RADemics

Transfer Learning and Domain Adaptation Techniques in Deep Reinforcement Learning for Cross-Domain Tasks

J. Jayaganesh, Vanathi M
GOVERNMENT ARTS AND SCIENCE COLLEGE,
Sathyabama Institute of Science and Technology

6. Transfer Learning and Domain Adaptation Techniques in Deep Reinforcement Learning for Cross-Domain Tasks

[1]J. Jayaganesh, Assistant Professor, Department of Computer Science, Government Arts and Science College, Perumbakkam, Chennai-600131, Tamil Nadu, India, everjays@gmail.com

[2]Vanathi M, Assistant Professor, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India, mercyvanu@gmail.com

Abstract

This book chapter explores the emerging synergy between Transfer Learning and Domain Adaptation techniques in Deep Reinforcement Learning (DRL), with a focus on their application to cross-domain tasks. As DRL continues to revolutionize various fields, the challenge of transferring learned knowledge across different domains remains a critical obstacle. This chapter delves into the theoretical foundations of transfer learning and domain adaptation, highlighting their significance in enhancing the generalization capabilities of DRL agents. It presents advanced methodologies for applying these techniques in complex real-world applications, including robotics, autonomous driving, and video game AI. The chapter examines hybrid models that integrate both approaches to improve performance in unfamiliar environments, offering insights into their potential and limitations. By addressing existing challenges and exploring future research directions, this chapter aims to provide a comprehensive understanding of how these techniques can advance DRL applications in dynamic, cross-domain settings.

Keywords:

Transfer Learning, Domain Adaptation, Deep Reinforcement Learning, Cross-Domain Tasks, Hybrid Models, Generalization.

Introduction

Deep Reinforcement Learning (DRL) has demonstrated impressive capabilities in training agents to perform complex tasks, ranging from robotic control to playing strategic games [1]. However, one of the major hurdles that DRL faces was the challenge of transferring knowledge learned in one domain to another, often drastically different domain [2,3]. Transfer Learning (TL) addresses this by allowing models to leverage previously learned representations and knowledge from one task or environment and apply them to a new, related task [4]. Similarly, Domain Adaptation (DA) focuses on overcoming differences between the source and target domains, specifically adjusting for distribution shifts between the two [5,6]. These two techniques are critical in enabling DRL systems to operate effectively in real-world applications where training data was often limited, and environments can differ from the ones seen during training [7,8]. The ability to apply transfer learning and domain adaptation in DRL can significantly reduce training time, improve efficiency, and enhance the performance of agents in complex, dynamic environments [9-12].

While the theoretical foundations of transfer learning and domain adaptation are well established, their application in cross-domain tasks within DRL remains a challenging area of research [13-15]. Real-world domains often exhibit significant differences in state spaces, action spaces, and reward structures, making it difficult for DRL agents to generalize learned policies [16]. The variability between source and target domains, such as environmental dynamics or task-specific characteristics, complicates the direct transfer of knowledge [17]. For instance, a DRL agent trained in a simulated environment struggle when deployed in a real-world setting due to the domain gap in visual inputs, sensor data, or interaction patterns [18]. Domain adaptation techniques, particularly adversarial training and domain randomization, have been developed to mitigate these gaps by making the learned policy more invariant to domain-specific variations [19,20]. However, the full integration of transfer learning and domain adaptation was still an ongoing area of research, requiring novel methodologies to address challenges such as negative transfer, where transferring knowledgeactually harm performance [21,22].

Recent advancements in DRL have seen the rise of hybrid models that combine both transfer learning and domain adaptation to overcome challenges in cross-domain task learning. These hybrid approaches aim to simultaneously transfer knowledge between domains while adapting the models to specific environmental conditions [23]. By integrating methods such as multi-task learning, where a model was trained on multiple related tasks, hybrid models are capable of achieving better generalization across domains [24,25]. The hybrid approach also leverages fine-tuning techniques, which involve adjusting pretrained models to better fit the characteristics of the target domain. A critical aspect of hybrid models was their ability to enhance the learning process in environments where labeled data was scarce or difficult to acquire. These models have shown promise in a variety of applications, such as autonomous driving, where an agent must adapt its learned behaviors from simulation to real-world driving scenarios. In robotics, hybrid models can enable agents to transfer and adapt manipulation strategies across different tasks and environments, such as from one type of object to another or from one robot to another.