# Investigating Learning Methods and Environment Representation in the Construction of Player Agents: Application on FIFA Game

Matheus Prado Prandini Faria
*Computer Science Department*
*Federal University of Uberlândia*
Uberlândia, Brazil
matheusprandini.96@gmail.com

Rita Maria Silva Julia
*Computer Science Department*
*Federal University of Uberlândia*
Uberlândia, Brazil
ritasilvajulia@gmail.com

Lídia Bononi Paiva Tomaz
*Computer Science Department*
*Federal Institute of Triângulo Mineiro*
Uberaba, Brazil
ldbononi@gmail.com

***Abstract*—The objective behind this study is to investigate Machine Learning (ML) techniques combined with methods from Computer Vision (CV) for state representation by images, to produce agents capable of solving problems, in real time, in environments with complex properties. Such difficulties require agents to be highly efficient in their learning (and, consequently, decision-making) and environmental perception processes, without which they will not be successful. The digital game FIFA - soccer simulator - is used as a case study because it represents a realistic and challenging environment. The ML techniques are investigated in the context of the Deep Learning (DL) approach provided by Convolutional Neural Networks (CNNs), being: imitation learning, used here with the purpose of endowing the agent with the ability to solve problems in a way closer to human; by deep reinforcement, in which the agent is trained in an attempt to autonomously abstract an optimal decision-making policy. Regarding the environmental perception, the following state representations approaches are investigated in this study: raw images - with and without color information - and through Object Detection Techniques (ODT). In order to further improve the performance of the agents produced, genetic algorithm techniques are explored to automatically define a CNN architecture to be used as the player agents decision-making module. In addition to corroborating the excellent results that DL combined with CV has been producing in the context of ML (particularly in games), the present work shows the great potential of the application of ODT in the process of enhancing the environmental perception, which counts as a relevant counterpart to the fact that ODT demands computational procedures with a higher cost in relation to the representations based on raw images.**

***Index Terms*—Deep Learning, Imitation Learning, Deep Reinforcement Learning, Object Detection Techniques, Genetic Algorithms**

## I. Introduction

Deep Learning (DL) has consolidated itself as one of the pillars for the progress of Artificial Intelligence (AI) and Machine Learning (ML) techniques [1]. DL enables computers to learn from experience and understand the world in terms of a hierarchy of concepts [2]. Combined with Computer Vision (CV), DL has been providing notorious advances in solving a variety of practical tasks that the best conventional ML techniques have many limitations on solving. Thus, it is important to highlight the relevance of DL+CV researches in problems such as self-driving systems and digital games [3].

The digital games stand out as a case study in this context for the following reasons: in addition to presenting a high impact on the current economic scenario, they involve technical difficulties that are very similar to those faced by agents that deal with problems of the real-life [4]. So, digital games provide an excellent laboratory to develop and test ML techniques, from Imitation Learning (IL) methods to Deep Reinforcement Learning (DRL) approach. Among the very defying game case studies, the successful video game FIFA highlights. This game, despite being a worldwide popular game and presenting challenging aspects to the process of building player agents, is still an almost unexplored problem in terms of ML researches compared to other famous games such as Starcraft and Dota.

The main reasons for this are: 1) the fact that these games have specific user-friendly interfaces for the development of scientific research, a result of the partnership between the big companies *Google* and *OpenAI* with the producers of these games (*Blizzard* and *Valve* respectively). These interfaces are designed so that the developed agents are easily connected and coupled to the gaming platform. It is noteworthy that there is no similar existing interface integrated into *FIFA*; 2) as there is no such interface for that game, the main source of representation of the environment are the observations produced through the manipulation and processing of game images. This is justified by the fact that the real game states are not accessible, since relevant information about them underlies the internal game mechanism (for example, the real position of a player is located in some position of memory although it is possible to identify such a player in the game image).

In contrast to the facts mentioned above, *FIFA* presents a great variety of interesting scenarios with a great potential to be explored by ML researches. It has a main game mode consisting of a competitive environment between two agents in which each one of them controls a soccer team composed of 11 players. In addition to this mode, the game has a large

The authors thank CAPES for financial support.

number of different environments aimed at improving the game skills of users. They involve tasks related to exchanging passes between players, finishing on goal, taking free kicks, dribbling with static and dynamic obstacles, among others. It is important to highlight that these environments provide different aspects, objectives and levels of complexity to be taken into account in the process of building intelligent agents.

