Curriculum Learning in Google Research Football

Emir Sahin
McGill University
Montréal, Québec
emir.sahin@mail.mcgill.ca

Dmytro Fedorenko
McGill University
Montréal, Québec
dmytro.fedorenko@mail.mcgill.ca

Angadh Verma
McGill University
Montréal, Québec
angadh.verma@mail.mcgill.ca

Abstract

Reinforcement Learning (RL) has seen substantial progress in the past decade, particularly through the incorporation of deep learning methods. However, significant challenges persist, such as computational complexity and sample inefficiency, especially in complex environments. This project explores the use of curriculum learning within the Google Research Football environment, a realistic and dynamic simulation of football matches, to address these challenges. By progressively structuring tasks of increasing complexity, we aim to enhance learning efficiency, stability, and performance of RL agents. Our findings indicate that curriculum-based strategies can significantly accelerate agent training and improve overall performance, highlighting a promising direction for future RL research.

1 Introduction

Reinforcement learning (RL) has emerged as a powerful approach for training agents to perform tasks through interaction with their environment. Recent advances, particularly in deep reinforcement learning, have enabled remarkable achievements in diverse fields such as gaming, robotics, and autonomous driving. Despite these successes, RL still faces considerable barriers, notably the issues of high computational demands and sample inefficiency. These issues become especially pronounced in environments that simulate real-world complexities.

The Google Research Football (GRF) environment presents an ideal testbed to address these challenges due to its realism, complexity, and relevance to strategic and coordinated decision-making [2]. GRF is modeled after popular video games and simulates a football match, capturing intricate aspects of the sport, including team coordination, strategic gameplay, real-time decision-making, and tactical execution (Figure 1). Given its complexity, training RL agents in GRF often requires substantial computational resources and extensive training durations.

Curriculum learning, which structures the training process by progressively increasing task difficulty, has shown promise in addressing inefficiencies in various RL scenarios [1, 5]. This project investigates curriculum learning within GRF to systematically evaluate its impact on training efficiency and overall agent performance. By doing so, we aim to demonstrate a robust method to mitigate computational and sample inefficiency barriers, thereby making RL techniques more applicable to complex scenarios such as GRF.

The subsequent sections of this paper review related literature, describe the methodology and experimental setup, present and analyze results, and discuss future directions for research.

Preprint. Under review.

Figure 1: GRF supports all the major football rules such as kickoffs (top left), goals (top right), fouls, cards (bottom left), corner and penalty kicks (bottom right), and offside [2].

2 Related Work

Reinforcement Learning (RL) has evolved as a robust paradigm for training agents through interactions with environments by maximizing cumulative rewards. The foundational framework introduced by Sutton and Barto provides a comprehensive overview of standard RL algorithms [9]. Deep reinforcement learning, which integrates deep neural networks into traditional RL frameworks, has substantially extended the applicability of RL methods. Further advancements, including policy gradient methods such as Proximal Policy Optimization (PPO) and Actor-Critic architectures, have provided stable and efficient solutions suitable for continuous and complex action spaces [8, 4].

Curriculum learning in supervised learning contexts has been adopted within RL to train on progressively harder learning tasks, improving both training efficiency and final performance [1]. The use of curriculum learning in RL has also been explored, highlighting its capacity to accelerate training by systematically adjusting task complexity [6, 7].

The Google Research Football environment has become a popular testbed for RL algorithms due to its realistic simulation of football matches [2]. Modeled after popular football video games, the environment provides a physics-based 3D football simulation where agents have to control their players, learn to pass between them, and overcome their opponent's defense to score goals. GRF provides a variety of challenging scenarios that effectively evaluate both single-agent and multi-agent RL algorithms. It simulates an 11 vs. 11 football match, featuring customizable difficulty levels and reward systems, performing strategic plays, and dynamic defensive actions. The environment also provides the Football Academy, a set of progressively harder and diverse reinforcement learning scenarios which makes this environment a good fit for curriculum learning.

Recent works like TiZero, which is a state-of-the-art multi-agent framework combine progressive curriculum learning with large-scale distributed self-play [3]. While TiZero achieves exceptional performance, it requires hundreds of CPUs and extensive engineering, making it difficult to reproduce or scale to lower-resource settings. Other works like COST (Curriculum-Oriented Skills and Tactics) take a different approach by employing an automated teacher–student framework to construct a task curriculum based on varying team sizes [10]. The student agent, equipped with a hierarchical policy and population-invariant communication, learns transferable skills across scenarios.

