

A Comparative Analysis of Deep Reinforcement Learning Approaches in Symbolic Optimization Tasks: The Case of DQN, QT-Opt and Samuel

Cem Özkurt¹, Ahmet Kutey Küçükler², Murat Karslıoğlu³ and Ruveyda Nur Özdemir⁴

*Department of Computer Engineering, Sakarya University of Applied Sciences, Sakarya, Türkiye.

ABSTRACT This study aims to comparatively analyze the performance of three reinforcement learning algorithms-DQN, QT-Opt, and Samuel's checkers algorithm-on the symbolic matrix multiplication task. The experiments were conducted using a customized simulation environment, MatrixMultiplyDiscoveryEnv, where each agent generates outer product-based symbolic actions to perform matrix multiplication with minimal error and computational cost. The reward function incorporates the Frobenius norm, operation count, and symbolic complexity. Based on 50,000 episodes, the QT-Opt algorithm demonstrated a highly stable reward profile, maintaining reward values close to zero throughout training. Samuel's algorithm showed rapid early learning, improving from -300 to around -100, but exhibited fluctuations in the later stages. In contrast, DQN's reward varied drastically, occasionally falling below -3000, indicating instability and sensitivity to environmental uncertainty. Regarding matrix error (Frobenius norm), Samuel's algorithm minimized its error to nearly zero in early training and maintained this performance. QT-Opt also performed well but showed occasional spikes in error. In terms of operation cost, QT-Opt consistently operated within 50-100 units, showing the highest efficiency. Samuel started with costs near 300, but reduced them gradually, converging towards QT-Opt's performance. DQN, however, showed wide and erratic cost distributions. In conclusion, QT-Opt achieved the most stable and efficient learning, particularly in continuous action domains. This paper provides a unique perspective by comparing classical and modern reinforcement learning methods within a unified experimental framework, highlighting both their historical significance and practical performance.

KEYWORDS

Reinforcement learning
QT-Opt
Symbolic optimization

INTRODUCTION

Reinforcement Learning (RL) is a powerful decision-making paradigm that enables an agent to discover optimal action policies based on reward signals through interaction with its environment. In this framework, three prominent methods in the literature, Deep Q-Network (DQN), QT-Opt, and Samuel's learning systems, are notable for their historical development, structural architecture, and application areas.

DQN was introduced by Mnih *et al.* (2015) and is one of the first deep reinforcement learning models that can make successful decisions in high-dimensional state spaces by integrating an experience repetition buffer and a target network structure. The theoretical foundation of DQN is later analyzed by Fan *et al.* (2020), who present the algorithmic convergence behavior and statistical error structure of this method. The contribution of mechanisms

such as experience repetition and target network to success is theoretically justified. Furthermore, extended variants such as DQfD (Deep Q-learning from Demonstrations) developed by Hester *et al.* (2018) accelerate the learning process with human demonstrations, increasing the usability of DQN for real-world applications.

QT-Opt, developed by Kalashnikov *et al.* (2018), is a scalable and distributed deep reinforcement learning algorithm that achieves high success in continuous control problems such as robotic grasping based on visual inputs. In particular, based on closed-loop control logic, it allows the robot to dynamically update its grasping strategy with visual feedback. It achieved %96 grasping success with RGB camera data only. QT-Opt's high-parameter Q-function is characterized by its capacity to learn from MAP inference and real-world experiences.

From a historical perspective, Samuel (1959) checkers playing program is more than just a game analysis. It is the theoretical foundation for modern reinforcement learning approaches by incorporating value function estimation, heuristic search and self-play. Samuel's work demonstrates that a computer program can improve itself through experience, not just rules, and was among the first to use the term "machine learning".

This paper aims to comparatively analyze the three aforementioned methods DQN, QT-Opt and Samuel's learner systems in a single experimental and conceptual framework. DQN and QT-Opt,

Manuscript received: 22 November 2025,

Revised: 21 December 2025,

Accepted: 14 January 2026.

¹cemozkurt@subu.edu.tr (Corresponding author).

²23010903094@subu.edu.tr

³24010903047@subu.edu.tr

⁴24010903035@subu.edu.tr

as contemporary deep learning-based reinforcement learning algorithms, are successful in high-dimensional and continuous action spaces. Samuel's method provides a historical reference point with its algorithmic simplicity and heuristic foundations.

