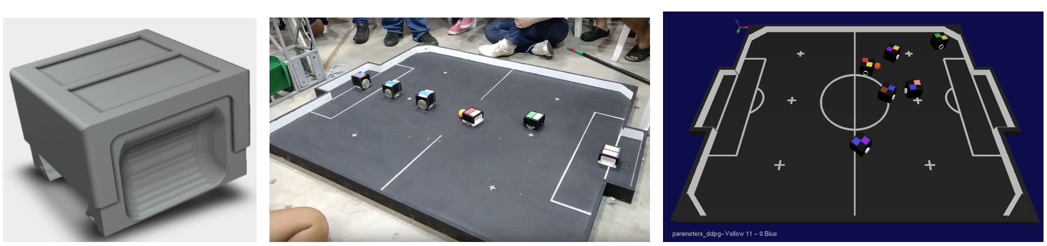
Learning to Play Soccer by Reinforcement and Applying Sim-to-Real to Compete in the Real World

Anonymous Author(s)
Affiliation
Address
email

1 Introduction

This work presents an application of *Reinforcement Learning* (RL) for the complete control of real soccer robots of the *IEEE Very Small Size Soccer* (VSSS) [1], a traditional league in the *Latin American Robotics Competition* (LARC). In the VSSS league, two teams of three small robots play against each other. We propose a simulated environment in which continuous or discrete control policies can be trained, and a Sim-to-Real method to allow using the obtained policies to control a robot in the real world. The results show that the learned policies display a broad repertoire of behaviors which are difficult to specify by hand. This approach, called VSSS-RL, was able to beat the human-designed policy for the striker of the team ranked 3rd place in the 2018 LARC, in 1-vs-1 matches.



(a) (b) (c)

Figure 1: (a) 3D model of a VSSS robot; (b) Real-world game setup; and (c) Simulation [2].

2 Research Problem

The VSSS robots are usually programmed to behave adequately in every situation identified by the programmers, employing path planning, collision avoidance, and PID control methods [7]. However, it is extremely hard to foreseen and tackle every possible situation in a dynamic game such as soccer. Therefore, it is clear the need for data-oriented approaches such as RL.

However, several barriers exist for applying RL successfully in the real world [5], as the large amounts of interactions required by the agents to achieve adequate performance are impractical due to degradation of hardware, energy consumption and time required. Thus, the research problem considered in this work is the application of the Sim-to-Real approach, in which the agents are trained in simulation and policies learned are transferred to the real robots.

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

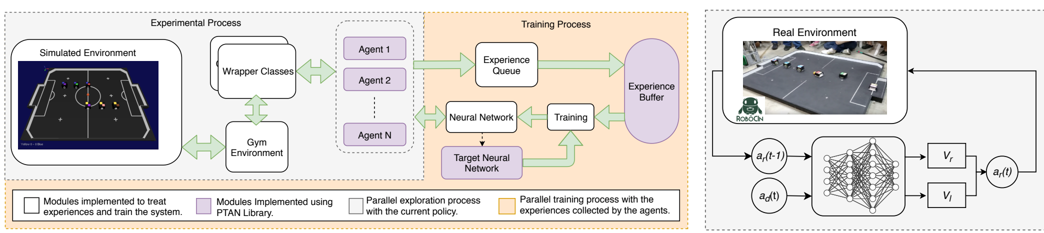
3 Motivation

Deep RL is a suitable approach for learning control and complex behaviors by interacting with the environment since it requires only the specification of a reward function that expresses the desired goals. In the literature of robot soccer, RL has been applied for learning specific behaviors, such as kicking [10] and scoring penalty goals [6].

Recently, two RL soccer simulation environments have been proposed: MuJoCo Soccer [12] and Google Research Football [8]. However, they are not suitable for the study of Sim-to-Real, because they either do not consider important physical and dynamical aspects or represent a very complex scenario that is not achievable by current robotics technology. Therefore, the need for such an adequate environment, allowing the study of the combination of RL with Sim-to-Real in dynamic, multi-agent, competitive, and cooperative situations, is the main motivation behind this work.

