Advancements, Challenges and Future Prospects of Reinforcement Learning in Healthcare

Meiyi Feng[a]
School of Business, Macau University of Science and Technology, Macau Special Administrative Region, China

Keywords: Reinforcement Learning, Healthcare Application, Machine Learning.

Abstract: Reinforcement Learning (RL) has become a groundbreaking approach in machine learning, significantly impacting healthcare by providing solutions to intricate decision-making challenges. This comprehensive review examines the current state of RL in healthcare, focusing on dynamic treatment protocols, automated diagnostic systems, resource allocation, as well as privacy and security issues. RL's ability to adapt and optimize treatment plans dynamically, enhance diagnostic accuracy, and manage healthcare resources efficiently underscores its potential to revolutionize clinical practices. However, the implementation of RL in healthcare is fraught with challenges, including the need for extensive, high-quality datasets, difficulties in interpreting complex models, and significant data privacy concerns. To mitigate these challenges, recent innovations have been introduced. Transitional Variational Autoencoders (tVAEs) are used to generate realistic patient data, enhancing the simulation capabilities of RL models. Federated learning frameworks have been developed to ensure data privacy by enabling collaborative model training without sharing raw data. Additionally, transfer learning and domain adaptation techniques improve the generalization of RL models across diverse healthcare settings. This review provides a thorough analysis of these advancements and their implications for healthcare, offering a detailed understanding of RL's current applications and limitations. Future research directions are proposed to address existing challenges, aiming to ensure the robust, transparent, and ethical integration of RL technologies into clinical settings, thereby maximizing their potential to improve healthcare outcomes.

1 INTRODUCTION

Reinforcement Learning (RL) is an approach within the field of machine learning based on reward mechanisms, which has gained widespread attention and application in the healthcare sector in recent years. RL interacts with the environment to continuously adjust strategies to achieve the goal of maximizing long-term rewards. Its unique features make it particularly suitable for addressing complex decision-making problems in healthcare, such as optimizing disease diagnosis and treatment plans (Yu, Liu, Nemati & Yin, 2021; Coronato et al., 2020). In healthcare, RL applications range from dynamic treatment regimens to automated medical diagnosis. For example, in chronic disease management and intensive care, RL can help optimize treatment plans by dynamically adjusting drug dosages and treatment strategies to improve therapeutic outcomes (Yu, Liu, Nemati & Yin, 2021; Coronato et al., 2020). Additionally, RL is used in medical image analysis and disease prediction, such as early detection of lung cancer and vessel centerline tracking (Yu, Liu, Nemati & Yin, 2021; Yang et al. 2024).

Despite the promising prospects of RL in healthcare, several challenges and gaps remain in current research. First, training RL models often relies on large amounts of high-quality data, which are costly and ethically challenging to obtain in the healthcare domain (Yu, Liu, Nemati & Yin, 2021). Second, existing RL models exhibit limitations in handling multi-class imbalanced data, which may lead to poor predictive performance for minority classes in practical applications (Yang et al. 2024). Moreover, the application of RL in healthcare faces issues with model generalization. Many current models perform well on specific datasets but have not been validated across different patient populations

[a] https://orcid.org/0009-0003-7693-933X

Feng, M.
Advancements, Challenges and Future Prospects of Reinforcement Learning in Healthcare.
DOI: 10.5220/0013205700004568
In *Proceedings of the 1st International Conference on E-commerce and Artificial Intelligence (ECAI 2024)*, pages 39-44
ISBN: 978-989-758-726-9

and medical settings (Yu, Liu, Nemati & Yin, 2021). Additionally, data privacy and security are crucial research directions, especially when using distributed learning methods like federated learning (Otoum et al. 2021).

To address these research gaps, recent studies have proposed several improvements. For instance, Transitional Variational Autoencoders (tVAEs) are used to generate more realistic patient trajectories, enhancing the model's ability to simulate patient data. Furthermore, RL frameworks combined with federated learning are explored to strengthen data protection and confidentiality measures in healthcare IoT systems. This method allows model training without sharing raw data, thereby protecting patient privacy (Otoum et al. 2021).