In recent years, there has been a growing body of work in the literature analyzing DQN and QT-Opt either individually or in pairwise comparisons (Hester *et al.* 2018; Kalashnikov *et al.* 2018). However, comprehensive triple comparisons in which these methods are evaluated together from a historical and structural perspective are almost non-existent. In this context, our study aims to bring together classical and modern reinforcement learning approaches on the same ground, both to make the historical evolution visible and to empirically reveal the performance differences between the methods.

THEORETICAL REFERENCE

Reinforcement learning has emerged as an important paradigm for optimizing decision-making processes based on interaction with the environment in robotic systems. Over time, this field has evolved from classical algorithms to deep learning-based artificial intelligence systems. One of the most fundamental milestones in this evolution is Arthur Samuel's classical approach based on reinforcement learning principles. While Samuel's method offered a primitive learning process where decisions were shaped based on experience, today this structure has merged with deep structures capable of learning in multidimensional state and action spaces (Liao *et al.* 2024). Accordingly, modern algorithms built on Samuel's methodological foundations have reached the capacity to learn complex tasks with the contribution of big data and visual signals (Gao 2024).

The Deep Q-Network (DQN) algorithm is one of the revolutionary developments in the field of reinforcement learning. It is widely preferred in critical applications such as autonomous navigation and collision avoidance (AlMahamid and Grolinger 2025; Chen *et al.* 2024; de Sousa Bezerra *et al.* 2023), especially because it enables decision making under limited sensor data and high uncertainty environments. The advantage of DQN is that it can learn in continuous state spaces by modeling the Q-value function with a deep neural network. However, the classical DQN structure has certain limitations. To overcome these problems, variants such as Double DQN, Dueling DQN and Noisy DQN have been developed. These variants have produced effective results, especially in high-risk environments such as autonomous underwater vehicles and confined space robots (Al-Hamadani *et al.* 2024; Gao *et al.* 2023; Chen *et al.* 2024).

The integration of attention mechanisms into the DQN structure has further strengthened the mission success by increasing the environmental awareness of the system. Thanks to this integration, more responsive and adaptive solutions have been developed for collision avoidance missions in UAV systems (Al-Hamadani *et al.* 2024). DQN has also been shown to offer a satisfactory alternative in terms of robustness and information sharing in challenging tasks such as decentralized multi-robot control (Wu and Suh 2024). Thus, it provides a flexible basis for both the interaction of individual robots with the environment and the coordination of multi-agent systems.

The QT-Opt algorithm is another prominent approach, especially in tasks such as robotic arm manipulation with high-dimensional visual data. By combining value-based methods with policy optimization, QT-Opt makes the data collection process both safe and cost-effective thanks to its structure suitable for offline learning (Zhang *et al.* 2025). The integration of this algorithm

with offline learning makes it easier for the robot to adapt to new situations by generalizing from previously observed data [16]. Especially in visual manipulation problems, high-performance policy learning can be achieved without directly interacting with the environment. In addition, recent work on behavior correction policies extends the safety boundaries of QT-Opt and ensures stable learning even in unsupervised scenarios (Dong *et al.* 2024).

The evolutionary impact of Samuel's classical method is not only of historical importance, but the principles derived from it are also evident in the design of modern artificial intelligence systems. For example, the integration of representation-based learning and DQN enables rapid adaptation to both static and dynamic conditions in tasks such as robotic arm control (Gao 2024). Such integrated structures not only improve task performance but also offer sustainable solutions for general AI architectures.

Multi-robot systems are another important application area of reinforcement learning. In these systems, real-time task sharing and coordination of agents with different abilities becomes quite complex with traditional methods, while multi-agent deep reinforcement learning approaches make this complexity manageable (Gao *et al.* 2023). In particular, the coordinated play of heterogeneous roles enables the system to perform adaptive task assignment and promotes collaborative learning (Pal *et al.* 2025). In addition, automatic task generation in task-diverse scenarios and learning reusable strategies for these tasks increase the scalability and flexibility of robotic systems (Jansson *et al.* 2024; Mao *et al.* 2023).