4 Technical Contribution

We propose a simulated environment called VSSS-RL[1], which supports continuous or discrete control policies. It includes a customized version of the VSS SDK simulator [2] and builds a set of wrapper modules to be compatible with the OpenAI Gym standards [4]. It consists of two main independent processes: the experimental, and the training process. In the first, an OpenAI Gym environment parser was developed, and wrapper classes were implemented to communicate with the agents. In the latter, the collected experiences are stored in an experience buffer that is used to update the policies, as illustrated in Fig. 2(a).



(a) (b)

Figure 2: VSSS-RL: (a) Environment Architecture for training high-level control policies. (b) Low-level control training processes to enable Sim-to-Real transfer.

We also proposed a Sim-to-Real method to transfer the obtained policies to a robot in the real world. It is a Domain Adaptation method [3], consisting of a Feed-Forward Neural Network which learns to map the desired high-level actions $a_d(t) = \{v, \omega\}$ (linear and angular speeds) to low-level control commands for the wheel speeds (V_R and V_L) (Fig. 2(b)).

4.1 Experimental Results

The results, submitted to ICRA2020, show that the two baseline RL methods evaluated, *Deep Deterministic Policy Gradient* (DDPG) [9] and *Deep Q Network* (DQN) [13], were able to learn suitable policies in simulation when applying reward shaping [11]. The learned polices display rich and complex behaviors[2] extremely difficult to specify by hand as well as to identify the correct moments when they should be applied. Moreover, the proposed Sim-to-Real method employed allowed us to achieve similar results in the real world in terms of average steps to score a goal (547.2 ± 233.6 in simulation and 456.8 ± 147.2 in the real world).

Finally, the complete approach was evaluated in 1-vs-1 matches against the striker of RoboCIn VSSS team, 3rd place on the LARC 2018. The final scores of the matches were 19 for VSSS-RL and 13 for RoboCIn in the first game, and 22 for VSSS-RL approach and 17 for RoboCIn in the second. These wins highlight the capabilities of the proposed approach.

[1] Source code will be available soon at: `https://github.com/robocin/vss-environment`

[2] See the video available at: `https://youtu.be/a9dTMtanh-U`

References

[1] Very Small Size Soccer Rules. http://www.cbrobotica.org/wp-content/uploads/2014/03/VerySmall2008_en.pdf, 2019. [Online; accessed 26-May-2019].

[2] VSS SDK. https://vss-sdk.github.io/book/general.html, 2019. [Online; accessed 5-June-2019].

[3] M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, et al. Learning dexterous in-hand manipulation. *arXiv preprint arXiv:1808.00177*, 2018.

[4] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. Openai gym, 2016.

[5] G. Dulac-Arnold, D. Mankowitz, and T. Hester. Challenges of real-world reinforcement learning. *arXiv preprint arXiv:1904.12901*, 2019.

[6] T. Hester, M. Quinlan, and P. Stone. Generalized model learning for reinforcement learning on a humanoid robot. In *2010 IEEE International Conference on Robotics and Automation*, pages 2369–2374. IEEE, 2010.

[7] J.-H. Kim, D.-H. Kim, Y.-J. Kim, and K. T. Seow. *Soccer robotics*, volume 11. Springer Science & Business Media, 2004.

[8] K. Kurach, A. Raichuk, P. Stańczyk, M. Zajac, O. Bachem, L. Espeholt, C. Riquelme, D. Vincent, M. Michalski, O. Bousquet, et al. Google research football: A novel reinforcement learning environment. *arXiv preprint arXiv:1907.11180*, 2019.

[9] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[10] M. Riedmiller and T. Gabel. On experiences in a complex and competitive gaming domain: Reinforcement learning meets robocup. In *2007 IEEE Symposium on Computational Intelligence and Games*, pages 17–23. IEEE, 2007.

[11] R. S. Sutton, A. G. Barto, et al. *Introduction to reinforcement learning*, volume 2. MIT press Cambridge, 1998.

[12] E. Todorov, T. Erez, and Y. Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012.

[13] M. Volodymyr, K. Kavukcuoglu, D. Silver, A. Graves, and I. Antonoglou. Playing atari with deep reinforcement learning. In *NIPS Deep Learning Workshop*, 2013.