

A Comprehensive Review on Advance in Deep Reinforcement Learning for Health Care Optimization

Jannatul Ferdous Disha, Kazi Md. Tanvir Anzum, Istiak Mahmud, MD Faiiaz Zaman Dehan and MD. Mahfuzur Rahman Masum

Department of Industrial Engineering and Management

Khulna University of Engineering & Technology, Khulna-9203, Bangladesh

jf.disha2001@gmail.com, tanvir@iem.kuet.ac.bd, istiaknoman12007021@gmail.com,

fz.dehan7815070987@gmail.com, rahmanmahfuz848@gmail.com

Abstract

Over the recent years, Deep Reinforcement Learning (DRL) has been an essential practice in enhancing profitability in healthcare sector to deliver customized interventions in hospitals, optimize resources, and make predictive models, especially in emerging economies. Nevertheless progress, challenges such as model interpretability and real-world integration remain. This study covers optimization approaches in DRL, with a focus on Proximal Policy Optimization (PPO) for its stability, efficiency, and application in healthcare. Beyond Proximal Policy Optimization (PPO), we will go into a larger repertoire of algorithms such as Soft Actor-Critic (SAC) and Twin Delayed Deep Deterministic Policy Gradient (TD3) and Deep Q-Networks (DQN) and Multi-Agent DRL (MADRL). We critique their use in various fields of healthcare that include adaptive radiation therapy, chronic illness treatment, and detection of diseases at a premature stage. As part of our methodology, we performed a broadened systematic review of 2004-2024 with the combination of grey literature and varying databases. As indicated in the evidence, DRL has great potential in improving the outcomes of patients and in efficiency of care. Future research directions include emerging frontiers in Explainable DRL, Federated DRL, and Digital Twin-based interventions as one of their future directions. The paper tries to advise on how DRL can be adopted in healthcare systems with a goal to achieve patient-centered, efficient, and data-driven clinics decision support.

Keywords

Deep Reinforcement Learning, Optimization Techniques, Healthcare Applications, Personalized Medicine, Explainable AI, Digital Twin

1. Introduction

Deep Reinforcement Learning (DRL) is a method of combining Deep Learning and Reinforcement learning to be used by systems to learn the best course of action by using the trial and error in more complex environments (Arulkumaran et al. 2017). DRL in healthcare can enable personalized care as well as optimal resources usage and predictive modeling. Despite such advantages, challenges or barriers do still exist which are: model interpretation, the calculation requirements and incorporation into clinical processes. In the recently examined review, newer DRL models and optimization methods are considered, which are intended to bridge the gap between the theoretically provided advances and the introduction of these into healthcare.

The major research questions addressed in this paper are the following ones:

- 1) Which optimization methods are now used in deep reinforcement learning?
- 2) What is the role of such methods in healthcare and what can they do to improve clinical outcomes?

3) Which empirical case studies can demonstrate the usefulness of optimized DRL models in the perspective of real-life healthcare?

This paper will be able to help in the overall initiative of incorporating the state-of-the-art method of machine learning into the healthcare domain by reading through methods of optimization implemented in the deep reinforcement learning application to the system level. The paper is structured in the following way: Section I provides the background of the deep reinforcement learning and its application to the healthcare sector. The second section is an overview of some important optimization strategies such as Proximal Policy Optimization (PPO) (Schulman et al. 2017). Section III contains details of the methodology followed during conducting the systematic review. Section IV contains the results and the outcomes and provides an analytical analysis of the efficacy of different optimization procedures in medical sectors. Section V contains hints on the directions of future research. Section VI provides some conclusions, and Section VII provides the list of all works that have been mentioned.

2. Literature Review

The given section covers the usage of deep reinforcement learning in healthcare and especially focuses on optimization methods used in its application and their optimal use.

2.1 DRL in Healthcare

Deep Reinforcement Learning (DRL) is a method, where agents learn to behave best in a dynamic environment with an objective of maximum cumulative rewards (Abdellatif et al. 2023a). In contrast to supervised learning where fixed datasets are encountered, DRL constantly evolves, and as such, it is very applicable in real-life problems such as those in the field of healthcare where decision-making is involved. In this respect, DRL may assist prescribing the individual care plan, patient monitoring, automation of clinical tasks, resource management, and dose prescriptions of chronic health conditions like diabetes, and hypertension (Yu et al. 2021a).

There are three main challenges that affect effective DRL:

1) *Exploration–Exploitation Trade-off*: Agents should learn new behaviors to acquire knowledge and at the same time exploit what they know to achieve immediate rewards. Over exploration slows down learning whereas over exploitation may result in convergence to in-optimal policies (Sohaib et al. 2024).

