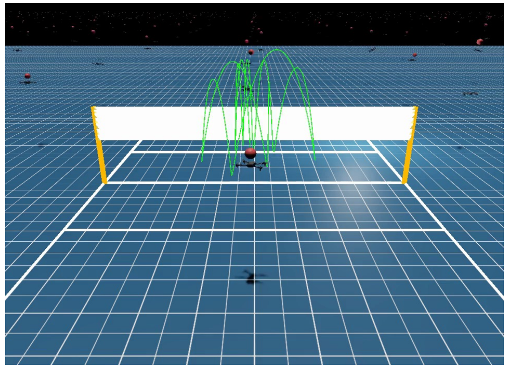
Environment
Entity
drone
Iris with racket
radius=0.2m
ball
mass=0.005kg
radius=0.1m
court
L=18m, W=9m
net H=2.43m
Observation
self state
pos, rot, vel, ...
ball obs
pos, vel, ...
other obs
pos, vel, ...
Action
thrust
CTBR
Reward
misbehave penalty
task reward
shaping reward
Backends
NVIDIA OMNIVERSE
USD
PyTorch
Tasks
Single-Agent
Multi-Agent Cooperation
Multi-Agent Competition
Algorithms
RL Algorithms
DQN
DDPG
TD3
SAC
PPO
MARL Algorithms
QMIX
MADDPG
MAT
MAPPO
HAPPO
Game-Theoretic Algorithms
SP
$PSRO_{uniform}$
FSP
$PSRO_{Nash}$

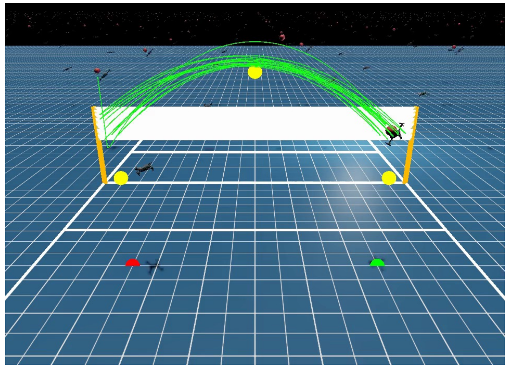
Tasks

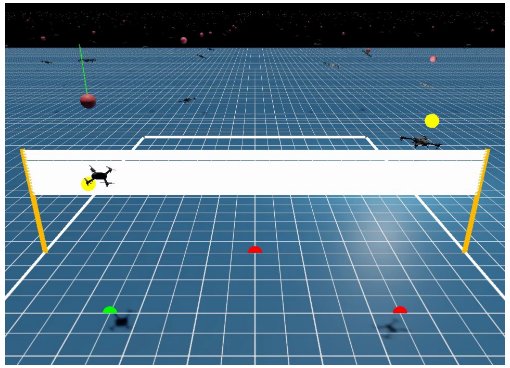

Single-Agent
Multi-Agent Cooperation
Multi-Agent Competition
Back and Forth
Hit the Ball
Solo Bump
Bump and Pass
Set and Spike (Easy)
Set and Spike (Hard)
1 vs 1
3 vs 3
6 vs 6
Legends:
Court
Net
Our Drone
Opponent Drone
Ball
Our Drone Trajectory
Opponent Drone Trajectory
Ball Trajectory
Target Ball Position
Defense Board