There are many studies in the literature where DQN and QT-Opt algorithms have achieved high performance on various robotic tasks. Samuel's method also provides a theoretical underpinning for these advanced algorithms, but comparative analysis of these methods in terms of performance, safety, sample efficiency and multi-task generalization within the same study is very limited. This paper presents a holistic view of the strengths and weaknesses of DQN, QT-Opt and Samuel methods from both theoretical and practical perspectives, especially in multi-tasking and high-risk robotic scenarios. Thus, it aims to provide a guiding reference for both academic and industrial applications.

MATERIALS AND METHODS

In this work, we use Deep Q-Network (DQN), QT-Opt and Arthur Samuel's checkers algorithm for reinforcement learning-based solutions to the symbolic matrix multiplication task. The experimental environment is based on the MatrixMultiplyDiscoveryEnv simulation environment designed to optimize symbolic processing efficiency. This environment vectorizes two input matrices and presents them to agents, which generate solutions through outer product-based symbolic actions. The reward function has a multi-component structure penalizing the Frobenius norm, number of operations and symbolic complexity. Model training is supported by classical and quantum critical architectures, and the algorithms are comparatively tested under the same conditions. DQN performed better in discrete action domains, while QT-Opt performed better in continuous action domains. Arthur Samuel's algorithm is included as a reference to establish historical context.

Data Set and Environment Definition

The experimental data set and the environment used in this study are based on an open-source reinforcement learning framework called Quantum Matrix RL. It is built on MatrixMultiplyDiscoveryEnv, a specialized simulation environment for the discovery of symbolic algorithms for basic linear algebra operations such

as matrix multiplication. The environment combines two input matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ in a flattened form to generate an observation vector. The agent generates symbolic actions based on this observation. The agent's goal is to approximate the product $C = AB$ by multiplying the given matrices A and B with minimum number of operations and low symbolic cost. Actions defined in this environment represent symbolic operations such as low-rank outer product. Each action updates the agent's current estimate C_t . The reward function used as a feedback mechanism, $\|C_t - AB\|_2$ includes both an error metric based on the Frobenius norm and penalties for number of operations and symbolic complexity. Through this structure, the agent is incentivized to optimize not only accuracy but also computational efficiency.

The model training process is supported by classical or quantum-based critic structures. The data used in this study consists of the outputs of the models trained in simulation and allows the comparison of each algorithm under the same experimental conditions. Thus, the learning behavior of different reinforcement learning algorithms on symbolic optimization tasks can be objectively analyzed.

The DQN and QT-Opt algorithms are chosen for comparative analysis because they have different action space types. DQN is a classical method that is efficient in discrete action spaces and provides a strong foundation for discrete tasks such as symbolic matrix multiplication. On the other hand, the QT-Opt algorithm enables the exploration of more complex strategies thanks to its direct optimization capability in continuous action domains. These two methods provide meaningful counterexamples to evaluate the performance of different reinforcement learning paradigms in a common framework.

Furthermore, Arthur Samuel's checkers algorithm is included in the study as a historical reference. This method sheds light on the theoretical foundations of modern algorithms, especially with its value function-based learning approach, and provides an important perspective to contextualize the process of methodological evolution.

Deep Q-Network (DQN)

DQN combines classical Q-learning with deep neural networks to enable efficient learning in high-dimensional state spaces. The expected total reward (Q-value) of each state-action pair is estimated through a neural network. The learning process is based on minimizing the difference between the target Q-values and the predicted values based on the Bellman optimality equation.

$$L(\theta) = \mathbb{E}_{(s,a,r,s') \sim \mathcal{D}} \left[r + \gamma \max_{a'} Q_{\theta'}(s', a') - Q_{\theta}(s, a) \right]^2 \quad (1)$$

In this formula, s represents the current state, a represents the action taken, r represents the immediate reward, s' represents the next state, a' represents the next possible actions, θ represents the parameters of the main network, θ' represents the fixed parameters of the target network and γ represents the discount coefficient chosen between 0 and 1. Experience replay memory and goal network techniques increase the stability of learning.