2) *Learning Stability*: Stability guarantees that the performance of agents advances with time, particularly in sophisticated environments such as health care. To ensure that the learning process is stable, some learning algorithms provide the stability through policy smoothing like Proximal Policy Optimization (PPO) and Trust Region Policy Optimization (TRPO) (Yu et al. 2021b).

3) *Computational Efficiency*: During the training of DRL models, colossal amounts of computational powers are required. This can decrease this burden with the help of function approximation and parallel computation use, making DRL viable in a clinical practice environment (Abdellatif et al. 2023b).

2.2 DRL Algorithms of Optimization

1) *Proximal Policy Optimization (PPO)*: PPO is a policy gradient technique, and it aims at stable and reliable learning. It restricts the degree of change the policy may have throughout training, too much change can cause instability in learning. PPO has been communal in adaptive treatment planning, where insignificant changes in patient care, which can be achieved through real-time feedback, are important in keeping safe and effectual (Siboo et al. 2023).

2) *Soft Actor-Critic (SAC)*: SAC is a non-policy algorithm and works well on continuous action space. It maximizes expected reward and entropy combination, which results in exploration. Since it is also sample efficient and robust it is perfectly suited to continuous control, e.g. real time dosage adjustment or dynamic ventilator settings in ICU treatment (Liu et al. 2024).

3) *Twin Delayed Deep Deterministic Policy Gradient (TD3)*: The TD3 annihilates overestimation bias in value functions and postpones policy update to enhance policy stability than DDPG. Special application of this algorithm

may be resource optimization, e.g. ICU bed utilization optimization or medical equipment utilization in hospitals (Wang et al. 2024).

4) *Deep Q-Network (DQN)*: DQN goes further to enhance Q-learning by leveraging on the power of deep neural networks to cope with complex applications of the state spaces. It primarily fits discrete decision-making tasks, like early diagnosis alerts, triage prioritization or suggestion of diagnostic tests (Liu et al. 2020).

5) *Multi-Agent DRL (MADRL)*: MADRL is a multi-agent set of learning or interactions taking place in a common environment. It is appropriate on complex systems such as hospitals where agents can be required to handle logistics, schedules of staff or a multi-patient schedule, the coordinated effort leading to an increase of the efficiency of the whole system (Ao et al. 2023).

3. Methodology

The approach to review is a thorough examination of academic papers in which the authors focused on the period between 2004 and 2024 and include such sources as Google Scholar and Science Direct.

3.1 Keyword Search

The aim of the keyword-based search was to find the emerging trends in the optimization methods used in deep reinforcement learning. The literature search is comprehensive and exhaustively done through selection of carefully chosen keywords.

("Reinforcement Learning") AND ("Optimization Techniques") AND ("Healthcare Applications" OR "Health care Applications") AND ("Predictive Modeling") AND ("Personalized Medicine")

3.2 The Review Process

The review was conducted following the requirements of the PRISMA, which is why it was started by a search in databases resulting in 500 articles. Of the 35 articles located through the identified search terms, there were 25 articles after removal of duplicates and screening of the titles and abstracts. Using exclusion criteria non-English articles, off-topic articles and articles feeding off contextual information rather than deep reinforcement learning optimization methods such as PPO, 12 articles were discarded. The remaining 13 articles were read in entirety to enable in detail examination. This way of strict selection ensured the relevancy of the approaches, as well as this provided the targeted discussion on DRL optimization in healthcare (Wang et al. 2023).

3.3 Extraction of information

The main information about each study was given in a structured form and contained the names of authors, objectives, methods, task types, data characteristics, and evaluation criteria. Common patterns were established using thematic analysis, which gave an insight on DRL optimization in healthcare (Aliyu, Akhir, et al. 2024). Results are presented around base research questions to make them simple and specific (Figure 1).

