Article

On Renormalization Group Based Deep Q-Network

Guman Garayev [*] and Azar Alili [*]

Article

On Renormalization Group Based Deep Q-Network

Guman B. Garayev [1,2] and Azar M. Alili [3]

[1]   Pasha Holding IT Excellence Centre, Azerbaijan
[2]   Institute of Molecular Biology and Biotechnologies MSERA, Azerbaijan
[3]   Pasha Development, Azerbaijan
[*]   Correspondence: guman.garayev@gmail.com

Abstract: In This paper we introduce the integration of Renormalization Group (RG) methods with Deep Q-Networks (DQNs) to improve reinforcement learning in high-dimensional state spaces. RG methods provide multi-scale analysis, enhancing state representation, learning stability, and exploration. The proposed RG-DQN algorithm uses hierarchical Q-value estimation and multi-scale representations, demonstrating superior performance on synthetic genomic data compared to traditional DQNs.

Keywords: DQN; renormalization group; AI; loss functions

1. Introduction

Deep Q-Learning (DQL) represents a cornerstone of modern reinforcement learning, integrating the principles of Q-learning with the powerful function approximation capabilities of deep neural networks [1]. The foundational idea of Q-learning is to learn a policy that maximizes the expected cumulative reward by approximating the optimal action-value function, which represents the maximum expected return of taking an action in a given state and following the optimal policy thereafter [6].

Deep Q-Networks (DQNs) extend Q-learning by using deep neural networks to estimate the action-value function. This allows DQNs to handle high-dimensional state spaces effectively. Key innovations in DQNs include experience replay, which breaks the correlation between consecutive experiences by storing and randomly sampling past transitions, and the use of a target network to stabilize training by providing consistent target values [2]. These innovations have enabled DQNs to achieve remarkable success in complex environments, notably attaining human-level performance in various Atari games [1].

The motivation for integrating RG methods into DQNs arises from the complementary strengths of these approaches. Traditional DQNs, despite their success, face challenges in efficiently exploring high-dimensional state spaces and capturing hierarchical structures within the data [3]. RG methods, with their multi-scale analysis capabilities, offer a promising avenue to address these limitations by providing a framework to understand and manipulate the system at various scales [3]. By applying RG principles, we aim to enhance state representation, improve learning stability, and optimize exploration.

Multi-scale representations can capture features at different levels of granularity, providing a richer and more structured view of the state space [4]. Hierarchical abstraction through RG can stabilize learning by focusing on significant features at each scale, reducing noise and variability in training [5,7,8]. Moreover, RG-inspired exploration strategies can efficiently navigate complex environments by leveraging the hierarchical structure of the state space [3].

The integration of Renormalization Group methods with Deep Q-Networks offers a promising enhancement to the traditional DQN framework. By leveraging the multi-scale analysis capabilities of RG, we can address key challenges in DQNs, such as high-dimensional state spaces and learning stability . This paper explores this novel integration, presenting both theoretical foundations and empirical results to demonstrate the potential benefits of this approach.

2. Methodology

Renormalization Group (RG) methods provide a powerful framework for understanding systems at multiple scales. Originating from theoretical physics, RG methods involve iteratively simplifying a

system by integrating out short-range fluctuations, leaving a model that captures long-range behavior [3]. This multi-scale analysis can be particularly useful in reinforcement learning environments characterized by high-dimensional state spaces and complex dynamics. By leveraging RG methods, we aim to enhance state representation, improve learning stability, and optimize exploration strategies in Deep Q-Networks (DQNs).

The first step in integrating RG methods with DQNs is to create a multi-scale representation of the state space. This involves decomposing the state into different scales to capture features at various levels of granularity. Techniques such as wavelet transforms can be employed to achieve this decomposition, resulting in a set of states $S_0, S_1, \ldots, S_n$, where each S_k represents the state at scale k. Mathematically, this can be represented as [4]:

$$S_k = \mathcal{T}(S_{k-1})$$

where $\mathcal{T}$ is a transformation function that aggregates or simplifies the state representation at scale $k-1$ to obtain the state at scale k.

2.1. Hierarchical Q-Value Estimation

With multi-scale state representations, we estimate Q-values at each scale. The Q-values for each scale k are denoted as $Q_k(s, a; \theta_k)$. These Q-values are updated using the Bellman equation adapted for each scale [6]:

$$Q_k(s, a; \theta_k) \leftarrow Q_k(s, a; \theta_k) + \alpha \left[r + \gamma \max_{a'} Q_k(s', a'; \theta_k^-) - Q_k(s, a; \theta_k) \right]$$

Where:

- α is the learning rate.
- γ is the discount factor.
- s' is the next state.
- θ_k^- represents the parameters of the target network at scale k.