This paper endeavours to offer an exhaustive analysis of the main research findings and recent advancements in the application of RL in healthcare. This paper explores a wide range of RL applications, including dynamic treatment regimes, automated medical diagnosis, and healthcare resource management. The review includes a detailed examination of existing models, their limitations, and the innovative solutions proposed to address these challenges. Additionally, this review is structured to explore the different facets and impacts of RL in healthcare, With the aspiration of offering an all-encompassing overview pertaining to the current framework and prospective developments of RL research in this significant sector.

2 METHODS

2.1 Introduction to Reinforcement Learning

Reinforcement learning, representing a sophisticated branch within the broader domain of machine learning, emphasizes enabling an agent to learn optimal decision-making through interactions with its environment. The fundamental principle involves the agent performing actions in different states with the aim of optimizing aggregate rewards over an extended period. Through the acquisition of feedback in the guise of rewards or penalties corresponding to its actions, the agent is steered towards formulating an optimal policy.

Reinforcement Learning challenges are commonly framed within the construct of Markov Decision Processes (MDPs), which are defined by the following key tuple S, A, P, R, γ (Amparore et al., 2013):

- s: A set of states
- a: A set of actions
- P: A transition likelihood matrix $P(s'|s,a)$, defining the likelihood of shifting to states from state s after executing an action a.
- R: A reward function $R(s,a)$, providing the instantaneous reward after action a in state s.
- γ: A discount factor gamma in $[0, 1]$, which quantifies the significance of future rewards.

The agent's paramount goal is to devise a strategy $\pi(s)$ that maximizes the anticipated aggregated reward, commonly known as the return. This return is calculated as the sum of discounted rewards accrued over time, reflecting both immediate and future benefits:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (1)$$

A commonly employed algorithm in RL is Q-learning, which updates the value of state-action pairs (Q-values) using the Bellman equation:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[R + \gamma max_{a'} Q(s',a') - Q(s,a)] \quad (2)$$

where α is the learning rate.

Through continuous interaction with its environment, the agent progressively learns to select actions that maximize long-term cumulative rewards. This ability renders RL especially effective for handling complex decision-making tasks across various fields, such as healthcare.

2.2 Dynamic Treatment Regimes

2.2.1 Enhancing DRL with Transitional Variational Autoencoders

In their study, Baucum and Khojandi introduce Transitional Variational Autoencoders (tVAEs) to improve Deep Reinforcement Learning (DRL) applications in healthcare. Temporal Variational Autoencoders (tVAEs) are sophisticated generative neural network models designed to create an explicit correlation between the configurations of clinical parameters across sequential temporal intervals, utilizing retrospective patient data. These models facilitate the accurate reconstruction of temporal patterns in clinical datasets by capturing the underlying distributional dynamics over time. One significant benefit of the tVAE model is its minimal reliance on distributional assumptions while maintaining consistent training and testing architectures. By utilizing tVAEs, researchers can create more realistic patient trajectories, facilitating the development of effective treatment policies (Baucum et al., 2020).

2.2.2 Closed-Loop Healthcare Processing with DRL and Human Body Simulators

Dai et al. propose a closed-loop healthcare processing method using DRL combined with Deep Neural Networks (DNNs) to build human body simulators. These simulators can dynamically accept interventions and produce observations. By integrating conceptual embedding techniques with DRL, the study investigates effective healthcare strategies. The system specifically consists of a virtual human physiological simulator seamlessly integrated with an advanced Deep Reinforcement Learning (DRL)-driven therapeutic intervention module. The treatment module diagnoses latent health states from high-dimensional radiographic observations and selects therapeutic actions to restore the simulated physiological model to an optimal healthy condition, creating a highly adaptive, autonomously self-regulating healthcare informatics architecture (Dai et al., 2022).

2.2.3 Personalized Healthcare

Coronato et al. provide a comprehensive review of RL's role in healthcare, emphasizing its potential to support personalized treatments in precision medicine. The document explores several RL applications, such as Dynamic Treatment Regimes (DTRs) for chronic illnesses like HIV, oncology, hypertension, and anemia. Utilizing the Sequential Multiple Assignment Randomized Trial (SMART) approach, RL aids in developing DTRs by employing sequential decision-making processes that consider patient responses at each stage of treatment. This approach aims to personalize drug dosages and treatment schedules to maximize therapeutic outcomes and minimize adverse effects (Coronato et al., 2020).