While these methods demonstrate the power of curriculum-based training, they primarily focus on full-team control and rely on substantial computational resources. In contrast, single-agent training using curriculum learning in GRF, to the best of our knowledge, remains underexplored, particularly in resource-constrained environments. Our work addresses this gap by investigating whether a single-agent controller, trained using a well-structured curriculum, can achieve competitive performance in GRF, using only CPU resources and limited training time.

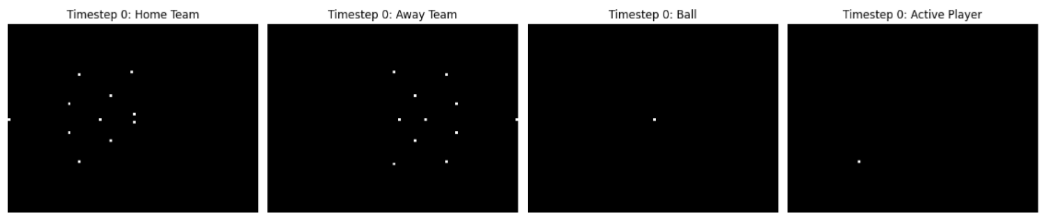

Figure 2: Environment Representation: each observation features four binary 2D channels, each depicting the positions of the: home team, away team, ball, and the active player.

3 Methodology

Our methodology involves integrating curriculum learning within the GRF environment football academy to progressively train reinforcement learning agents from simpler tasks to more complex scenarios. The steps involved in our approach are outlined below:

3.1 Environment Setup

The GRF environment is configured to simulate various football scenarios, ranging from simple tasks like scoring from a static position to more complex tasks involving dynamic team coordination. The environment settings include adjustable difficulty levels and customizable parameters such as opponent strength, number of controlled players, and specific scenario goals.

We use the stacked representation of the game with checkpoint rewards as it gave the best results in the original paper. Example can be seen in Figure 2. Observations are structured as tensors of shape (72, 96, 16), capturing spatial and temporal features of the game environment. Actions were discrete, encompassing 19 distinct possibilities covering various football-specific maneuvers. Throughout training, we focus on a single agent: the player currently in possession of the ball.

3.2 Curriculum Design

Our training curriculum progresses through the following stages, each building in complexity:

1. empty goal (close)
2. empty goal
3. run to score
4. run to score with keeper
5. pass and shoot with keeper
6. run pass and shoot with keeper
7. 3 vs 1 with keeper

These scenarios are available through the GRF Academy. This structured progression helps the agent master basic skills before tackling complex, dynamic play.

3.3 Training and Evaluation

The agent is trained using the Proximal Policy Optimization (PPO) algorithm with up to 5 million steps per academy stage, and an early-stopping mechanism was implemented to speed up the process. Other hyperparameters are the same as in the original paper [2] to maintain the comparability of results. The agent is evaluated based on the number of goals scored. This methodology allows us to systematically assess the effectiveness of curriculum learning in enhancing learning efficiency and overall agent performance in the complex and realistic GRF environment. The Football Academy features a total of 13 different scenarios, but due to the computational and time limitations we focused on the first seven scenarios, presented above.

4 Experiments and Results

We trained a single-agent model using the PPO algorithm and the deep learning model architecture presented in the original work by Kurach *et al.* We used the checkpoint reward system and set an early stopping threshold based on the running average of the last 200 training episodes. For the first 4 scenarios the early stopping threshold was set to 1.9, however, during the "Pass and Shoot with Keeper" scenario, training platoed at the reward values of around 1.7, leading us to drop the early stopping threshold to that value. The number of training steps per each of the individual scenarios is presented in the Table 1. To quantify the performance of our algorithm and to compare it with the previously presented results, we measured the average goal difference scores from a 100 episodes after the training was complete; these results are presented in the Figure 3.

Table 1: Number of Training Episodes spent on each of the Football Academy Scenarios, along with the running number of total steps trained through the curriculum. The last column indicates how our model compares to the baseline.

