Improving FIFA Free Kicks Player Agent Performance through Object Detection Techniques

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*—Video games stand out as an important case study in the Machine Learning and Computer Vision based researches for the following reasons: in addition to presenting challenging technical difficulties, they have a high impact on the current economic scenario. Although several of these researches use Deep Reinforcement Learning (DRL) algorithms as a learning approach, few of them use Object Detection Techniques (ODTs) to improve their state representation strategy. Motivated by these arguments, the present paper proposes two DRL+ODT based agents to cope with free kicks in the FIFA digital game. The main contributions here are: to test the performance of the enhanced ODT version named MobileNetV2 in dynamic environments; and improving the performance of the free kicks agents through the inclusion of the barrier element as a new object to be detected by MobileNetV2 ODT. The results confirm the efficiency of MobileNetV2 in terms of the performance of the agents in dynamic environments.**

***Index Terms*—Object Detection, State Representation, Deep Reinforcement Learning, Games**

I. Introduction

Machine Learning (ML), with emphasis on the Deep Reinforcement Learning (DRL) algorithms, has been an outstanding scientific field of research and significant progress has been made in solving challenges of everyday life using such methods [1]. Several problems addressed by these agents require the Computational Vision (CV) ability as an additional resource to be coupled to their learning processes, such as digital games and self-driving systems [2]. The Object Detection Techniques (ODTs) represent a relevant CV auxiliary tool to improve the agents' ability to perceive the environment in which they actuate, which enhances their decision-making capacity [3]. Noteworthy here is that few works applied ODTs to the learning process of the player agents [4].

In this context, Trivedi proposes a DRL-based agent for the task of free kicks in a complex 3D football simulation Video Game developed and published by Electronic Arts (EA Games) - FIFA - considering a simplified scenario that does not include a goalkeeper [5]. In that work, the author trained an ODT based on the Single Shot MultiBox Detector (SSD) classifier [6] and on the MobileNetV1 [7] feature extractor to identify relevant objects like the ball, the player and the goal. This model was used to generate the state representation to be processed by the Deep Q-Network (DQN) RL algorithm. However, the results obtained showed that it is possible to significantly improve the developed approach.

The work [8] proposes an improved version of MobileNetV1 feature extractor, named MobileNetV2, and evaluates the performance of this new version only in a static scenario of object detection (testing in databases).

Motivated by this, the general objective of the present paper is to extend both approaches proposed in [5] and [8], to improve the training time and the performance of DRL+ODT based FIFA player agents.

Thus, the main contributions here are: 1) Implementing two new versions of FIFA automatic players that improve the agent proposed in [5] by replacing its MobileNetV1 based feature extractor with the enhanced version MobileNetV2. These new architectures allow for: evaluating the performance of MobileNetV2 in a dynamic scenario, which extends [8]; and mitigating the training time of player agents; 2) Enhancing the agent's environment perception by including the barrier among the objects detected through ODT, which also extends and improves Trivedi work.

The experiments confirmed that the improvements in the agent's learning ability provided by the approach proposed herein resulted in positive effects in terms of the following evaluative parameters: score of well-succeeded kicks (rate of goals) and training time.

The next sections are structured as following: Section II resumes the background; Section III describes the approach proposed in this paper; Section IV shows the experimental results; finally, Section V presents the conclusion and future works.

II. Background

This section presents a summary of the SSD MobileNet, the DQN model, and FIFA's environment.

A. SSD MobileNet

SSD is a method for detecting objects in images using a single deep neural network that is faster and significantly more accurate than the previous state-of-art for single shot detectors (for example, YOLO) [6]. The present paper uses a version of SSD with MobileNetV1 as the base network and another SSD

The authors thank CAPES for financial support.

version composed of MobileNetV2 [8]. The base network in this scenario works as a feature extractor. Shortly, MobileNet is a family of lightweight deep neural networks based on an architecture that uses depthwise separable convolutions [7].

Sandler et al. [8] described a new mobile architecture, denominated MobileNetV2, that improves the state of the art performance of mobile models on multiple tasks, including object detection. The authors evaluated and compared the performance of MobileNetV1 and MobileNetV2 as feature extractors with a modified version of the SSD in a static scenario (using COCO dataset [9]) through two evaluative metrics: mean average precision (mAP) and running time (in milliseconds). The results showed that the version implemented with the MobileNetV2 achieved a similar level of performance to MobileNetV1, but with faster running time.