QT-Opt

QT-Opt performs policy learning through direct optimization in continuous action spaces. This method offers a more flexible and efficient solution space for continuous control problems that classical DQN has difficulty solving. Especially preferred in robotics

applications, this method allows for the development of finer discriminative action strategies. Unlike DQN, in QT-Opt the optimal action is determined by choosing an argument that maximizes the Q-value.

$$a^* = \arg \max_a Q(s, a; \theta) \quad (2)$$

Where s is the current state, a is the feasible action, $Q(s, a; \theta)$ is the function that estimates the expected reward of the action in that state, and a^* is the action that maximizes this value. The QT-Opt algorithm also learns using Bellman-based target Q-values. This structure allows future reward expectations to be linked to present decisions.

$$y = r + \gamma \max_{a'} Q_{\theta'}(s', a'; \theta') \quad (3)$$

Where y is the target Q-value, r is the immediate reward, s' is the next state, a' is the possible next actions, γ is the discount coefficient, and $Q(s', a'; \theta')$ is the Q-value estimated with the target network parameters. QT-Opt balances exploration and exploitation in action selection with strategies such as ϵ -greedy.

Arthur Samuel's Checkers Algorithm

Samuel's algorithm is an early example of value function-based learning. This approach was one of the first systems to allow a computer to learn its own strategies by playing games. This model laid the foundations for traditional AI applications and played a key role in the historical development of modern reinforcement learning algorithms. The system evaluates each game state with a weighted sum of certain features.

$$V(s) = \sum_i w_i x_i(s) \quad (4)$$

In this formula, $x_i(s)$ defines the i th attribute of the state (e.g. number of pieces, position, etc.), w_i is the weight assigned to this attribute, and $V(s)$ is the total score of the state. The learning process is performed by updating these weights based on game experience. This structure makes it possible to make sense of previously unseen situations through generalizable evaluation functions.

RESULTS AND DISCUSSIONS

In this section, the performance of DQN, QT-Opt and SAMUEL algorithms over 50,000 episodes is analyzed comparatively according to three main criteria: reward, matrix error (Frobenius norm) and operational cost. The findings are analyzed in detail with graphs.

Reward Evaluation

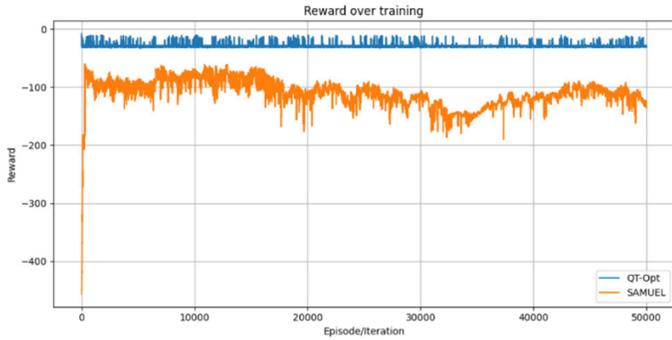


Figure 1 QT-Opt and SAMUEL Reward Over Training

As can be seen in Figure 1, the QT-Opt algorithm has a very stable reward profile throughout the training process and these values are generally close to zero. This shows that QT-Opt makes very stable and reliable decisions in the applied environment. SAMUEL, on the other hand, rises rapidly from negative values in the early stages and then fluctuates around -100. This shows that SAMUEL reaches a certain performance limit in the learning process, but it is not as stable as QT-Opt.

Matrix Error (Frobenius Norm)

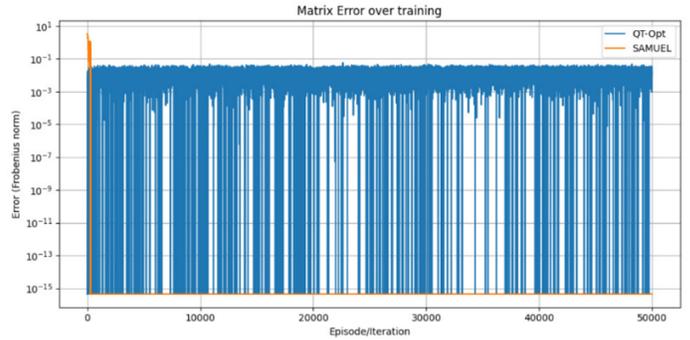


Figure 3 QT-Opt and SAMUEL Matrix Error Over Training

As seen in Figure 3, when evaluated in terms of matrix error values, the SAMUEL algorithm has demonstrated a remarkable success. It minimized its error at the beginning of the training process and reached almost zero levels and maintained this stable structure throughout the training period. Although QT-Opt tries to keep its error at low levels with relatively high frequency fluctuations, sudden jumps in these values indicate instability in the learning process.