Football Academy Scenario	**N Steps / Scenario**	**Total Steps**	**Ours VS Baseline**
Empty Goal Close	139,728	139,728	$>$1M, $\approx$ 50M
Empty Goal	1,371,408	1,511,136	$\approx$ 50M
Run to Score	188,864	1,700,000	$\approx$ 1M
Run to Score with Keeper	2,501,792	4,201,792	$\gg$ 50M
Pass and Shoot with Keeper	5,000,000	9,201,792	$\approx$ 5M
Run, Pass and Shoot with Keeper	879,088	10,080,880	$\approx$ 5M
3 vs 1 with Keeper	4,325,792	14,406,672	$\approx$ 5M

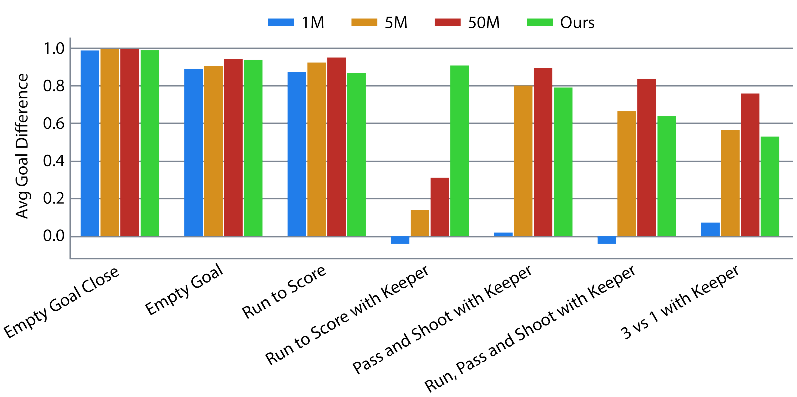

Figure 3: Average goal difference on *Football Academy*, comparison of our curriculum learning approach (green) with the baseline PPO training results featuring three models trained for a different number of steps each: 1M, 5M, and 50M (blue, yellow, red) presented in Kurach *et al.* [2].

5 Conclusion and Future Work

In this paper, we explore the effectiveness of curriculum learning for training a single-agent policy in the Google Research Football environment under tight resource constraints. By leveraging a carefully curated progression of Football Academy scenarios, we demonstrate that a standard PPO agent can achieve competitive performance, matching or even surpassing more computationally-heavy baselines. Remarkably, our results also show greater training consistency between the different curricula, enabling previously untrainable scenarios, such as Run to Score with Keeper, to be learned efficiently using a fraction of compute resources. Our results show that with an efficient training curriculum, single-agent control can learn transferable football skills and strategic behaviors typically associated with much larger-scale multi-agent training regimes.

Several promising directions emerge from this work, like investigating whether the learned single-agent policy could serve as a strong initialization for multi-agent systems, where individual agents are bootstrapped with pre-trained skills before learning to coordinate. This modular transfer could significantly reduce the training burden in full 11v11 settings. Another direction could be replacing the fixed curriculum with an adaptive system that adjusts task difficulty based on the agent's performance.

References

[1] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 41–48, 2009.

[2] Karol Kurach, Anton Raichuk, Piotr Stańczyk, Michał Zając, Olivier Bachem, Lasse Espeholt, Carlos Riquelme, Damien Vincent, Marcin Michalski, Olivier Bousquet, et al. Google research football: A novel reinforcement learning environment. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 4501–4510, 2020.

[3] Fanqi Lin, Shiyu Huang, Tim Pearce, Wenze Chen, and Wei-Wei Tu. Tizero: Mastering multi-agent football with curriculum learning and self-play, 2023.

[4] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei Rusu, Joel Veness, Marc Bellemare, Alex Graves, Martin Riedmiller, Andreas 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*, 518:529–33, 02 2015.

[5] Sanmit Narvekar, Bei Peng, Matteo Leonetti, Jivko Sinapov, Matthew E Taylor, and Peter Stone. Curriculum learning for reinforcement learning domains: A framework and survey. *Journal of Machine Learning Research*, 21(181):1–50, 2020.

[6] Sanmit Narvekar and Peter Stone. Learning curriculum policies for reinforcement learning. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '19, page 25–33, Richland, SC, 2019. International Foundation for Autonomous Agents and Multiagent Systems.

[7] Sanmit Narvekar and Peter Stone. Generalizing curricula for reinforcement learning. In *4th Lifelong Machine Learning Workshop at ICML 2020*, 2020.

[8] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms, 2017.

[9] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.

[10] Rundong Wang, Longtao Zheng, Wei Qiu, Bowei He, Bo An, Zinovi Rabinovich, Yujing Hu, Yingfeng Chen, Tangjie Lv, and Changjie Fan. Towards skilled population curriculum for multi-agent reinforcement learning, 2023.