B. DQN

DQN [10] is a value-function based DRL method that achieved scores across a wide range of classic Atari 2600 video games that were comparable to that of a professional video games tester [11]. This method combines the advantages of deep learning using Convolutional Neural Networks (CNN) for abstract representation with the Q-learning method [12] to learn an optimal policy based on the screen pixels that represent the game states exclusively.

C. FIFA Environment

FIFA is a football simulation game, developed and published by EA Game. In particular, this work explores the free kicks mode that deals with 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 environment is stochastic, which is a major challenge in the learning process of an agent.

In [5], an object detection model using the SSD MobileNetV1 technique is trained in order to identify the following elements: ball, goal, and players. In this way, MobileNetV1 processes the game screenshots giving a 128-dimensional flattened feature map as state representation to the DQN model considering the identified elements. The resulting model achieved a 50% rate of goals with 1000 training epochs in a GPU GTX-1070. It is important to note that the model was not trained to detect some other relevant objects to the free kicks, such as barriers, which could improve the performance of the DQN model.

III. DRL Agents based on Object Detection

This section presents the object detection framework used in the present paper. Figure 1 shows the general architecture of the DRL-based agents. Basically, these agents are benefited by the state representation provided by the object detection model in order to perform their decision-making through the DQN model upon the FIFA free kicks task.

The object detection model is the module refined by the contributions proposed in the present work, for this reason, its details will be presented in subsection III-A. The DQN model used in this work is the same employed in [5]. It receives a 128-dimensional flattened feature map as input. In addition, it has two dense layers containing 512 neurons each. Finally, the output layer is composed of 4 neurons (corresponding to the actions that can be performed). The FIFA free kicks environment is constituted by the scenario described in subsection II-C. In this way, the agent has the task to score as many free kicks as possible. Considering this context, it is possible to formulate the free kicks as a reinforcement learning problem as follows:

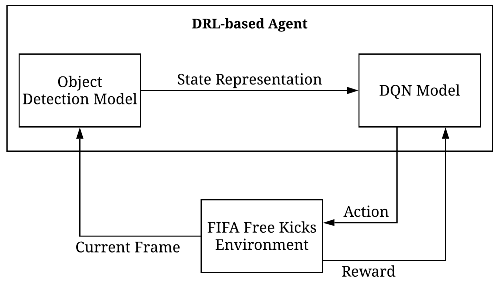

Fig. 1. General architecture of the DRL-based agents.

- **States**: the real game state is not fully represented in its corresponding image, since relevant information concerning it underlies within the game's engine, which is not accessible. Thus, the performance of the learning process strongly depends on state representation. In this paper, the state representation explored is presented in subsection III-A.
- **Actions**: there are four possible legal actions to be executed: *move left*, *move right*, *low kick* and a *high kick* at an established height (which usually ranges from the height of the barrier and the height of the goal). It is important to note that the kick power is considered constant in both possibilities of kicking, being empirically defined by one of the authors with advanced expertise.
- **Rewards**: after the execution of a kick action, if the agent scores a goal, then the reward is equal to one. Otherwise, the reward for the agent is equal to minus one. For intermediate actions (*move left* and *move right) the reward is equal to zero.*

A. Object Detection Model

The object detection model aims to generate the state representation to be processed by the agents for their decision-making along the situations faced in the free kicks, as shown in Figure 2. The **Mobile Feature Extractor** module, which corresponds to a CNN, process this frame and generates a 128-dimensional feature map, named **High Level Feature Map**. This feature map contains valuable information about the elements desired for detection and also general information about the frame. Then, it feeds an SSD model to perform the detection of bounding boxes that outline these relevant elements. It is important to note that the **SSD** module does not influence the state representation explored by the present paper, being important only for the training of the feature extractor.

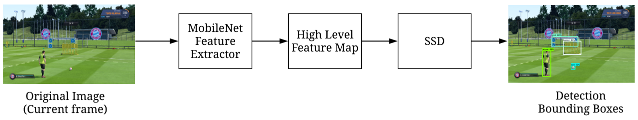

Fig. 2. Object Detection Architecture [13].

In [5], the elements considered were ball, player, and goal. Additionally, the MobileNetV1 was used as a feature extractor. The agent generated by the mentioned work is named here as *A1*. In this way, the present work improves [5] as follows: 1) updating the feature extractor from MobileNetV1 to an improved version, MobileNetV2 (**Improvement 1**). It is expected that this new version is better suited to handle the task of object detection by improving the quality of a DRL-based agent both in terms of training time and in-game performance. This improvement generated an agent named here as *A2*; 2) enhancing the state representation by adding barrier - an extremely relevant element in the context of free kicks - to the objects processed by MobileNetV2 (**Improvement 2**). This improvement generated an agent named here as *A3*. Increasing the quality of a state representation is essential in the learning process of DRL-based agents. Table I summarizes the main characteristics of the agents proposed here.