Thus, considering the points presented here, the general objective of this work was to investigate appropriate ML techniques - IL and DRL approaches - combined with environment representation techniques from CV, to produce agents capable of solving problems, in real-time, in fully observable, stochastic and unknown environments that vary according to the following properties: single or multi-agent, static or dynamic. Therefore, the present work uses three distinct *FIFA* game scenarios as a laboratory for the study of such techniques for the development of automatic agents. Thus, each of the scenarios covered by this work were investigated from two perspectives: 1) Analysis of the most appropriate learning strategy (supervised - through IL or DRL); 2) Investigation of the most appropriate way to represent the environment (raw images of the game scenario or derived from ODTs).

## II. FIFA Game

FIFA is a football simulation video game developed and published by EA Games. In particular, this work explores two free-kicks scenarios: 1) **Single-Agent Scenario**: involves situations with and without barriers simulating players blocking the goal. Despite the difficulty of scoring goals in the presence of these barriers, it does not include a goalkeeper. This scenario is illustrated in Fig. 1; 2) **Multi-Agent Scenario**: involves situations that include barriers and a goalkeeper. This scenario is illustrated in Fig. 2. In this context, the goalkeeper represents an opponent who tries to minimize the actions of the free-kicks taker agent. Furthermore, this work also explores the **Confrontation Mode**, showed in Figure 3, that deals with situations involving two soccer players controlled by the user (in yellow uniform) against two soccer players manipulated by the FIFA's engine (in orange uniform).

Fig. 1. Example of free-kicks situation in the single-agent scenario.

## III. General Architecture of FIFA Player Agents

In particular, as shown in Figure 4, this work considers that an intelligent agent is composed of three main modules: Representation; Decision-Making; and Learning. The first module is strictly related to the agent's perception of the environment,

Fig. 2. Example of free-kicks situation in the multi-agent scenario.

Fig. 3. Example of confrontation mode situation.

processing the information captured by the sensors in order to generate an adequate representation (explained in section IV). The second module is responsible for deciding the actions to be performed in the environment based on the representations generated by the first module. In this research, this module is represented by Deep Neural Networks - more specifically, CNNs and Multilayer Perceptrons (MLPs). It is noteworthy here that this work proposed the Minimum CNN-GA (MCNN-GA) method, which automatically defines CNN architectures through a policy that minimizes the weigh vector dimensions (or number of parameters) and the classification error rate of the CNNs. Finally, the last module has the function of allowing the agent to have the ability to improve their decision making based on examples provided by a supervisor or from their own experience (described in section V).

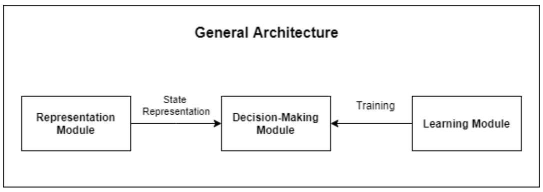


Fig. 4. General architecture of intelligent agents considered in this work

## IV. Environment Representations

The present work considered three different environment (or state) representations: raw images without color information; raw images with color information; from ODTs. In the raw images without color information, the Representation Module transforms the original image (representing the current frame) into a smaller size image and converts it to grayscale (removes the color information). Analogously, this same process is performed for the raw images with color information, keeping, however, the color information (RGB channels). Finally, in the ODT representation, the original image (current frame) is processed by a feature extractor model (in the case of this work, MobileNetV2) to generate a simplified representation (feature maps) which is used by a SSD model [5] to perform the detection of bounding boxes that outline the objects to be identified (the coordinates of these bounding boxes are then used to represent the game state).

## V. Learning Strategies

The Learning Module of the agents produced in this work was composed of DRL or IL approaches. The former is represented by the Deep Q-Learning (DQN) algorithm and the latter is represented by the Direct Imitation Learning (DIL) and Deep Active Imitation (DAI) methods.

### A. Deep Reinforcement Learning

DQN [6] is a value-function based DRL algorithm which achieved scores across a wide range of classic Atari games that were comparable to that of a professional players [7]. DQN combines the advantages of deep learning using deep CNNs for abstract representation with the Q-learning method [8] in order to learn an optimal policy based, exclusively, on the images that represent the game states.

### B. Imitation Learning

DIL via behavioral cloning [9] is one of the main methods to cope with an IL problem. This method usually assumes that an agent receives demonstrations consisting of observation-action pairs [10]. The goal is to produce a policy that direct maps observations to actions from the demonstrations in the training data. It consists of the following processes: 1) collecting demonstrations from the expert; 2) supervised training of a CNN from the data set composed of the collected demonstrations.

DAI is a method for learning from demonstrations using active learning and deep neural networks proposed in [11]. This method extends DIL approach with the following process: use of active learning to refine the initially learned policy, making the policy more robust (it queries the expert the optimal action under situations with low confidence). These active samples are then collected and used to improve the policy weaknesses [11].