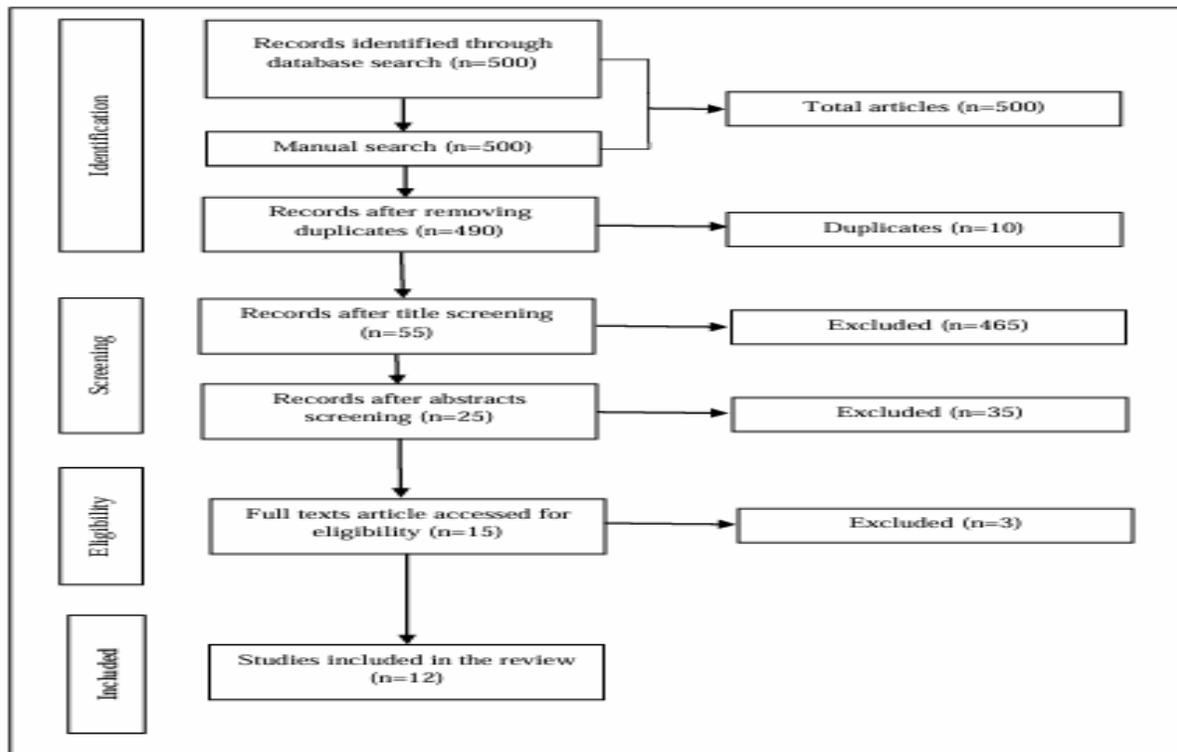


Figure 1. PRISMA-based visual selection summary

4. Results

4.1 RQ1: Which DRL optimization methods are applied?

The review features the many different DRL methods that help to maximize clinical efficiency:

- 1) *Cloud-based DRL backed by the use of a Kalman filter*: This is a variant of the cloud-based DRL which makes the overall healthcare service selection more effective due to the efficient use of cloud-based environment resources. This would help in the delivery of services in an improved way as the system is designed to dynamically adapt to such changes in load and demand (Zhong et al. 2024).
- 2) *DRL with Tumor-Response Models*: Ensures the greatest benefit of adaptive radiation therapy (ART) by appropriately adapting the plans of the treatments based on responses of the tumor. It becomes more specific and safe because of the personal dose situation (Tseng et al. 2017).
- 3) *Dual-DRL at Health Interventions*: Makes the type, timing and frequency of an intervention on chronic disease management specific. Such models enable a real-time adjustment of digital health strategies as a way of enhancing the engagement of the patients (Xu and Zhang 2023).
- 4) *Multi-Agent DRL in Home Care*: Minimizes health risks caused by the wrong administration of medication and helps with the care of the elderly due to their establishment as cooperating agents. The interconnections of the multi-agent frameworks organize various activities in the care field, which increases the quality and independence of home care (Nahom Abishu et al. 2025).
- 5) *DQN in managing Sepsis and Diabetes*: Invoked to discover the best series of treatment procedures based on the history of the patient retrieved in Electronic Health Records (EHRs). The adaptive policies of DQN contribute to prompt and efficient decisions made in relation to the clinics (Wu et al. 2023).

6) *Time2Vec-based Attention Mechanisms*: The method offers negative glycemc events estimation in health data by enhancing the activity of usage of features. The application of such practices permits preventative control of glucose (López-Andreu et al. 2023).

7) *Contextual Bandits*: To personalized care morphing the patient context to the intervention to adjust chronic diseases outcomes. The potential is that the approach focuses on the optimal care decision-making in real-time that balances pursuit of expansion and deepening (Zhou et al. 2023).

8) *DRL on PK/ PD Modeling*: Personalization of warfarin by personal pharma-co-dynamic/ pharma-co-kinetic modeling of warfarin. It reduces complications that pertain to the administration of medication since dosing is customized according to the reaction of individual patients (Tosca et al. 2024) (Table 1).

Table 1. Comparison of Optimization Methods

Method	Type	Strength	Application	Limitation
PPO	On-policy	Stable, easy to tune	Treatment planning	Sample inefficient
SAC	Off-policy	High sample efficiency	Dosage optimization	Complex tuning
TD3	Off-policy	Reduces overestimation	Resource allocation	Longer training
DQN	Off-policy	Handles discrete actions	Early alerts	Poor in continuous spaces
MADRL	Multi-agent	Scalability	Hospital logistics	High complexity

4.2 RQ2: How are these techniques used and what is their effect on health care outcomes?

1) *Cloud Optimization*: DLR merges with the concept of cloud computing in order to maximize the potential of the high quality healthcare services that can respond and can provide cost effective services by means of choosing and delivering service. The outcome is that there is improved service reliability to key services within healthcare and latency is reduced (Timčenko 2020).