TABLE I
DRL-BASED AGENTS IMPLEMENTED FOR FREE KICKS

Agent	Feature Extractor	Objects
A1	MobileNetV1	Ball, Goal and Player
A2	MobileNetV2	Ball, Goal and Player
A3	MobileNetV2	Ball, Barrier, Goal and Player

IV. Experiments and Results

The experiments have two objectives: 1) to evaluate **Improvement 1**: the impact of updating the ODT based on MobileNetV1 to the ODT based on MobileNetV2 on the DRL-based agent performance (subsection IV-A); 2) to evaluate the impact on the learning process quality of the agents considering the state representation used in [5] and the one proposed in this work (**Improvement 2**), which extends the former by adding the barrier as a relevant element to the object detection model (subsection IV-B). Both evaluations were performed employing comparative tests based on two parameters: the rate of goals and training time.

To reach the objectives, the DRL agents were trained with the four actions described in section III (move to left and right, low, and high kicking). It is important to note that each agent was trained in the course of 1000 epochs. An epoch ends when a kick action is executed. Ten sessions were then performed with the trained agents, each composed of 100 epochs, in order to verify whether there was statistical significance in the performance obtained by means of *independent-samples t-test* [14] using the software package SPSS.

The experiments were executed in an architecture composed of a machine with a GPU Nvidia GeForce GTX-745 and 16 GB RAM.

A. Evaluating MobileNetV1 x MobileNetV2 as feature extractor in a dynamic scenario

To deal with the first objective, the agent produced in [5], *A1*, which uses the ODT based on MobileNetV1, is compared with a newer agent version with ODT based on MobileNetV2, *A2*. Thus, the main objective of this first experiment is to investigate the impact on the learning process using the MobileNetV2 as a feature extractor (**Improvement 1**).

TABLE II
PERFORMANCE *A1* AND *A2* REGARDING THE TRAINING PHASE OVER 1000 EPOCHS.

Agent	Average rate of goals (in %)	Training time (in minutes)
A1	62.1	360
A2	63.4	270

Table II shows that *A2* presented a similar level of play in terms of the average rate of goals compared to *A1* in the training phase over 1000 epochs. However, it performed much better with respect to training time. In fact, *A2* demanded 90 minutes less (25% less time) than *A1* to complete the training phase and with a slightly higher rate of goals.

TABLE III
PERFORMANCE OF *A1* AND *A2* REGARDING THE TEST PHASE OVER 10 SESSIONS.

	1	2	3	4	5	6	7	8	9	10	Mean	Standard Deviation
A1	76	77	77	81	72	77	72	71	78	82	76.3	3.713
A2	77	76	81	79	82	74	85	75	73	77	77.9	3.814

Thus, a final test was conducted to compare the performance of such trained agents by performing 10 sessions composed of 100 epochs each. Table III shows, respectively, the rate of goals (in %), the total mean, and the standard deviation of both agents over the sessions. An *independent-samples t-test* with a *significance level* ($\alpha = 0.05$) was conducted to compare the performance of *A1* and *A2*. In this manner, the following null hypothesis H_0 was created: *A1* holds the same level of performance as *A2*.

The statistical test indicated that the mean rate of goals for *A2* is greater than the mean for *A1* (*t-value* = -0.951). However, since *p-value* $= 0.354$ is not lower than the *significance level* ($\alpha = 0.05$), the null hypothesis cannot be rejected. These results suggest that changing the feature extractor from MobileNetV1 to MobileNetV2 does not have a significant effect on the rate of goals of a DRL-based agent in the free kicks task. It should be noted that these results were consistent with those presented in [8], which compared the same feature extractors in a static scenario, since both had similar performance differentiating only in the processing time. Therefore, it is concluded that the great advantage of *A2* over *A1* is the smaller training time.

B. Evaluating the enhanced state representation

In this experiment, the agent *A2* produced in the previous step, which uses the MobileNetV2 as feature extractor, is compared with a new agent version implemented with the same ODT, but with the differential of detecting the barrier element (**Improvement 2**) - incorporated in the state representation generated by the object detection model - *A3*. Thus, the main objective of this second experiment is to investigate the impact on the learning process by adding a new element, the barrier. In fact, this element is extremely relevant in this context, since it is considered the main obstacle in the free kicks task, providing a great challenge to the learning process of player agents.

