Deep Reinforcement Learning for Intelligent Video Surveillance

Rizea Valentina
AI Multimedia Lab
National University of Science and Technology
POLITEHNICA Bucharest
Bucharest, Romania

Bogdan Ionescu
AI Multimedia Lab
National University of Science and Technology
POLITEHNICA Bucharest
Bucharest, Romania
bogdan.ionescu@upb.ro

***Abstract*—This PhD research aims to explore and develop the application of Deep Reinforcement Learning (DRL) for intelligent video surveillance. DRL combines advanced neural networks with reinforcement learning (RL) techniques to create an intelligent agent capable of identifying and responding to unusual or violent situations in a dynamic and complex environment, such as in video surveillance. In the context of this research, DRL leverages the strengths of Deep Learning (DL) to extract complex features from raw video data and RL algorithms to develop action strategies based on the rewards received. These technologies have the potential to significantly enhance the detection of anomalies, violent behaviors, and other risky situations in video surveillance systems.**

***Index Terms*—Deep Reinforcement Learning (DRL), Reinforcement Learning, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Prioritized Experience Replay (PER)**

I. Introduction

This PhD research aims to explore and develop the application of Deep Reinforcement Learning (DRL) for intelligent video surveillance. DRL combines advanced neural networks with reinforcement learning (RL) techniques to create an intelligent agent capable of identifying and responding to unusual or violent situations in a dynamic and complex environment, such as in video surveillance.

In the context of this research, DRL leverages the strengths of Deep Learning (DL) to extract complex features from raw video data and RL algorithms to develop action strategies based on the rewards received. These technologies have the potential to significantly enhance the detection of anomalies, violent behaviors, and other risky situations in video surveillance systems.

II. Activities completed so far

In this initial phase of the research, the focus has been on developing a deeper understanding of the theory behind reinforcement learning algorithms and their extensions in the DRL domain. The activities carried out thus far have concentrated on the following aspects:

A. Theoretical Foundations of Reinforcement Learning (RL)

The first chapter of the report elaborated on the fundamentals of reinforcement learning, including the formal structure of the Markov Decision Process (MDP) and the use of Bellman equations. These concepts are essential because they define how an agent can make optimal decisions in stochastic environments to maximize cumulative reward. The following topics were analyzed:

- State-value function (V) and action-value function (Q), which assess the quality of a state or action in a specific context.
- The exploration-exploitation dilemma, crucial for RL algorithms, as they need to balance between exploring new actions and exploiting the known ones.

Among the algorithms studied were Q-learning and SARSA, which are used for efficiently estimating value functions. Q-learning is an off-policy algorithm, meaning that Q-values can be updated independently of the current policy, providing flexibility in exploring optimal strategies. In contrast, SARSA is an on-policy algorithm that updates Q-values based on the current policy, reflecting more accurately the agent's actual behavior at the time.[1]

B. Integration of Neural Networks and the Extension to DRL

To address the complexity of video data, we extended the concept of reinforcement learning by integrating convolutional neural networks (CNNs), which are used in DRL. In the documentation created, we analyzed in detail several DRL algorithms and architectures, including:

- Deep Q-Network (DQN), developed by DeepMind, which uses neural networks to learn the Q-value of possible actions in a given video state.
- Double DQN, an improvement over DQN that reduces the risk of Q-value overestimation.
- Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM), which can store and process past information, making it possible to capture temporal relationships in video data.

This chapter also covered the exploration - exploitation dilemma and the use of techniques like noisy nets and prioritized experience replay (PER), which allow for more diverse exploration of actions. Furthermore, we identified the importance of Rainbow DQN, which combines several DQN improvements, making it suitable for modeling complex video surveillance scenarios.[2]

1st Workshop on Artificial Intelligence for Multimedia,
AI Multimedia Lab, CAMPUS Research Institute, POLITEHNICA Bucharest,
Bucharest, Romania, 8 November, 2024

C. Challenges of DRL in Video Surveillance

We detailed the unique challenges of DRL in video surveillance, such as:

Credit assignment problem – DRL faces difficulties in learning from rare or delayed rewards, common in anomaly detection where critical events are sparse.

Partial observability – The DRL agent does not have access to complete information at all times; using LSTM-based recurrent layers allows the agent to maintain a longer context to compensate for the lack of information.[3]

Slow learning – Most DRL algorithms require large amounts of data to learn robust policies, which is a limitation in video surveillance systems, where processing speed and response time are crucial.[4]

III. Future Research Directions

The next phase of the research will involve both practical testing of the DRL algorithms on real-world video datasets and the development of optimized architectures for detecting violence and other anomalies. Specifically, future activities will include:

A. Testing DRL Algorithms on Specific Datasets

An important objective is to test the studied algorithms on real video surveillance datasets. We will start by testing variants of DQN and Actor-Critic to assess:

- Accuracy and precision in detecting unusual events.
- Generalization to new environments and scenarios.
- Response time, which is essential for real-time detection.

B. Development of an Optimized Hybrid DRL Architecture

We will integrate CNNs and RNNs (especially LSTMs) into a hybrid DRL architecture to effectively extract and interpret both spatial and temporal features from video scenarios, improving the precision of detecting violent behaviors. The goal of this hybrid architecture is to improve:

- Response to rapid movements and detection of aggressive behaviors.
- Stability and memory capacity of the algorithm, thus reducing detection errors.

C. Implementation of Optimization and Regularization Techniques

To reduce computational costs and enhance efficiency, we will implement advanced techniques such as:

- Noisy nets – for more diverse and efficient exploration of the action space, reducing the risk of being trapped in a local maximum.
- Prioritized Experience Replay (PER) – which will allow the algorithm to prioritize relevant experiences, speeding up the learning process and improving the robustness of event detection.

D. Performance Evaluation in Detecting Violent Behaviors

Evaluation will be performed using datasets containing various surveillance scenarios to assess performance in:

- Detecting violent behaviors and rapid movements.
- Adaptability to environmental changes, such as different lighting conditions and unexpected object movements.
- Scalability in complex environments, with the aim of creating a DRL model capable of real-time operation.

Conclusions and Perspectives

So far, the research has established a solid theoretical foundation for applying DRL in intelligent video surveillance, providing a deep understanding of reinforcement learning algorithms and their extensions through deep neural networks. The study of algorithms such as DQN, Double DQN, Actor-Critic, and Rainbow has been essential for developing efficient solutions for recognizing abnormal behaviors in video surveillance contexts.

Future directions will focus on the practical implementation of DRL algorithms, testing them in real-world environments and on specific video datasets. The aim of these tests is to assess and improve the ability of the algorithms to recognize violent behaviors and other risky situations, with a focus on accuracy, processing speed, and adaptability to various scenarios.

By developing an optimized hybrid architecture and applying advanced regularization techniques, this research has the potential to contribute significantly to the field of video surveillance, making surveillance systems more effective, autonomous, and capable of operating in real-time.

In the long term, the project may open new research avenues in DRL, particularly in the use of artificial intelligence for autonomous surveillance, proactive detection, and automated interventions, with a positive impact on security in public and private spaces.

References

[1] Sutton Barto Book: Reinforcement Learning: An Introduction."
[2] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, May 2015
[3] V. Mnih et al., "Human-level control through deep reinforcement learning", Nature, vol. 518, Feb. 2015
[4] A. D. Tijsma, M. M. Drugan, M. A. Wiering, "Comparing exploration strategies for Q-learning in random stochastic mazes". In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 2016, pp. 1–8