2) *Chronic Disease Management*: DRL gives personalized reminders and preventive measures to patients leading to a huge decrease in drugs mistakes and better medication compliance that stimulates better patient outcomes and decreases hospitalization frequency (Reynolds et al. 2018).

3) *Adaptive Radiation Therapy*: DRL dynamic shaping of radiation intensity in accordance with the characteristics of boundaries of the tumor object, in order to ensure maximum tumor control after treatment and maximum sparing of healthy tissues. This brings about better treatment and better tolerance to the patient (Lim-Reinders et al. 2017).

4) *Early Diagnosis*: With the help of the large-scale IoT data analysis, DRL enables the early diagnostics of diseases such as lung cancer, making it more likely to be overcome and thus eliminating the complications of the late diagnosis (Khajuria and Sarwar 2022).

5) *Treatment Pathway*: DRL uses a personalized approach to treating sepsis and diabetes and makes the treatment plan case-specific, leading to a better glucose control and an increased resource utilization and healthcare efficiency (Keerthika et al. 2024).

6) *Medication Optimization*: DRL maximizes therapy dosing, including such drugs as warfarin and diabetes therapies, to improve therapeutic value and mitigate the occurrence of adverse drug events and increase patient safety (Aliyu, Patah Akhir, et al. 2024).

7) *Asthma Forecasting*: DRL helps to accurately anticipate asthma attacks with the help of past patient data and risk score analysis, taking preventive action before the attack happens and positively impacting the quality of life and the need to use emergency treatments (Do et al. 2019).

4.3 RQ3: What are the case studies to develop an idea of the application of optimized reinforcement learning models in healthcare environments?

The capability of DRL to individualize care provision and produce an improved clinical decision making process has made it a game-changer to the health sector. The section will discuss a set of case studies that illustrate how DRL may be used in conducting individualized medication management, adaptive radiation therapy, early disease detection, optimization of treatment pathways and precision dosing. The abovementioned examples demonstrate the potential of the DRL to transform patient outcomes and make the healthcare systems more efficient overall.

1) Management of Chronic disease and medication adherence

Case Study 1: An DRL-based system sent a personalized alert on drug on-time schedules to ensure that they received it with reduced errors and adherence and found to be effective in the elderly and cognitively impaired patients (Balkrishnan 2005).

Case Study 2: Dual DRL models have been used to control an intervention (type, frequency, timing) in a process of chronic diseases management to optimize patient engagement and outcomes by supplying personalized patient support at the right time (Brummel and Carlson 2016).

2) Cancer-Adaptive Radiotherapy

Case Study 3: In lung cancer, an DRL-based Adaptive Radiation Therapy (ART) modified doses according to the response of the tumor and maximized the tumor BED, decreased variation, and spared healthy tissue, as well (Morgan and Sher 2020).

Case Study 4: A Deep perturbation and integrated learning on deep canonical correlation (Deep RL) model adjusted the dose in scrolling CT of lung cancer patients in real time, which provided higher level of dose precision and treatment success as compared to uniform dose protocols (Brock 2019).

3) Early Detection and Prediction

Case Study 5: DRL integrated with IoT data facilitated the early detection of lung cancer and thus the performance of the diagnostic process as well as the prognosis of survival via timely interventions (Schaefer et al. 2004).

4) Optimization of treatment traces

Case Study 6: A DRL algorithm was applied to the EHR data to tailor the sequencing of treatments to various chronic diseases such as sepsis and diabetes to produce better clinical outcomes (Ngo et al. 2012).

Case Study 7: An DRL model that pays particular attention was used to predict glycemic events on diabetics, enabling early intervention and glycemic control (Ngo et al. 2012).

5) Patient-Wise Optimal Dosing of Drugs

Case Study 8: DRL had streamlined the warfarin dosing of virtual patients, streamlining errors and increasing the effectiveness of the therapy of the drug with a narrow safety margin (Houy and Le Grand 2019).

6) Forecasting of the Attack of Asthma

Case Study 9: DRL (Deep Q-Learning) forecasted asthma attacks based on old data, which allowed preventing them and made them less severe and frequent (Kim et al. 2020).

4.4 DRL application in Healthcare

1) Personalized Chronic Disease Management: Tailor-made intervention in diabetes and hypertension was implemented with the help of SAC and PPO (Abdellatif et al. 2023c).

2) Adaptive Radiation Therapy: TD3 and PPO enhance a more optimality in the intake of dosage of radiations daily (Abdellatif et al. 2023c).