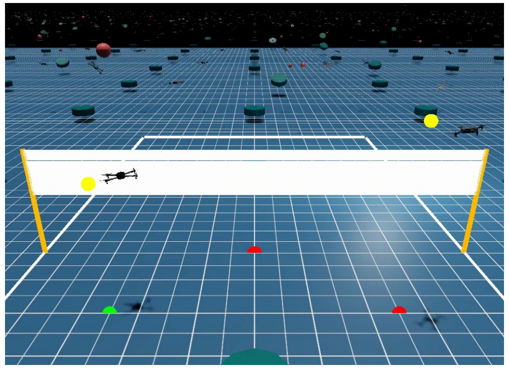

Back and Forth

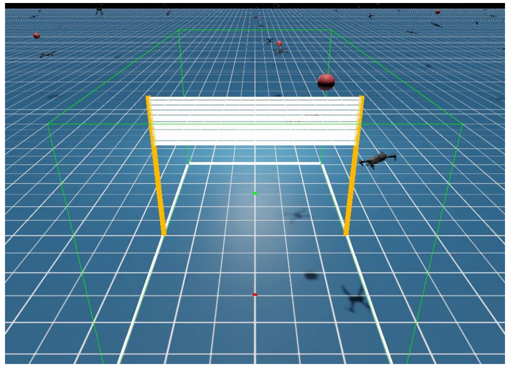

Hit the Ball

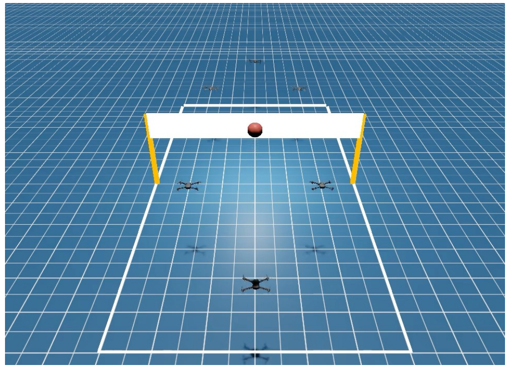

Solo Bump

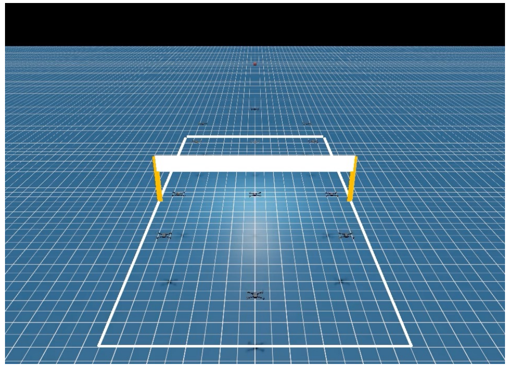

Bump and Pass


Set and Spike (Easy)


Set and Spike (Hard)


1 vs 1


3 vs 3


6 vs 6

Benchmark results

Single-agent tasks

	Back and Forth		*Hit the Ball*		*Solo Bump*	
	CTBR	PRT	CTBR	PRT	CTBR	PRT
DQN	0.00 ± 0.00	0.00 ± 0.00	0.39 ± 0.02	1.88 ± 0.34	0.00 ± 0.00	0.00 ± 0.00
DDPG	1.14 ± 0.34	0.83 ± 0.23	2.87 ± 0.55	3.98 ± 1.08	0.44 ± 0.34	0.67 ± 0.32
TD3	1.12 ± 0.68	0.99 ± 0.01	3.00 ± 0.52	3.91 ± 0.35	3.68 ± 1.43	5.29 ± 1.28
SAC	0.90 ± 0.12	0.83 ± 0.25	3.76 ± 1.46	3.87 ± 2.34	0.54 ± 0.27	1.36 ± 0.60
PPO	$\mathbf{9.25 \pm 0.31}$	$\mathbf{10.04 \pm 0.20}$	$\mathbf{10.48 \pm 0.08}$	$\mathbf{11.40 \pm 0.06}$	$\mathbf{8.58 \pm 0.79}$	$\mathbf{10.83 \pm 1.24}$

- Comparing different action spaces, the results indicate that PRT slightly outperforms CTBR in most tasks.
- With a single set of hyperparameters, on-policy RL methods maintains consistently strong performance across multiple tasks, demonstrating superior robustness compared to off-policy methods.

Benchmark results

Multi-agent cooperative tasks

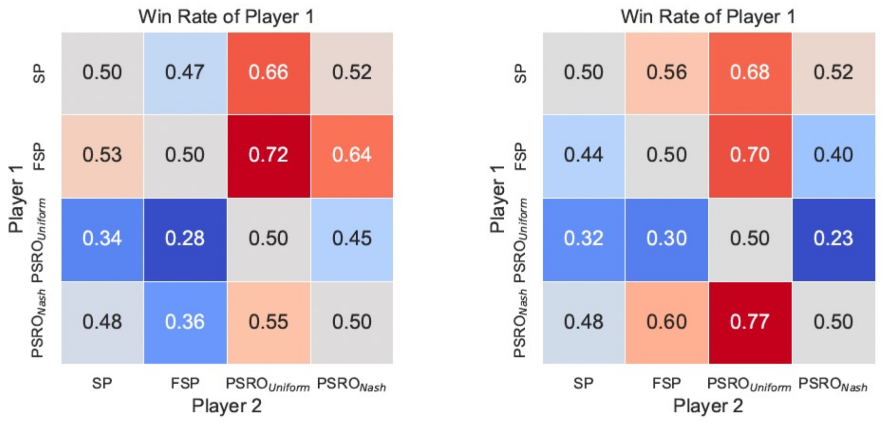
	Bump and Pass		*Set and Spike (Easy)*		*Set and Spike (Hard)*	
	w.o. shaping	w. shaping	w.o. shaping	w. shaping	w.o. shaping	w. shaping
QMIX	0.09 ± 0.01	0.09 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00
MADDPG	0.79 ± 0.15	0.84 ± 0.09	0.22 ± 0.02	0.23 ± 0.01	0.22 ± 0.02	0.22 ± 0.02
MAPPO	$\mathbf{11.32 \pm 0.91}$	$\mathbf{13.71 \pm 0.58}$	$\mathbf{0.25 \pm 0.00}$	$\mathbf{0.99 \pm 0.00}$	$\mathbf{0.25 \pm 0.00}$	0.75 ± 0.01
HAPPO	7.95 ± 3.67	12.14 ± 0.83	$\mathbf{0.25 \pm 0.00}$	0.98 ± 0.00	$\mathbf{0.25 \pm 0.00}$	0.79 ± 0.10
MAT	7.39 ± 6.00	13.11 ± 0.43	$\mathbf{0.25 \pm 0.00}$	0.89 ± 0.13	$\mathbf{0.25 \pm 0.00}$	$\mathbf{0.80 \pm 0.11}$