2.2. Aggregated Q-Values

To integrate the Q-values from different scales, we introduce a weighting function w_k that assigns appropriate weights to the Q-values at each scale. The overall Q-value is a weighted sum of the multi-scale Q-values:

$$Q(s, a; \theta) = \sum_{k=0}^{n} w_k Q_k(s, a; \theta_k)$$

Where:

- w_k are the weights assigned to each scale k, with $\sum_{k=0}^{n} w_k = 1$ [4].

2.3. Multi-Scale Loss Function

The loss function for each scale k is defined as the mean squared error between the target Q-value y_j and the predicted Q-value $Q_k(s_j, a_j; \theta_k)$ [2]:

$$\mathcal{L}_k(\theta_k) = \mathbb{E}\left[\left(y_j - Q_k(s_j, a_j; \theta_k)\right)^2 \right]$$

Where:

- $y_j = r_j + \gamma \max_{a'} Q_k(s'_j, a'; \theta_k^-)$

The combined loss function across all scales is:

doi:10.20944/preprints202407.0953.v1

3 of 4

$$L(\theta) = \sum_{k=0}^{n} w_k \mathcal{L}_k(\theta_k)$$

To integrate RG methods with DQN, we propose a new training algorithm that begins with initialization steps. Initially, we set up the replay buffer D with a capacity of N. Following this, we initialize the Q-networks $Q_k(s, a; \theta_k)$ for each scale k using random weights θ_k. We also initialize the target networks $Q_k(s, a; \theta_k^-)$ with weights set as $\theta_k^- = \theta_k$. For each episode, we start by initializing the state s. For each step of the training process, the following actions are performed:

1. Select an action a using an epsilon-greedy policy based on $Q(s, a; \theta)$.
2. Execute the selected action a and observe the reward r and the next state s'.
3. Store the experience (s, a, r, s') in the replay buffer D.
4. Sample a mini-batch of experiences (s_j, a_j, r_j, s'_j) from D.
5. Compute the target y_j for each scale k as:

$$y_j = r_j + \gamma \max_{a'} Q_k(s'_j, a'; \theta_k^-)$$

6. Compute the loss $\mathcal{L}_k(\theta_k)$ for each scale k as:

$$\mathcal{L}_k(\theta_k) = \mathbb{E}\left[\left(y_j - Q_k(s_j, a_j; \theta_k) \right)^2 \right]$$

7. Perform gradient descent on the combined loss function $L(\theta)$:

$$\theta_k \leftarrow \theta_k - \alpha \nabla_{\theta_k} L(\theta)$$

8. Update the state $s \leftarrow s'$.

By following these steps, the RG-enhanced DQN algorithm leverages multi-scale representations and hierarchical Q-value estimation, resulting in improved learning stability and efficiency in high-dimensional state spaces [3]. This integration of RG methods provides a robust framework for enhancing the performance of traditional DQNs.

3. Simulation and Results

To test the effectiveness of the RG-Enhanced DQN compared to the standard DQN, we generated a synthetic genomic dataset. The data comprises sequences of nucleotide bases, each represented by integers ranging from 0 to 3. We introduced structural variations at different scales to create a challenging classification task. The dataset was divided into training and testing sets.We trained both the standard DQN and the RG-Enhanced DQN on the synthetic genomic data. The standard DQN utilizes the raw input sequences, while the RG-Enhanced DQN applies a wavelet transform to extract multi-scale features from the sequences. Both models were trained for five iterations to ensure quick execution.

The evaluation was performed on the testing set, with the accuracy of both models recorded.The results of the simulation showed a significant improvement in the accuracy of the RG-Enhanced DQN compared to the standard DQN. The standard DQN achieved an accuracy of 0.50, while the RG-Enhanced DQN achieved an accuracy of 0.72, demonstrating the superior performance of the RG-Enhanced approach in handling multi-scale genomic data.

4. Conclusion

The significant improvement in accuracy observed with the RG-Enhanced DQN can be attributed to its ability to capture multi-scale features in the genomic data. The wavelet transform used in the RG-Enhanced DQN decomposes the input sequences into different scales, allowing the model to learn

both fine-grained and coarse-grained patterns. This multi-scale representation is particularly effective for genomic data, where structural variations occur at various levels.

In conclusion, the application of RG methods to enhance DQNs provides a robust approach to improving classification accuracy in tasks involving complex, multi-scale data, such as genomic sequences with structural variations.

References

1.   Mnih, V., et al. (2015). "Human-level control through deep reinforcement learning." *Nature*.
2.   Hasselt, H. van, et al. (2016). "Deep reinforcement learning with double Q-learning." *AAAI*.
3.   Wilson, K.G., & Kogut, J. (1974). "The renormalization group and the ϵ expansion." *Physics Reports*.
4.   Mallat, S. (1989). "A theory for multiresolution signal decomposition: The wavelet representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
5.   Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM.
6.   Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.
7.   Dietterich, T.G. (2000). "Hierarchical reinforcement learning with the MAXQ value function decomposition." *Journal of Artificial Intelligence Research*.
8.   Hinton, G.E., et al. (2006). "A fast learning algorithm for deep belief nets." *Neural Computation*.