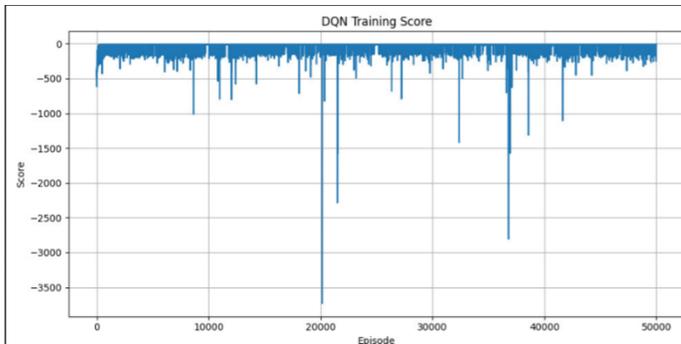


Figure 2 DQN Training Score

As seen in Figure 2, when the DQN algorithm is examined, it is seen that the performance is quite variable and there are serious performance losses that sometimes fall below -3000. This high variance shows that the model has difficulty learning a stable policy and is not robust against environmental uncertainties.

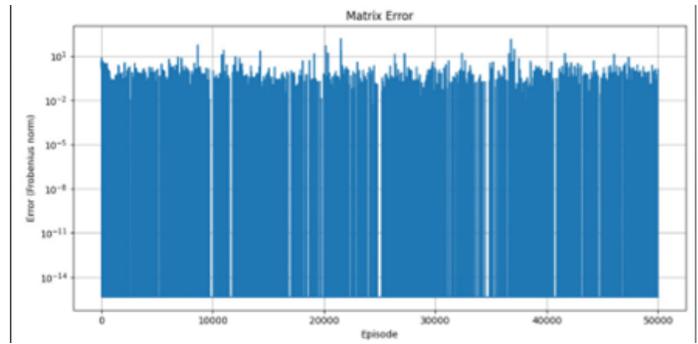


Figure 4 DQN Matrix Error

As seen in Figure 4, it is observed that in the DQN algorithm, the errors are constantly changing and reach very high values in certain sections. This shows that the algorithm cannot fully provide numerical stability in weight updates and therefore does not converge stably.

Operation Cost Analysis

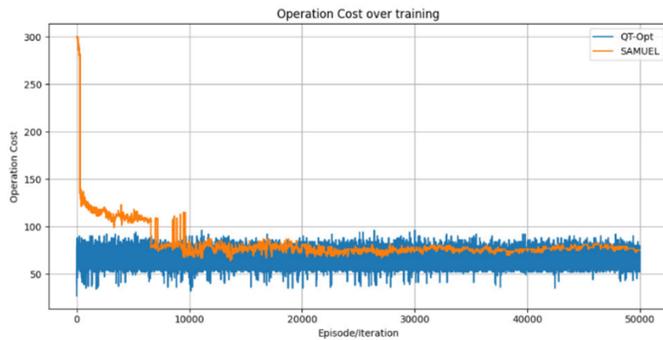


Figure 5 QT-Opt and SAMUEL Operation Cost Over Training

As seen in Figure 5, the QT-Opt algorithm exhibits a very successful performance in terms of operational costs. During most of the training, costs fluctuated between 50 and 100 units, and this fluctuation remained low amplitude. Although the SAMUEL algorithm started with high costs at the level of 300 at the beginning, it managed to reach a performance close to QT-Opt by systematically reducing its costs in the following sections. This reveals SAMUEL's long-term learning success and adaptability.

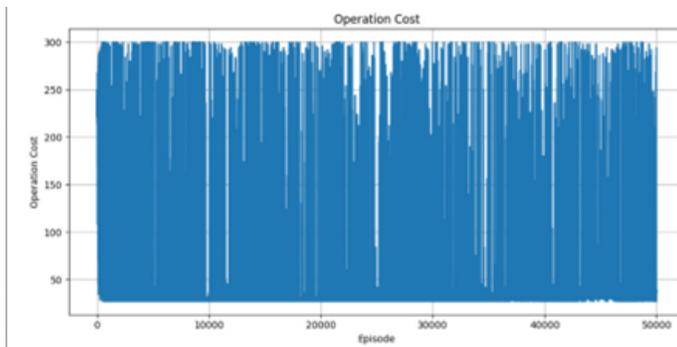


Figure 6 DQN Operation Cost

As seen in Figure 6, the operating costs of the DQN algorithm show an extremely irregular and wide distribution. Serious jumps and fluctuations were observed during the training period; this shows that the model could not achieve permanent and sustainable learning in terms of cost reduction.