3) Early Diagnosis of the Disease: With the use of DQN, preemptive warnings of asthma and cancer can be provided with the help of IoT information (Abdellatif et al. 2023c).

4) Resource Allocation: SAC and better MADRL is used in the administration of the ICU beds and also the scheduling (Abdellatif et al. 2023c).

5. Future Research Directions

Scalability and Clinical Validation of DRL Models: Real-world clinical trials are also necessary to confirm DRL models on the population of different patients and various healthcare environments. The adoption needs to be ensured by making it scalable in the controlled set-ups into the complex and real-time clinical operation.

1) *Electronic health records (EHR) integration*: The main direction of future studies regarding DRL algorithms should implement their smooth integration into an EHR system, allowing access to real-time data, individual treatment recommendations, and patient monitoring through their lifetime data and ensuring data integrity and interoperability (Tiase et al. 2021).

2) *Framework of Standardized Benchmarks Development*: Defining consistent databases, assessment parameters and test conditions of DRL in healthcare will make possible to conduct comparative testing of various algorithms and guarantee consistency and comparability of evaluation (Otis-Wilborn et al. 2000).

3) *Ethical and Regulatory Issues*: To make reasonable use of DRL systems in the health sector, research should investigate such ethical concerns as bias, transparency, and accountability and align the findings with regulatory strategies such as HIPAA or GDPR (Pesapane et al. 2018).

4) *Explainabilities and Trust in AI decisions*: It is important to come up with explainable models of DRL that can be employed by clinicians. Explainable AI (XAI) research that can be specifically applied to DRL should be conducted to facilitate the usability (Shin 2021).

5) *Real-Time deployment and computational Efficiency*: Additional lightweight algorithm and hardware acceleration innovation (e.g. GPU, or edge computing) is required to enable real-time and applications taking place at the bedside using DRL (Koubaa et al. 2022).

6) *Stability to Data Sway and Uncertainty*: The nature of missing, noisy, or biased clinical data and how DRL models may be used to address it will also be an important research area on the way to the field of clinical-related healthcare (Raymakers et al. 2005).

7) *Multi Agent Collaboration in Complex Systems*: Along with a larger scale of usage as multi-agent frameworks, DRL can better contribute to results at a system level in hospital operations, emergency response, and harmonized patient care (Herrera et al. 2020).

6. Conclusion

The use of deep reinforcement learning (DRL) in healthcare can be considered an enormous step towards better patient outcomes, and the potential of this approach to healthcare is tremendous. Incorporating more advanced optimization algorithms, namely Proximal Policy Optimization (PPO), SAC, TD3, and MADRL, DRL framework can easily overcome the inherent complexities of clinical setups, thus facilitating the substantial and efficient decision-making process. Empirical case studies in all spheres- such as medication adherence and radiation therapy- have shown the ability of DRL to improve treatment accuracy and ensure quality of care in general. Nevertheless, the future prospects of DRL utilization are rather bright despite the existing challenges, most of which concern the interpretability of models and the complexity of healthcare data. It preconditions development of efficient, evidence-based and patient-oriented healthcare systems.

Acknowledgements

The authors would like to extend their appreciation to the individuals who participated in this study. Their assistance and cooperation were indispensable in making this research possible.

References

Abdellatif, Alaa Awad, Naram Mhaisen, Amr Mohamed, Aiman Erbad, and Mohsen Guizani. "Reinforcement Learning for Intelligent Healthcare Systems: A Review of Challenges, Applications, and Open Research Issues." *IEEE Internet of Things Journal* 10 (24): 21982–2007. 2023a.