2.2.4 Dual-Agent RL Model for Optimizing Dynamic Treatment Regimes

Blumrosen et al. propose a Dual-Agent Reinforcement Learning (DaRL) model utilizing medical and natural agents to optimize healthcare interventions. The medical agent employs external data and sensors to adjust treatments, while the natural agent's rewards are based on internal biological states like neural circuits and dopamine levels. This integration enhances treatment accuracy and efficiency, dynamically adjusting drug dosages and reducing incorrect interventions by continuously monitoring patient health (Blumrosen et al., 2019).

2.3 Automated Medical Diagnosis

2.3.1 Automated Diagnosis with Machine Learning Models

Structured medical data, such as physiological signals, images, and lab tests, benefit from RL methods for tasks like feature extraction and image segmentation. Key techniques include Q-learning for prostate segmentation in ultrasound images, Trust Region Policy Optimization (TRPO) for surgical gesture segmentation, and Deep Q-Network (DQN) for automatic landmark detection in MRI and CT images, achieving robustness and accuracy. Unstructured medical data, such as clinical notes and reports, use RL techniques for diagnosis inference. Key methods include DQN to improve diagnosis accuracy using external evidence, adaptive online learning combining supervised learning for risk assessment and RL for decision making, and symptom checking systems using DQN to enhance efficiency and accuracy in diagnosis (Yu, Liu, Nemati & Yin, 2021).

2.3.2 Deep Reinforcement Learning for IoT-Enabled Smart Healthcare Systems

Jagannath et al. present a novel IoT-enabled smart healthcare system that utilizes DRL to automate medical diagnosis and decision-making. The system architecture is composed of four tiers: patient data acquisition, edge processing, data conveyance, and cloud computation. Patient data is gathered through various sensors in a Body Area Network (BAN) and sent to data centers using IoT protocols. DRL algorithms, like Deep Q-Network (DQN), are utilized to conduct intricate data analysis and formulate sophisticated diagnostic and therapeutic determinations. The system was evaluated with synthetic data from BAN sensors, showing high accuracy in estimating hidden health states and making decisions that closely align with those of a physician. This method not only improves diagnostic accuracy but also provides an efficient telemedicine solution (Jagannath et al., 2022).

2.3.3 Dueling Double Deep Q-Network for Multi-Class Imbalance in Healthcare Applications

Yang, EI-Bouri et al. introduce a dueling double deep Q-network (D-DDQN) to tackle the issue of multi-class imbalance in reinforcement learning, particularly for healthcare applications. The Q-value function is optimized through policy iteration. The dueling architecture separates value and advantage estimation, enabling better state-action representation. Double deep Q-learning mitigates overestimation by using separate networks for action selection and value estimation. This approach is particularly beneficial in healthcare for optimizing resource allocation and improving diagnostic accuracy in class-imbalanced medical datasets (Yang et al., 2024).

2.4 Health Resources Allocation and Scheduling and Health Management

2.4.1 Enhanced Intelligent Clustering-Oriented Routing Framework for 5G-Integrated Smart Healthcare Solutions

Ahad et al. introduce an sophisticated Clustering-based Routing Protocol (CRP-GR) that integrates game theory with reinforcement learning to improve resource allocation within a 5G-driven smart healthcare paradigms. This protocol is designed to optimize energy usage and extend network lifespan by selecting energy-efficient cluster heads and determining the optimal multipath routes for data transmission. By using reinforcement learning, the protocol adapts to the heterogeneous and dynamic nature of smart healthcare networks, improving the quality of service (QoS) and ensuring efficient use of resources (Ahad et al., 2021).

2.4.2 EPRAM: Enhancing Smart Healthcare with Fog Computing

Talaat et al. introduce the Effective Prediction and Resource Allocation Method (EPRAM) within fog computing architectures for advanced cognitive healthcare applications enhances real-time resource management and prediction using deep reinforcement learning and probabilistic neural networks. EPRAM achieves low latency, high resource utilization, and effective load balancing, outperforming traditional algorithms (Talaat et al., 2022).

