Date of publication December 1, 2023.

Reinforcement Learning for Optimizing Delivery Paths in Hospital Settings: A Review

AAILA ARIF[1], AND JIACUN WANG[2]
[1]The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[2]The Macao Institute of Systems Engineering, Macau University of Science and Technology, Macao 999078, China
[1]Honors School, Monmouth University, W. Long Branch, New Jersey, USA
[2]Department of Computer Science and Software Engineering, Monmouth University, W. Long Branch, New Jersey, USA.

Corresponding author: Jiacun Wang (e-mail: jwang@monmouth.edu).

"This work was partially supported by 0093/2023/RIA2 the Science and Technology Development Fund, Macau SAR under Grant 0050/2020/A1."

ABSTRACT Reinforcement learning is a branch of machine learning that facilitates the interaction of autonomous agents with their environments. This is done by "teaching" an agent efficient decision-making through iterative processes of exploration and trial-and-error. This review article focuses on the application of reinforcement learning within the healthcare industry. We review recent publications that address the optimization of the pickup and delivery processes for essential supplies and medications with mobile robots, and the integration of these two key technologies to improve hospital operations efficiency. We also investigate the gap between research results and real-world applications, and point out directions for future work.

INDEX TERMS Reinforcement learning, robot navigation, healthcare, medication delivery.

I. INTRODUCTION

This review article focuses on the application of reinforcement learning (RL) within the healthcare industry. The goal is to create a model that improves the decision-making of autonomous agents in a simulated hospital environment. In this context, we review recent publications that address the optimization of the pickup and delivery processes for essential supplies and medications within hospital settings.

The significance of this work lies in its potential to address critical challenges within the healthcare industry. Refinement learning techniques make it possible to improve efficiency in hospital operations by developing an autonomous agent capable of optimizing pickup and delivery processes. By automating such processes, a positive impact could be achieved on both the patient and caregiver well-being. The research was relevant in the context of applying machine learning techniques to solve real-world problems [1] [2], which would have abundant benefits for both healthcare providers and patients. Therefore, the significance of this research lay in its potential to bring about positive and practical changes in the healthcare sector.

The rationale for this study came from the knowledge gap that exists within the field of healthcare robotics, specifically the optimization of processes within hospital systems through the use of reinforcement learning. Common challenges faced in hospitals included the shortage of personnel, workplace hazards, and crowded facilities [3] [4] [5]. In particular, nurse burn-out has been a great concern [6], [7]. These challenges can be addressed by creating robotic agents capable of navigating through a hospital, in particular patient rooms in a hospital, and performing tasks that were repetitive or dangerous to do in certain scenarios, such as administering medicine.

The distribution of hospital inpatient rooms depends on several factors, including the size of the hospital, specialty areas, patient demographics, and operational considerations. Hospitals typically have multiple floors, each dedicated to specific departments or medical specialties. Inpatient rooms are distributed on these floors based on the types of patients they serve. For example, surgical inpatient units may be located on one floor, while medical inpatient units are on another. Inpatient rooms are typically located near essential hospital services, such as nursing stations, medication rooms, imaging facilities, and operating rooms. This proximity ensures quick access to necessary resources and medical staff.