TABLE IV
PERFORMANCE OF *A2* E *A3* REGARDING THE TRAINING PHASE OVER 1000 EPOCHS.

Agent	Average rate of goals (in %)	Training time (in minutes)
A2	63.4	270
A3	66.7	290

Table IV shows that *A3* presented a higher level of play in terms of the average rate of goals compared to *A2* in the training phase over 1000 epochs. The average rate of goals scored by *A3* was 3.3% greater than the average rate of goals scored by *A2*. However, in terms of training time, the performance was similar.

TABLE V
PERFORMANCE OF *A2* AND *A3* REGARDING THE TEST PHASE OVER 10 SESSIONS.

	1	2	3	4	5	6	7	8	9	10	Mean	Standard Deviation
A2	77	76	81	79	82	74	85	75	73	77	77.9	3.814
A3	85	81	87	83	85	83	87	82	83	80	83.6	2.366

Thus, a final test was conducted to compare the performance of such trained agents by performing 10 sessions composed of 100 episodes each. Table V shows, respectively, the rate of goals (in %), the total mean and the standard deviation of both agents over the sessions. Analogously to experiment 1, an *independent-samples t-test* with a *significance level* ($\alpha = 0.01$) was conducted to compare the performance of the agents *A2* and *A3*. In this manner, the following null hyphotesis H_0 was created: *A2* holds the same level of performance as *A3*.

The statistical test indicated that the mean rate of goals for *A3* is significant greater than the mean for *A2* (*t-value* = -4.016)). Since *p-value* $= 0.001$ is lower than the *significance level* ($\alpha = 0.01$), the null hypothesis can be rejected. These results suggest that the new state representation (considering the barrier) does have a significant effect on the rate of goals of a DRL-based agent with a 99% confidence interval for the mean difference. Therefore, it is concluded that *A3* has a better performance over *A2* in the free kicks.

V. CONCLUSION AND FUTURE WORKS

This paper improves the gain in the learning process of DRL-based agents - in terms of training time and in-game performance - through ODTs in the dynamic scenario of free kicks in FIFA automatic players. The results confirmed that replacing a MobileNetV1 based ODT with a MobileNetV2 version allowed for enhancing the performance of both agents proposed herein. Further, the most complete of these agents, *A3*, which includes the barrier element in its state representation, proved to be the best version produced in this work. It is important to point out that the state representation explored by the present paper can be generalized to any other problem, as long as the object detection model is trained with an adequate database. As future works, the authors intend to study new DRL-based strategies to cope with the remaining modes of FIFA, such as: inserting the goalkeeper in the scene of the game and detecting appropriate dynamics of passe exchanging among the players. Besides, the authors are working in an alternative state representation based on ODT with greater interpretability of the generated data than the **High Level Feature Map** to consider only information referring to elements relevant to the game scenario defined as objects to be detected and, thus, to improve the results obtained here. The source code of this work is available at https://github.com/matheusprandini/FifaFKObjectDetection.

REFERENCES

[1] P. Henderson, R. Islam, P. Bachman, J. Pineau, D. Precup, and D. Meger, "Deep reinforcement learning that matters," in *AAAI Conference on Artificial Intelligence*, 2018, pp. 3207–3214.

[2] A. D. A. Voulodimos, N. Doulamis and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational Intelligence and Neuroscience*, 2018.

[3] Z. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–21, 2019.

[4] A. Venkatesh, "Object tracking in games using convolutional neural networks," Master's thesis, California Polytechnic State University, San Luis Obispo, 6 2018.

[5] C. Trivedi, "Using deep q-learning in fifa 18 to perfect the art of free-kicks," 2018. [Online]. Available: https://github.com/ChintanTrivedi/DeepGamingAI_FIFARL

[6] 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*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Springer International Publishing, 2016, pp. 21–37.

[7] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," 2017. [Online]. Available: https://arxiv.org/abs/1704.04861

[8] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 2018, pp. 4510–4520.

[9] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *ECCV*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer International Publishing, 2014, pp. 740–755.

[10] 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.

[11] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, 2017.

[12] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. The MIT Press, 2018.

[13] M. P. P. Faria, R. M. S. Julia, and L. B. P. Tomaz, "Deep active imitation learning in fifa free-kicks player platforms based on raw image and object detection state representations," 2019.

[14] A. M. Colman, *A dictionary of psychology / Andrew M. Colman*, 3rd ed. Oxford University Press Oxford ; New York, 2009.