- Using reward shaping leads to better performance.
- With a single set of hyperparameters, on-policy MARL methods also outperform off-policy MARL methods.

Benchmark results

Multi-agent competitive tasks

	1 vs 1			*3 vs 3*		
	Exploitability ↓	Win Rate ↑	Elo ↑	Exploitability ↓	Win Rate ↑	Elo ↑
SP	48.63	0.55	1072	**25.76**	0.59	1077
FSP	30.41	**0.63**	927	38.86	0.52	906
$PSRO_{Uniform}$	18.51	0.35	854	49.48	0.28	750
$PSRO_{Nash}$	**10.74**	0.47	**1147**	35.83	**0.61**	**1268**

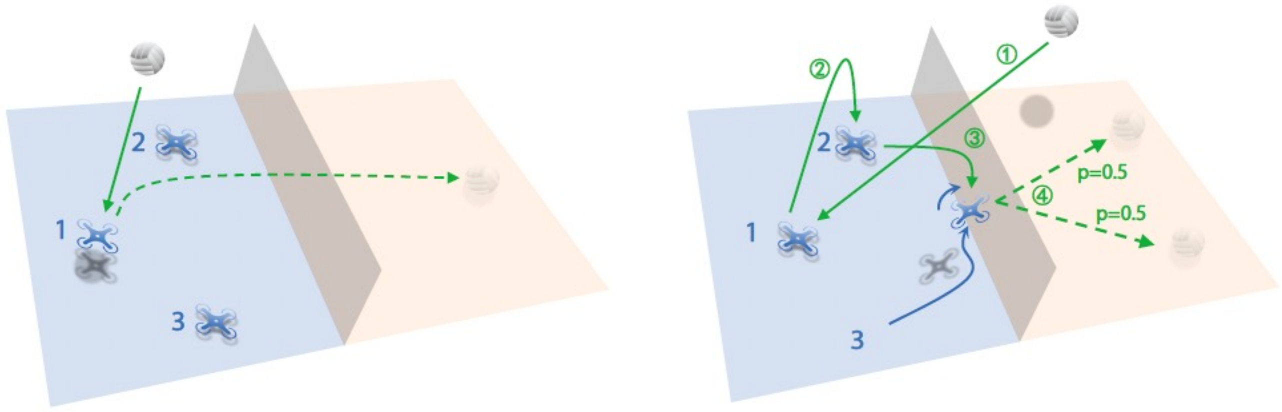


(a) *1 vs 1* (b) *3 vs 3*

- In the most difficult 6 vs 6 task, none of the methods converges to an effective strategy.
- Existing approaches also struggle in 1 vs 1 and 3 vs 3 tasks that combine motion control and strategic play.