2.5 Security and Privacy

2.5.1 Federated Reinforcement Learning-Augmented Intrusion Detection Framework for IoT-Enabled Healthcare Ecosystems

Otoum et al. discuss the implementation of an Intrusion Detection System (FRL-IDS) which is grounded in Federated Reinforcement Learning designed to enhance security and privacy in IoT-enabled healthcare systems. The proposed model leverages federated learning to maintain data privacy by ensuring that data remains localized while only sharing model parameters. This approach helps in detecting and mitigating cyber intrusions without compromising patient data. The system processes data from multiple healthcare entities and updates the global detection model collaboratively, ensuring robust and secure data communication across the network (Otoum et al., 2021).

2.5.2 Reinforcement Learning-Based Trajectory Design for WPT-Enabled UAV Healthcare Delivery

Merabet et al. explore the deployment of reinforcement learning methodologies to enhance the security and efficiency of healthcare delivery systems utilizing Unmanned Aerial Vehicles (UAVs). The proposed system employs Wireless Power Transfer (WPT) technology, enabling UAVs t to replenish their power reserves mid-mission, thereby extending their operational range. The RL algorithm is used to design optimal UAV trajectories that minimize travel time and energy consumption while ensuring secure data transmission during deliveries. By implementing RL, the system can dynamically adapt to changing environmental conditions and potential security threats, providing a reliable and secure healthcare delivery solution (Merabet et al., 2022).

3 DISCUSSIONS

3.1 Limitations and Challenges

3.1.1 Interpretability

RL models, particularly deep reinforcement learning, are often termed "opaque systems" owing to the intricate and enigmatic nature of their decision-making processes, which are not easily interpretable by humans. In healthcare, interpretability is crucial as medical professionals need to trust and understand the reasoning behind a model's recommendations. The lack of transparency can impede the adoption of RL in clinical settings, as practitioners may be hesitant to trust systems they cannot fully understand. This challenge necessitates the development of methods that can elucidate the internal workings of RL models, making them more accessible and trustworthy.

3.1.2 Applicability

The application of RL in healthcare is still in its infancy, and its practical use is limited by several factors. Firstly, RL requires extensive amounts of data to train models effectively, which can be a significant barrier in medical domains where data is often scarce and sensitive. Secondly, healthcare environments are highly complex and dynamic, making it difficult to design RL algorithms that can adapt to such variability. Finally, the integration of RL systems into existing healthcare workflows presents another layer of complexity, as these systems must be seamlessly incorporated into routine practices without disrupting care delivery.

3.1.3 Privacy

Patient data privacy is a paramount concern in healthcare. RL models require substantial amounts of data, which raises significant privacy issues. Ensuring that patient information is protected while using it to train RL models is a complex task. Methods such as data anonymization and secure data sharing protocols are necessary but may not be sufficient to address all privacy concerns. Additionally, the potential for data breaches and the misuse of sensitive information are ongoing risks that need to be continually managed and mitigated.

3.2 Future Prospects

3.2.1 The Solutions for Interpretability

To address the interpretability challenge, integrating expert systems with RL can provide a framework where the decision-making process is more transparent. Expert systems utilize heuristic-based methodologies to emulate the cognitive judgment capabilities of a domain-specific specialist. By combining this with RL, it is possible to create systems that not only learn from data but also incorporate domain-specific knowledge that is easier for practitioners to understand. Approaches such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) present promising methods for interpreting the outputs of complex models. SHAP assigns an importance value to each feature for a specific prediction, thereby elucidating the influence of each feature on the model's output. LINE approximates the model locally with a simpler interpretable model, providing insights into how the original model works in the vicinity of the instance being predicted. These methods can help demystify the decisions made by RL algorithms, thereby increasing their acceptance and trust among healthcare professionals.

3.2.2 The Solutions for Applicability

Cross-domain knowledge transfer and domain specific adaptation hold significant potential for enhancing the applicability of RL in healthcare. Transfer learning entails using a model pre-trained in one domain and fine-tuning it for a related but different domain, thus reducing the data needed for training in the new domain. This is especially beneficial in healthcare, where data is frequently scarce and costly to acquire. Domain adaptation, on the other hand, helps models generalize across various settings by aligning the data distributions between the source and target domains. This ensures that RL models can perform well even when there are differences between the training data and the real-world data they encounter in deployment. By applying these techniques, RL models can become more robust and effective in various healthcare environments, from hospitals in urban centers to rural clinics with different patient demographics.