Fig. 1 illustrates a simple layout of the inpatient rooms in

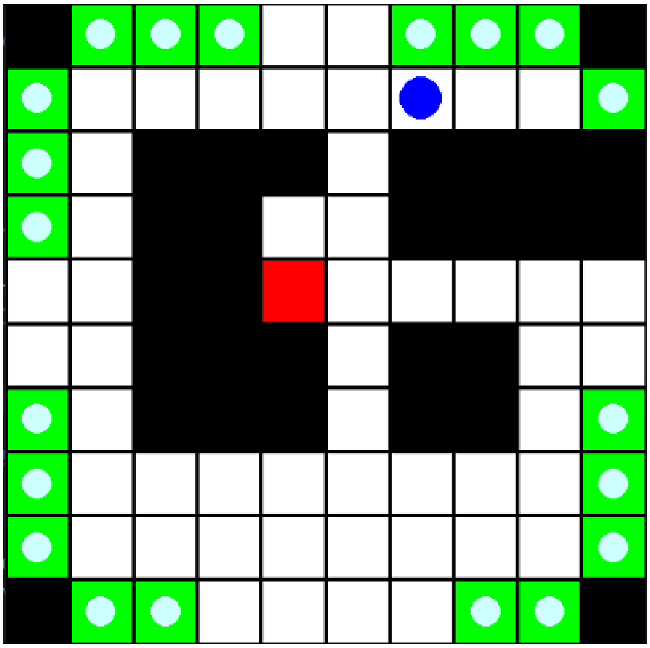


FIGURE 1. Distribution of patient rooms. The green boxes represent patient rooms, the white boxes represent hallways, the red box represents the robot's dock, and the black boxes represent areas inaccessible to the robot.

a hospital in one floor. The grid is 10 x 10 with each box representing one of the following: hallway, hospital room, inaccessible area, or docking station. The green boxes represent patient rooms. These are areas the agent will need to visit for deliveries. The red box represents the robot's dock, where it will return to after completing its route. The black boxes represent areas inaccessible to the robot, such as offices and closets. The white boxes represent hallways, where the robot can move freely on its delivery path. The agent is represented as the blue circle. The white circles are placeholders for the deliveries the robot is yet to finish. When the agent visits a room, it "picks up" the room's delivery circle to show that the delivery was made. Since there would not be doors between patient rooms themselves, the robot is unable to pass through from one green box to another directly. It must return to the whitespace of the hallway before entering another room. Of course, a real-world hospital patient room layout would be much more complex than this figure shows, not to mention that there are always patients and healthcare staff moving along the hallways.

In order to understand the status of research in the area, a comprehensive review of the literature that underpins the field of reinforcement learning and its applications in the healthcare sector is imperative. The subject of utilizing machine learning and algorithms within hospital systems to optimize the pickup and delivery of essential supplies and medications to improve the well-being of both the patient and the healthcare provider was approached through reinforcement learning, the Q-learning algorithm and the Python coding language. This review of the literature predominantly centers on an explanation of the task, the significance of such research, the rationale behind choosing reinforcement learning coupled with Q-Learning, applications of AI models and robots in healthcare settings, and gaps in research that were addressed with the work.

The selection of resources was done in a series of steps including a search of scholarly and peer-reviewed articles in digital libraries with filters based on the topic, a thorough examination of each of the papers, and a final selection of research papers. In the first step, ScienceDirect, IEEE Xplore, and Monmouth University's Hawkfind were utilized to select sources based on the topics of reinforcement learning, Q-learning, robotics and AI in healthcare, and gaps that existed in the research being done. In the second step, the papers were read through entirely to discover common themes that could be established between multiple papers. In the last step, a final selection of papers was conducted based upon sources that provided data pertaining to the project being done as well as containing lines of reasoning that could be linked to multiple other articles.

The rest of the paper is organized as follows: Section II provides a concise overview of the literature search criteria and selection methodology. Section III examines recent advancements in reinforcement learning, robotic navigation, and problem-solving applications within healthcare settings. Section IV explores key challenges and proposes potential solutions.