## VI. Experiments and Results

The purpose of the experiments is to evaluate the performance of the player agents generated in this work through the FIFA game scenarios explored here. For each of these scenarios, several agents were produced with different combinations of game state representation, decision-making module and learning method. The first and second test scenario evaluates the performance of player agents in the free kicks scenarios (**Single-Agent Scenario** and **Multi-Agent Scenario**) and **Confrontation Mode** respectively.

In order to execute the experiments, this work proposed an interface to connect these player agents to the complex and challenging environment of *FIFA*, as there is no interface integrated existing into this game [12]. Basically, through this connection interface, it is possible for the player agents to interpret the state of the environment, make their decisions (performing actions in the game) and measure the impact of such decisions in order to improve their learning.

### A. Test Scenario 1

The objective of this test scenario is to evaluate the performance of the agents produced for both free kicks scenarios (**Single-Agent Scenario** and **Multi-Agent Scenario**).

Firstly, agents based on the DQN algorithm were evaluated on the Single-Agent (SA) scenario considering different environment representations and decision-making modules. Table I shows the results obtained in terms of the convergence time (number of episodes) for agents to achieve an average of 85% goals in the last 100 episodes. These results were retrieved considering 10 training sessions. It is important to note that the best performance was from $DQN\text{-}SA\text{-}Agent$ + **ODT** + **MLP** variation, which achieved the specified average of goals in 782 episodes approximately (outperforming agents based on the raw images representations). Considering the Multi-Agent (MA) scenario, all DRL-based agents converged to worse goal rates than random agents due to the difficulty of the environment (since there is a goalkeeper). This fact motivated the study of IL-based agents.

TABLE I
RESULTS OF $DQN\text{-}SG\text{-}Agent$ AND ITS VARIATIONS IN TERMS OF THE CONVERGENCE TIME (NUMBER OF EPISODES)

| Agent | Mean |
|---|---|
| $DQN\text{-}SA\text{-}Agent$ + **Grayscaled Images** + **CNN** | 1577,10 |
| $DQN\text{-}SA\text{-}Agent$ + **Colored Images** + **CNN** | 912,40 |
| $DQN\text{-}SA\text{-}Agent$ + **ODT** + **MLP** | 781,80 |

Table II shows the results obtained with respect to the IL-based agents (built from the DAI method) in terms of the rate of goals (considering 10 training sessions). Regarding the SA scenario, agents obtained rates above 90% (it is important to highlight that the human supervisor agent obtained a 95% rate). Regarding the MA scenario, $DAI\text{-}MA\text{-}Agent$ + **Colored Images** + **CNN** stood out with rate of goals equal to 7% (very close to the rate obtained by the human supervisor, which was 7.5%).

TABLE II
RESULTS OF $DAI\text{-}SA\text{-}Agent$ AND $DAI\text{-}MA\text{-}Agent$ VARIATIONS IN TERMS OF RATE OF GOALS

| Agent | Mean |
|---|---|
| $DAI\text{-}SA\text{-}Agent$ + **Grayscaled Images** + **CNN** | 90,50 |
| $DAI\text{-}SA\text{-}Agent$ + **Colored Images** + **CNN** | 91,60 |
| $DAI\text{-}SA\text{-}Agent$ + **ODT** + **MLP** | 92,10 |
| $DAI\text{-}MA\text{-}Agent$ + **Grayscaled Images** + **CNN** | 5,80 |
| $DAI\text{-}MA\text{-}Agent$ + **Colored Images** + **CNN** | 7,00 |
| $DAI\text{-}MA\text{-}Agent$ + **ODT** + **MLP** | 1,70 |

### *B. Test Scenario 2*

The objective of this test scenario is to evaluate the performance of the agents produced for the **Confrontation Mode** (CM). This evaluation performs a comparative analysis between IL methods considering optimized decision-making modules (using MCNN-GA) and non-optimized ones.

Table III show the results obtained by the IL-based agents in terms of in-game score considering 50 matches against the FIFA's engine (the goal is to reach the highest score possible). The results indicate a very close performance among agents built without and with MCNN-GA. Although the agents built from MCNN-GA have not significantly improved performance in CM, it is interesting to note that they are much lighter in terms of the number of parameters compared to the agents that do not use MCNN-GA. The latter agents consist of a CNN with 3.181.400 parameters, while the CNNs built using MCNN-GA has approximately 20 times less parameters [13], which has a direct impact on memory (space complexity) and inference time. The results from Table III also indicate a superior performance of the variation that considers the **Colored Images** representation over **Grayscaled Images**, since the best performance is associated with *DAI-CM-Agent* + **Colored Images + MCNN-GA**, suggesting that raw images with color information have a greater capacity to generate relevant features for player agents' decision-making compared to raw images without color information.