In this study, the effect of the Cross Entropy method on three different reinforcement learning algorithms (DQN, QT-Opt and Samuel) was comparatively analyzed. The findings revealed that the QT-Opt algorithm, in particular, outperformed the other methods in terms of both reward stability and processing efficiency. This result is consistent with the findings of Kalashnikov et al., which showed that the closed-loop control structure proposed for QT-Opt can increase comprehension success up to 96% in high-dimensional visual tasks (Kalashnikov et al. 2018).

Another remarkable advantage of QT-Opt is its ability to achieve high success with offline learning. In the study of Wei et al., it was reported that significant improvements were achieved in learning processes by integrating advanced derivatives of DQN (NoisyNet, PER, Distributional DQN) into the QT-Opt structure (Wei et al. 2023). In our study, Frobenius norm and operational cost metrics showed that QT-Opt works more balanced and with low

errors. It is reported in the literature that advanced optimization contributions, such as the QT-Opt-PSO integration proposed by Zhang et al. (2025), similarly enhance error reduction.

On the other hand, the performance of the DQN algorithm is lower than QT-Opt due to the high variance reward profile and fluctuating matrix errors. This situation confirms the basic requirement of the improved DQN algorithms proposed by Chen et al. for indoor obstacle avoidance problems. Also, the Agile DQN architecture proposed by AlMahamid and Grolinger improved the performance of DQN under partial observations with attention mechanisms and timing-sensitive Q-estimation strategies (Gao 2024; Chen et al. 2024). Such improved architectures are promising to overcome the limitations of classical DQN, especially in scenarios with high visual complexity (e.g. UAV obstacle avoidance). In particular, the stability of DQN has been tried to be increased with the PER-D2MQN structure, and in this direction, it has been observed in our study that the classical DQN architecture is not robust to environmental uncertainties.

Additionally, as suggested by (Liao et al. 2024), combining Dueling and Noisy DQN architectures can increase success rates, especially in complex environment conditions (e.g., autonomous underwater vehicles), but such DQN variants have limited applicability in our experimental setting since they require more complex and computationally intensive systems than the original DQN structure (Liao et al. 2024). In this context, DQN has been shown to have room for improvement in terms of sample efficiency and learning stability.

Samuel's checker algorithm, with its low level of complexity and intuitive decision structure, showed a tendency for rapid learning, especially at the beginning of the training period. However, performance instability in the medium and long term and limited adaptation capacity limits the use of this method in modern tasks. However, as Sutton and Barto emphasize, Samuel's value function-based approach constitutes the conceptual origin of today's deep value-based systems and continues to be a source of inspiration (Sutton and Barto 1998).

As a result, Cross Entropy based optimization produces more efficient results in QT-Opt, which shows that this algorithm performs more effective and stable learning in continuous action spaces. Future studies can test the applicability of this method in multi-agent systems and representation-based learning scenarios, and hybridization of QT-Opt with current variants of DQN may offer new balance points between sample efficiency and computational cost (Wu and Suh 2024; Al-Hamadani et al. 2024).

Acknowledgments

The authors would like to thank Sakarya University of Applied Sciences Artificial Intelligence and Data Science Application and Research Center for supporting this study.

Ethical standard

The authors have no relevant financial or non-financial interests to disclose.

Availability of data and material

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that generative artificial intelligence (AI) tools were used during the preparation of this manuscript. Specifically, AI assistance was utilized for language editing, text refinement, and formatting purposes. The authors take full responsibility for the content and have carefully reviewed and verified all AI-assisted outputs.