- Abdellatif, Alaa Awad, Naram Mhaisen, Amr Mohamed, Aiman Erbad, and Mohsen Guizani. "Reinforcement Learning for Intelligent Healthcare Systems: A Review of Challenges, Applications, and Open Research Issues." *IEEE Internet of Things Journal* 10 (24): 21982–2007. 2023b.
- Abdellatif, Alaa Awad, Naram Mhaisen, Amr Mohamed, Aiman Erbad, and Mohsen Guizani. "Reinforcement Learning for Intelligent Healthcare Systems: A Review of Challenges, Applications, and Open Research Issues." *IEEE Internet of Things Journal* 10 (24): 21982–2007. 2023c. <https://doi.org/10.1109/JIOT.2023.3288050>.
- Aliyu, Dahiru Adamu, Emelia Akashah Patah Akhir, Nurul Aida Osman, Jabir Abubakar Salisu, Yahaya Saidu, and Jameel Shehu Yalli. "Optimization Techniques in Reinforcement Learning for Healthcare: A Review." 1–6. 2024.
- Aliyu, Dahiru Adamu, Emelia Akashah Patah Akhir, Nurul Aida Osman, Jabir Abubakar Salisu, Yahaya Saidu, and Jameel Shehu Yalli. "Optimization Techniques in Reinforcement Learning for Healthcare: A Review." *2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, August 23, 1–6. 2024. <https://doi.org/10.1109/ICCUBEA61740.2024.10774698>.
- Ao, Tianyong, Kaixin Zhang, Huaguang Shi, Zhanqi Jin, Yi Zhou, and Fuqiang Liu. "Energy-Efficient Multi-UAVs Cooperative Trajectory Optimization for Communication Coverage: An MADRL Approach." *Remote Sensing* 15 (2): 429. 2023.
- Arulkumar, Kai, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. 2017. "Deep Reinforcement Learning: A Brief Survey." *IEEE Signal Processing Magazine* 34 (6): 26–38.
- Balkrishnan, Rajesh. "The Importance of Medication Adherence in Improving Chronic-Disease Related Outcomes: What We Know and What We Need to Further Know." *Medical Care* 43 (6): 517–20. 2005. <https://doi.org/10.1097/01.mlr.0000166617.68751.5f>.
- Brock, Kristy K. "Adaptive Radiotherapy: Moving Into the Future." *Seminars in Radiation Oncology* 29 (3): 181–84. <https://doi.org/10.1016/j.semradonc.2019.02.011>. 2019.
- Brummel, Amanda, and Angeline M. Carlson. "Comprehensive Medication Management and Medication Adherence for Chronic Conditions." *Journal of Managed Care & Specialty Pharmacy* 22 (1): 56–62. <https://doi.org/10.18553/jmcp.2016.22.1.56>. 2016.
- Do, Quan, Son Tran, and Alexa Doig. "Reinforcement Learning Framework to Identify Cause of Diseases - Predicting Asthma Attack Case." *2019 IEEE International Conference on Big Data (Big Data)*, December, 4829–38. <https://doi.org/10.1109/BigData47090.2019.9006407>. 2019.
- Herrera, Manuel, Marco Pérez-Hernández, Ajith Kumar Parlikad, and Joaquín Izquierdo. "Multi-Agent Systems and Complex Networks: Review and Applications in Systems Engineering." *Processes* 8 (3): 312. 2020. <https://doi.org/10.3390/pr8030312>.
- Houy, Nicolas, and François Le Grand. "Personalized Oncology with Artificial Intelligence: The Case of Temozolomide." *Artificial Intelligence in Medicine* 99 (August): 101693. 2019. <https://doi.org/10.1016/j.artmed.2019.07.001>.
- Keerthika, K., M. Kannan, and T. Saravanan. "Clinical Intelligence: Deep Reinforcement Learning for Healthcare and Biomedical Advancements." In *Deep Reinforcement Learning and Its Industrial Use Cases*, 1st ed., edited by Shubham Mahajan, Pethuru Raj, and Amit Kant Pandit. Wiley. 2024. <https://doi.org/10.1002/9781394272587.ch7>.
- Khajuria, Rishi, and Abid Sarwar. "Reinforcement Learning in Medical Diagnosis: An Overview." In *Recent Innovations in Computing*, edited by Pradeep Kumar Singh, Yashwant Singh, Maheshkumar H. Kolekar, Arpan Kumar Kar, and Paulo J. S. Gonçalves, vol. 832. 2022. Lecture Notes in Electrical Engineering. Springer Singapore. https://doi.org/10.1007/978-981-16-8248-3_15.
- Kim, Dohyeong, Sunghwan Cho, Lakshman Tamil, Dae Jin Song, and Sungchul Seo. "Predicting Asthma Attacks: Effects of Indoor PM Concentrations on Peak Expiratory Flow Rates of Asthmatic Children." *IEEE Access* 8: 8791–97. 2020. <https://doi.org/10.1109/ACCESS.2019.2960551>.
- Koubaa, Anis, Adel Ammar, Anas Kanhouh, and Yasser AlHabashi. "Cloud Versus Edge Deployment Strategies of Real-Time Face Recognition Inference." *IEEE Transactions on Network Science and Engineering* 9 (1): 143–60. 2022. <https://doi.org/10.1109/TNSE.2021.3055835>.
- Lim-Reinders, Stephanie, Brian M Keller, Shahad Al-Ward, Arjun Sahgal, and Anthony Kim. "Online Adaptive Radiation Therapy." *International Journal of Radiation Oncology* Biology* Physics* 99 (4): 994–1003. 2017.
- Liu, Jianhua, Peng Xie, Jijia Liu, and Xiaoguang Tu. "Task Offloading and Trajectory Optimization in UAV Networks: A Deep Reinforcement Learning Method Based on SAC and A-Star." *CMES-Computer Modeling in Engineering & Sciences* 141 (2). 2024.