TABLE III
RESULTS OF IL-BASED AGENTS IN TERMS OF THE IN-GAME SCORE

| Agent | Mean |
|---|---|
| *DIL-CM-Agent* + **Grayscaled Images + CNN** | 1100,00 |
| *DIL-CM-Agent* + **Grayscaled Images + MCNN-GA** | 1298,00 |
| *DIL-CM-Agent* + **Colored Images + CNN** | 1752,00 |
| *DIL-CM-Agent* + **Colored Images + MCNN-GA** | 1710,00 |
| *DAI-CM-Agent* + **Grayscaled Images + MCNN-GA** | 1622,00 |
| *DAI-CM-Agent* + **Colored Images + MCNN-GA** | 2172,00 |

## VII. PUBLICATIONS AND AWARDS

The main results of this dissertation in terms of scientific publication are listed below:

- Abstract titled Improving FIFA Free Kicks Player Agent Performance through Object Detection Techniques published in the proceedings of XIII Computer Science Theses and Dissertations Workshop (WTDCC - UFU 2019), awarded as the second best master's research. An extended version of this work was published in **SBGames 2020 (Qualis B1)** [14].
- [12] **FLAIRS Conference (Qualis A4)**.
- [15] **ICTAI (Qualis A3)**.
- [16] **ICMLA (Qualis A3)**.
- [13] **BRACIS (Qualis A4)**.

## VIII. CONCLUSION AND FUTURE WORK

This work investigated ML techniques - both the DRL and IL approaches - combined with CV-based environment representation techniques - using CNNs and also ODTs to efficiently process raw game images - in the process of building agents that solve problems, in real time, in some FIFA games scenarios. Therefore, at first, a connection interface was built and validated. Regarding the agents built here, the learning methods were evaluated combined with the environment representations. Particularly, game states from colored images (RGB) had the best results overall despite attesting the potential that ODTs have in improving the quality of the perception of the environment (noting that their problem is the high processing time). Among the learning methods considered, those based on IL were fundamental to deal with the property of unknown inherent in FIFA. Regarding the MCNN-GA method, the results obtained showed the huge gain in the construction of automated CNNs in terms of the number of parameters, which is interesting in the context of real-time environments. As future work, the authors intend to investigate the use of games in the medical context.

## REFERENCES


[1] Y. LeCun, Y. Bengio, and G. E. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, book in preparation for MIT Press.

[3] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11 – 26, 2017. [Online]. Available: https://doi.org/10.1016/j.neucom.2016.12.038

[4] L. B. P. Tomaz, "Adaba: a new alpha-beta distribution approach - application to the checkers game domain," Ph.D. dissertation, Universidade Federal de Uberlândia, 2019.

[5] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 2016*, 2016, pp. 21–37.

[6] V. Mnih, K. Kavukcuoglu, D. Silver, and A. A. Rusu, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[7] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[8] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, 1st ed. Cambridge, MA, USA: MIT Press, 1998.

[9] M. Bain and C. Sammut, "A framework for behavioural cloning," in *Machine Intelligence 15*, 1995.

[10] Y. Liu, A. Gupta, P. Abbeel, and S. Levine, "Imitation from observation: Learning to imitate behaviors from raw video via context translation," *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1118–1125, 2018.

[11] A. Hussein, E. Elyan, M. M. Gaber, and C. Jayne, "Deep imitation learning for 3d navigation tasks," *Neural Computing and Applications*, vol. 29, no. 7, pp. 389–404, Apr 2018.

[12] M. P. P. Faria, R. M. S. Julia, and L. B. P. Tomaz, "A deep reinforcement learn-based fifa agent with naive state representations and portable connection interfaces," in *International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, 2019, pp. 282–285.

[13] ——, "Improving FIFA player agents decision-making architectures based on convolutional neural networks through evolutionary techniques," in *Intelligent Systems - 9th Brazilian Conference, BRACIS 2020, Rio Grande, Brazil, October 20-23, 2020, Proceedings, Part I*, ser. Lecture Notes in Computer Science, R. Cerri and R. C. Prati, Eds., vol. 12319. Springer, 2020, pp. 371–386.

[14] ——, "Improving fifa free kicks player agent performance through object detection techniques," in *Proceedings of SBGames 2020 - XIX Simpósio Brasileiro de Jogos e Entretenimento Digital*.

[15] ——, "Deep active imitation learning in fifa free-kicks player platforms based on raw image and object detection state representations," in *31st International Conference on Tools with Artificial Intelligence*, 2019.

[16] ——, "Evaluating the performance of the deep active imitation learning algorithm in the dynamic environment of fifa player agents," in *IEEE International Conference on Machine Learning and Applications*, 2019.