LITERATURE CITED

- Al-Hamadani, M. N., M. A. Fadhel, L. Alzubaidi, and B. Harangi, 2024 Reinforcement learning algorithms and applications in healthcare and robotics: A comprehensive and systematic review. *Sensors* **24**: 2461.
- AlMahamid, F. and K. Grolinger, 2025 Agile dqn: Adaptive deep recurrent attention reinforcement learning for autonomous uav obstacle avoidance. *Scientific Reports* **15**: 1–18.
- Chen, C., J. Yu, and S. Qian, 2024 An enhanced deep q-network algorithm for localized obstacle avoidance in indoor robot path planning. *Applied Sciences* **14**: 11195.
- de Sousa Bezerra, C. D., F. H. T. Vieira, and D. P. Q. Carneiro, 2023 Autonomous robotic navigation approach using deep q-network late fusion and people detection-based collision avoidance. *Applied Sciences* **13**: 12350.
- Dong, Q., T. Kaneko, and M. Sugiyama, 2024 An offline learning of behavior correction policy for vision-based robotic manipulation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 5448–5454.
- Fan, J., Z. Wang, Y. Xie, and Z. Yang, 2020 A theoretical analysis of deep q-learning. In *Proceedings of Learning for Dynamics and Control*, pp. 486–489.
- Gao, T., 2024 Optimizing robotic arm control using deep q-learning and artificial neural networks through demonstration-based methodologies: A case study of dynamic and static conditions. *Robotics and Autonomous Systems* **181**: 104771.
- Gao, Y., J. Chen, X. Chen, C. Wang, J. Hu, *et al.*, 2023 Asymmetric self-play-enabled intelligent heterogeneous multirobot catching system using deep multiagent reinforcement learning. *IEEE Transactions on Robotics* **39**: 2603–2622.
- Hester, T., M. Vecerik, O. Pietquin, M. Lanctot, T. Schaul, *et al.*, 2018 Deep q-learning from demonstrations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.
- Janssonie, P., B. Wu, J. Perez, and J. Peters, 2024 Unsupervised skill discovery for robotic manipulation through automatic task generation. In *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, pp. 926–933.
- Kalashnikov, D., A. Irpan, P. Pastor, J. Ibarz, A. Herzog, *et al.*, 2018 Scalable deep reinforcement learning for vision-based robotic manipulation. In *Proceedings of the Conference on Robot Learning*, pp. 651–673.
- Liao, X., L. Li, C. Huang, X. Zhao, and S. Tan, 2024 Noisy dueling double deep q-network algorithm for autonomous underwater vehicle path planning. *Frontiers in Neurorobotics* **18**: 1466571.
- Mao, J., T. Lozano-Pérez, J. B. Tenenbaum, and L. P. Kaelbling, 2023 Learning reusable manipulation strategies. In *Proceedings of the Conference on Robot Learning*, pp. 1467–1483.
- Mnih, V., K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, *et al.*, 2015 Human-level control through deep reinforcement learning. *Nature* **518**: 529–533.
- Pal, A., A. Chauhan, and M. Baranwal, 2025 Together we rise: Optimizing real-time multi-robot task allocation using coordinated heterogeneous plays. arXiv preprint arXiv:2502.16079.

Samuel, A. L., 1959 Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* **3**: 210–229.

Sutton, R. S. and A. G. Barto, 1998 *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, USA.

Wei, S., C. Li, J. Seyler, and S. Eivazi, 2023 Integration of efficient deep q-network techniques into qt-opt reinforcement learning structure. In *Proceedings of the International Conference on Agents and Artificial Intelligence*, volume 3, pp. 592–599.

Wu, B. and C. S. Suh, 2024 Deep reinforcement learning for decentralized multi-robot control: A dqn approach to robustness and information integration. In *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, volume 88636, p. V005T07A035.

Zhang, H., S. Zeng, Y. Hou, H. Huang, and Z. Xu, 2025 Improved qt-opt algorithm for robotic arm grasping based on offline reinforcement learning. *Machines* **13**: 451.

How to cite this article: Özkurt, C., Küçükler, A. K., Karslıoğlu, M. and Özdemir, R. N. A Comparative Analysis of Deep Reinforcement Learning Approaches in Symbolic Optimization Tasks: The Case of DQN, QT-Opt and Samuel. *Computational Systems and Artificial Intelligence*, 2(1),15-20, 2026.

Licensing Policy: The published articles in CSAI are licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