- Liu, Yuankun, Dongxia Zhang, and Hoay Beng Gooi. "Optimization Strategy Based on Deep Reinforcement Learning for Home Energy Management." *CSEE Journal of Power and Energy Systems* 6 (3): 572–82. 2020.
- López-Andreu, Francisco Javier, Juan Antonio López-Morales, Zaida Hernández-Guillen, et al. "Deep Learning-Based Time Series Forecasting Models Evaluation for the Forecast of Chlorophyll a and Dissolved Oxygen in the Mar Menor." *Journal of Marine Science and Engineering* 11 (7): 1473. 2023. <https://doi.org/10.3390/jmse11071473>.
- Morgan, Howard E., and David J. Sher. "Adaptive Radiotherapy for Head and Neck Cancer." *Cancers of the Head & Neck* 5 (1): 1. 2020. <https://doi.org/10.1186/s41199-019-0046-z>.
- Nahom Abishu, Hayla, Abegaz Mohammed Seid, Aiman Erbad, Sergio Márquez-Sánchez, Javier Hernandez Fernandez, and Juan Manuel Corchado. "Multiagent DRL-Based Demand Response Optimization for IoT-Based Smart Home Energy Management Systems." *IEEE Internet of Things Journal* 12 (14): 27003–20. 2025. <https://doi.org/10.1109/JIOT.2025.3562137>.
- Ngo, T. T., M. Bellalij, and Y. Saad. "The Trace Ratio Optimization Problem." *SIAM Review* 54 (3): 545–69. 2012. <https://doi.org/10.1137/120864799>.
- Otis-Wilborn, Amy, Judith Winn, Alison Ford, and Maureen Keyes. "Standards, Benchmarks, and Indicators: Designing a Framework for Professional Development of Preservice and Practicing Teachers." *TEACHING Exceptional Children* 32 (5): 20–28. 2000. <https://doi.org/10.1177/004005990003200503>.
- Pesapane, Filippo, Caterina Volonté, Marina Codari, and Francesco Sardanelli. "Artificial Intelligence as a Medical Device in Radiology: Ethical and Regulatory Issues in Europe and the United States." *Insights into Imaging* 9 (5): 745–53. 2018. <https://doi.org/10.1007/s13244-018-0645-y>.
- Raymakers, J.A., M.M. Samson, and H.J.J. Verhaar. "The Assessment of Body Sway and the Choice of the Stability Parameter(s)." *Gait & Posture* 21 (1): 48–58. 2005. <https://doi.org/10.1016/j.gaitpost.2003.11.006>.
- Reynolds, Rebecca, Sarah Dennis, Iqbal Hasan, et al. "A Systematic Review of Chronic Disease Management Interventions in Primary Care." *BMC Family Practice* 19 (1): 11. 2018.
- Schaefer, A. L., N. Cook, S. V. Tessaro, et al. "Early Detection and Prediction of Infection Using Infrared Thermography." *Canadian Journal of Animal Science* 84 (1): 73–80. 2004. <https://doi.org/10.4141/A02-104>.
- Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. "Proximal Policy Optimization Algorithms." *arXiv Preprint arXiv:1707.06347*. 2017.
- Shin, Donghee. "The Effects of Explainability and Causability on Perception, Trust, and Acceptance: Implications for Explainable AI." *International Journal of Human-Computer Studies* 146 (February): 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>. 2021.
- Siboo, Sanjna, Anushka Bhattacharyya, Rashmi Naveen Raj, and SH Ashwin. "An Empirical Study of DDPG and PPO-Based Reinforcement Learning Algorithms for Autonomous Driving." *Ieee Access* 11: 125094–108. 2023.
- Sohaib, Rana Muhammad, Syed Tariq Shah, Oluwakayode Onireti, and Muhammad Ali Imran. "Harnessing DRL for URLLC in Open RAN: A Trade-off Exploration." *arXiv Preprint arXiv:2407.17598*. 2024.
- Tiase, Victoria L, William Hull, Mary M McFarland, et al. "Patient-Generated Health Data and Electronic Health Record Integration: A Scoping Review." *JAMIA Open* 3 (4): 619–27. 2021. <https://doi.org/10.1093/jamiaopen/ooaa052>.
- Timčenko, Valentina V. "Cloud-Based Dynamic Line Rating: Architecture, Services, and Cyber Security." In *Cyber Security of Industrial Control Systems in the Future Internet Environment*. IGI Global. 2020.
- Tosca, Elena Maria, Alessandro De Carlo, Davide Ronchi, and Paolo Magni. "Model-Informed Reinforcement Learning for Enabling Precision Dosing Via Adaptive Dosing." *Clinical Pharmacology & Therapeutics* 116 (3): 619–36. 2024. <https://doi.org/10.1002/cpt.3356>.
- Tseng, Huan-Hsin, Yi Luo, Sunan Cui, Jen-Tzung Chien, Randall K. Ten Haken, and Issam El Naqa. "Deep Reinforcement Learning for Automated Radiation Adaptation in Lung Cancer." *Medical Physics* 44 (12): 6690–705. 2017. <https://doi.org/10.1002/mp.12625>.
- Wang, Can, Jiaheng Zhang, Aoqi Wang, et al. "Prioritized Sum-Tree Experience Replay TD3 DRL-Based Online Energy Management of a Residential Microgrid." *Applied Energy* 368: 123471. 2024.
- Wang, Yixiao, Huaming Wu, Rutvij H Jhaveri, and Youcef Djenouri. "DRL-Based URLLC-Constraint and Energy-Efficient Task Offloading for Internet of Health Things." *IEEE Journal of Biomedical and Health Informatics* 28 (6): 3305–16. 2023.
- Wu, XiaoDan, RuiChang Li, Zhen He, TianZhi Yu, and ChangQing Cheng. "A Value-Based Deep Reinforcement Learning Model with Human Expertise in Optimal Treatment of Sepsis." *Npj Digital Medicine* 6 (1): 15. 2023. <https://doi.org/10.1038/s41746-023-00755-5>.