Benchmark results

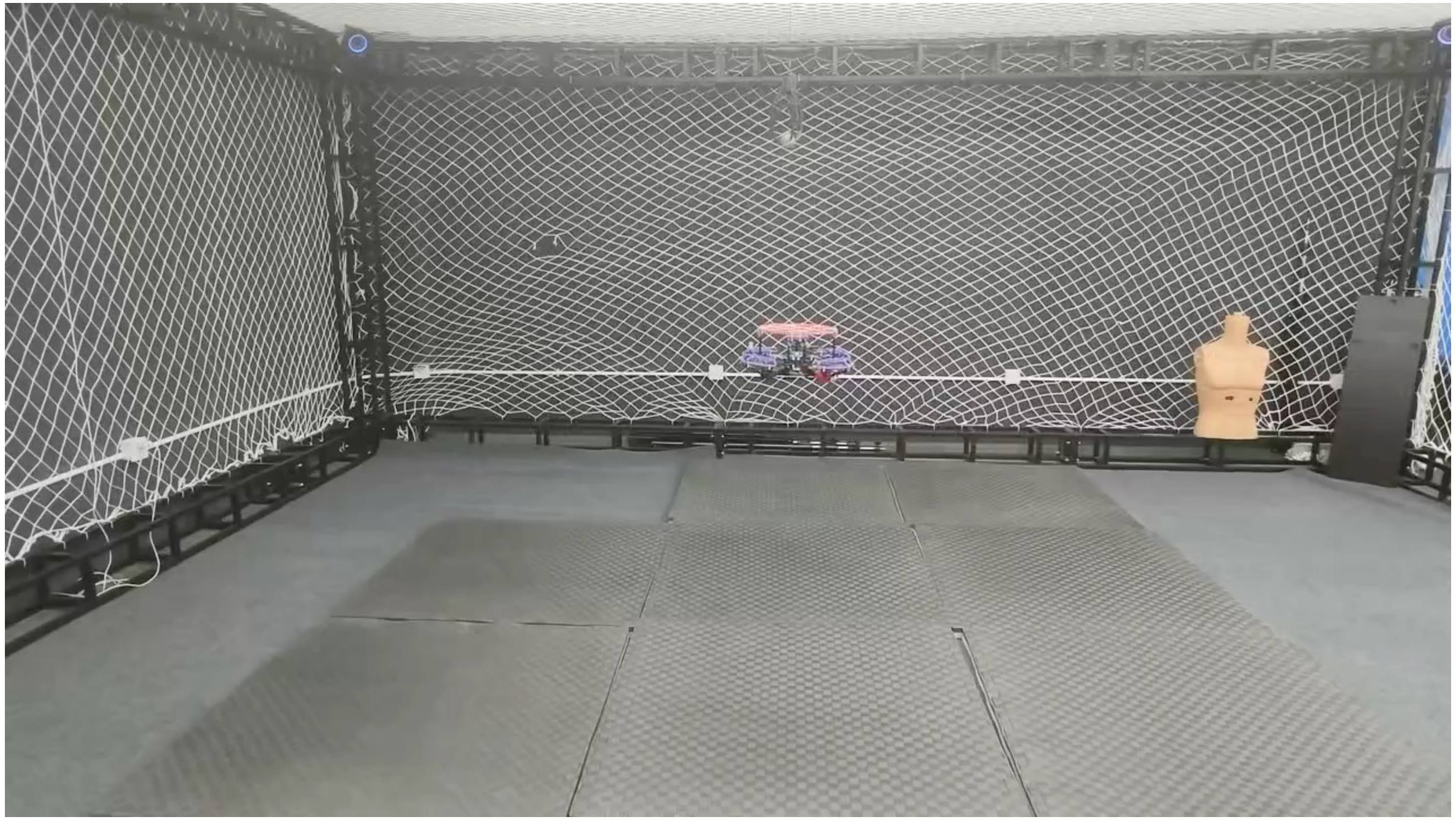
Multi-agent competitive tasks



(a) *Serve* (b) *Attack*

- Additionally, we design a rule-based hierarchical policy which achieves 69.5% win rate against the strongest baseline in the 3 vs 3 task, demonstrating its potential for tackling the complex interplay between low-level control and high-level strategy.

Sim-to-Real



We use the Solo Bump task as a demonstration of the policy's ability to zero-shot transfer to the real world. Experiment results show that the drone successfully performs bump tasks multiple times.

Highlights

- We introduce **VolleyBots**, a novel robot sports environment centered on drone volleyball, featuring mixed competitive and cooperative game dynamics, turn-based interactions, and agile 3D maneuvering.
- We release a curriculum of tasks, ranging from **single-drone** drills to multi-drone **cooperative** plays and **competitive** matchups, and baseline evaluations of representative (MA)RL and game-theoretic algorithms.
- We design a rule-based hierarchical policy that achieves a 69.5% win rate against the strongest baseline in the 3 vs 3 task, offering a promising solution for tackling the complex interplay between low-level control and high-level strategy.

Thanks for listening!

VolleyBots: A Testbed for Multi-Drone Volleyball Game Combining Motion Control and Strategic Play

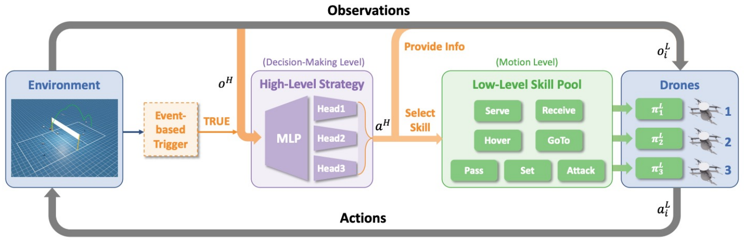
Paper

Code

Website

Series of Work Based on VolleyBots:

Mastering Multi-Drone Volleyball through Hierarchical Co-Self-Play Reinforcement Learning **[Accepted at CoRL 2025]**



JuggleRL: Mastering Ball Juggling with a Quadrotor via Deep Reinforcement Learning **[Submitted to ICRA 2026]**