3.2.3 The Solutions for Privacy

Federated learning offers a solution to the privacy challenge by enabling the training of RL models across multiple institutions without sharing raw data. This approach allows models to learn from a broader dataset while maintaining patient confidentiality. In federated learning, data remains localized at each institution, and only the model updates (gradients) are shared and aggregated to improve the global model. This decentralized approach to training models mitigates the risks associated with centralized data storage, such as data breaches and privacy violations. By aggregating insights from decentralized data sources, federated learning can help create more comprehensive and accurate RL models. For example, hospitals in different regions can collaborate to train an RL model on patient treatment outcomes without exposing sensitive patient data. This collective learning process can result in more effective and generalized healthcare solutions, potentially accelerating the adoption of reinforcement learning in the healthcare sector.

4 CONCLUSIONS

To summarize, this paper has provided a thorough analysis of RL applications, including dynamic treatment protocols, automated diagnostic systems, resource allocation, as well as privacy and security measures. The primary contribution of this study lies in highlighting RL's ability to enhance healthcare outcomes. However, RL faces significant limitations, such as data scarcity, model interpretability, and privacy concerns. Future research should aim to integrate expert systems for improved interpretability, leverage transfer learning to enhance applicability, and employ federated learning to address privacy issues. These advancements will be crucial in fully realizing the potential of RL in healthcare, ensuring robust, transparent, and ethical deployment in clinical settings.

REFERENCES

Ahad, A., Tahir, M., Sheikh, M. A., Ahmed, K. I., & Mughees, A. 2021. An intelligent clustering-based routing protocol (crp-gr) for 5g-based smart healthcare using game theory and reinforcement learning. *Applied Sciences*, *11*(21), 9993.

Amparore, E. G., & Donatelli, S. 2013. *States, actions and path properties in Markov chains* (Doctoral dissertation, PhD thesis, University of Torino, Italy).

Baucum, M., Khojandi, A., & Vasudevan, R. 2020. Improving deep reinforcement learning with transitional variational autoencoders: A healthcare application. *IEEE Journal of Biomedical and Health Informatics*, *25*(6), 2273-2280.

Blumrosen, G. 2019, October. Enhancing healthcare quality with reinforcement learning modeling. In *2019 IEEE SENSORS* (pp. 1-4). IEEE.

Coronato, A., Naeem, M., De Pietro, G., & Paragliola, G. 2020. Reinforcement learning for intelligent healthcare applications: A survey. *Artificial intelligence in medicine*, *109*, 101964.

Dai, Y., Wang, G., Muhammad, K., & Liu, S. 2022. A closed-loop healthcare processing approach based on deep reinforcement learning. *Multimedia Tools and Applications*, *81*(3), 3107-3129.

Jagannath, D. J., Dolly, R. J., Let, G. S., & Peter, J. D. 2022. An IoT enabled smart healthcare system using deep reinforcement learning. *Concurrency and Computation: Practice and Experience*, *34*(28), e7403.

Merabet, A., Lakas, A., & Belkacem, A. N. 2022, May. WPT-enabled UAV trajectory design for Healthcare delivery using reinforcement learning. In *2022 International Wireless Communications and Mobile Computing (IWCMC)* (pp. 271-277). IEEE.

Otoum, S., Guizani, N., & Mouftah, H. 2021. Federated reinforcement learning-supported IDS for IoT-steered healthcare systems. In *ICC 2021-IEEE International Conference on Communications* (pp. 1-6). IEEE.

Talaat, F. M. 2022. Effective prediction and resource allocation method (EPRAM) in fog computing environment for smart healthcare system. *Multimedia Tools and Applications*, *81*(6), 8235-8258.

Yang, J., El-Bouri, R., O'Donoghue, O., Lachapelle, A. S., Soltan, A. A., Eyre, D. W., ... & Clifton, D. A. 2024. Deep reinforcement learning for multi-class imbalanced training: applications in healthcare. *Machine Learning*, *113*(5), 2655-2674.

Yu, C., Liu, J., Nemati, S., & Yin, G. 2021. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)*, *55*(1), 1-36.