- Xu, Songlin, and Xinyu Zhang. “Augmenting Human Cognition with an AI-Mediated Intelligent Visual Feedback.” *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, April 19, 1–16. 2023. <https://doi.org/10.1145/3544548.3580905>.
- Yu, Chao, Jiming Liu, Shamim Nemati, and Guosheng Yin. “Reinforcement Learning in Healthcare: A Survey.” *ACM Computing Surveys (CSUR)* 55 (1): 1–36. 2021a.
- Yu, Chao, Jiming Liu, Shamim Nemati, and Guosheng Yin. “Reinforcement Learning in Healthcare: A Survey.” *ACM Computing Surveys (CSUR)* 55 (1): 1–36. 2021b.
- Zhong, Chongzhou, Mehdi Darbandi, Mohammad Nassr, Ahmad Latifian, Mehdi Hosseinzadeh, and Nima Jafari Navimipour. “A New Cloud-Based Method for Composition of Healthcare Services Using Deep Reinforcement Learning and Kalman Filtering.” *Computers in Biology and Medicine* 172: 108152. 2024
- Zhou, Tongxin, Yingfei Wang, Lu (Lucy) Yan, and Yong Tan. “Spoiled for Choice? Personalized Recommendation for Healthcare Decisions: A Multiarmed Bandit Approach.” *Information Systems Research* 34 (4): 1493–512. 2023. <https://doi.org/10.1287/isre.2022.1191>.

Biographies

Jannatul Ferdous Disha is pursuing B.Sc. Engineering degree in Industrial Engineering and Management from Khulna University of Engineering & Technology. Her research interests include smart manufacturing, green supply chain, lean manufacturing & six sigma, operation research, and Industry 4.0.

Kazi Md. Tanvir Anzum is a Lecturer in the Department of Industrial Engineering and Management at Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh. He earned his B.Sc. Engineering degree in Industrial and Production Engineering from KUET. Currently, he is pursuing a master’s degree in Industrial Engineering, further expanding his expertise in the field. Tanvir’s research interests include supply chain management, health systems engineering, operations research, lean manufacturing, and system development in healthcare. Enthusiastic about operations research, he actively collaborates with researchers worldwide, contributing to high-quality journal articles and conference papers. With a passion for integrating innovative methodologies into healthcare and industrial systems, he aims to enhance efficiency and sustainability across these domains.

Istiaq Mahmud is a undergraduate student in the Department of Industrial Engineering and Management at Khulna University of Engineering and Technology, Bangladesh. His research focus is data science, operations research, supply chain management and product design. Proficient in Solid-Works and Auto-Cad, he has designed a project on rotating bookshelf. He is also ardent that Industry 4.0 and additive manufacturing are implemented to encourage efficiency, sustainability, and innovations.

MD Faiaz Zaman Dehan is presently enrolled in B.Sc. Engineering degree in Industrial Engineering and Management from Khulna University of Engineering & Technology, Khulna, Bangladesh. He is interested in research in such areas as artificial intelligence, smart manufacturing, digital transformation, and quality management. He is proficient in Python and MATLAB and has been involved in projects in data-driven decision-making and predictive analyses of industrial systems. His interests lie in the pursuit of Industry 4.0 technologies; digital twin applications, and sustainable manufacturing processes to enhance efficiency, competitiveness and innovation in the developing economies.

MD. Mahfuzur Rahman Masum is an undergraduate in Department of Industrial Engineering and Management Khulna University of Engineering and Technology, Khulna in Bangladesh. His research interests include supply chain management & logistics, green supply chain, and operation research.