II. REINFORCEMENT LEARNING

Reinforcement learning, a branch of machine learning, is an approach in which an intelligent agent is trained to perform certain tasks [8]. This entity, called an agent, must interact with its environment and discover the most optimal route to accomplish its tasks [8] [9]. The agent(controller) undergoes positive or negative reinforcement in response to an action, accessed through the reward and punishment functions [10]. The agent operates in an unknown environment, observing its state at each time step and obtaining a reward or punishment after executing a chosen action [11]. The policy, which maps states to actions, is pivotal [12]. Reinforcement learning aims to continuously identify and develop the most efficient policy to maximize cumulative future rewards. One commonly adopted method to establish a connection between actions and rewards involves learning the anticipated quality of actions in a given state, referred to as Q-Learning [13]. Many RL algorithms employed this method to create a state-transition system to simulate potential agent movements. Following each movement, the RL agent accrues a numerical reward, either a positive or negative number. During the agent training phase, the Q values are computed and stored as the agent systematically explores the environment, creating estimates of immediate and long-term action value. This learning process directs the agent to select actions which will lead to the highest rewards and thus too proficient task completion. Based on the research that is done on reinforcement learning, its use in the project provided an efficient way to solve the problem of optimizing supply delivery in hospitals.

The foundation of Q-learning process is a Markov decision process (MDP), where the states represent the problem configurations, actions correspond to decisions that change the state, and rewards indicate the quality of the solutions obtained. By framing optimization problems in this way, RL algorithms can learn to navigate the solution space effectively.

A. MATHEMATICAL MODEL

Q-Learning process can be specified by:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha[r_{t+1} + \gamma max_a Q(s_{t+1},a_{t+1}) \\ -Q(s_t,a_t)]$$

where $Q(s_t,a_t)$ is the Q-value for taking action a_t in state s_t, r_t is the instant reward, α is the learning rate, and γ is the discount factor, which reflects the importance of future reward. The objective function of Q-Learning is to find the policy that maximize the expectation of cumulative reward:

$$max_\pi E[\sum_{t=0}^{T} \gamma^t r(s_t,a_t,s_{t+1})|\pi]$$

Q-Learning is a special type of Temporal-Difference (TD) learning [14], based on Bellman equation. In RL, the state value function $V_\pi(s)$ estimates the value of the current state. State value $V(s)$ is related to the value of the current state as well as all future states. The cumulative reward expectation weighs in the calculation of a state value s [15]. The corresponding Bellman equation of the state value function is as follows:

$$V_\pi(s) = E_\pi[r_{t+1} + \gamma V(s_{t+1})|s_t = s]$$

where γ is the attenuation coefficient or discount factor. It reflects the degree that the corresponding agent values the future states. The state action-value function of the cumulative optimal value function $V^*\ (s)$ and $Q(s,a)$ can be expressed as:

$$V^*\ (s) = max_\pi E[\sum_{t=0}^{H} \gamma^t R(S_t,A_t,S_{t+1})|\pi, s_0 = s]$$

$$Q_\pi(s,a) = E_\pi[r_{t+1}+\gamma r_{t+1}+\gamma^2 r_{t+1}+\cdots|A_t = a, S_t = s]$$

$$V(s_t) \leftarrow V(s_t) + \alpha[r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

Among them, $r_{t+1} + \gamma V(s_{t+1})$ is the TD objective function, while $r_{t+1} + \gamma\ V(s_{t+1}) - V(s_t)$ is the TD deviation. α is the rewarding decay coefficient of the decay rate γ. According to the updated formula of TD(0), we adopted, the Q value can be derived to obtain the updated formula of Q-Learning, which is also the Q-Learning mentioned above definition:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha[r_{t+1}+ \\ \gamma max_{a_{t+1}} Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)]$$

B. BASIC ELEMENTS FOR PATH OPTIMIZATION

1) Environment

The RL environment is composed of the learning agent, the state set S and action set A, and a few parameters such as the attenuation factor γ. Given strategy π, the algorithm goal is to solve the state value function $v(\pi)$. RL is a trial-and-error learning in the environment, in which the experience is gained through interaction with the environment. In the example study, the environment is an adaptation of OpenAI's GridWorld. GridWorld is one of the environments from OpenAI's Gymnasium library for RL-related projects. This interpretation of GridWorld was modeled after a potential hospital floor plan, including a nurse's station, where the robot is dispatched, hospital rooms, and space for private offices and supply rooms.

2) States

For different applications, the definition of states is different. As for the example mentioned above, the state is defined as the coordinate of the cell where the robot is located.

$$s \in \{i,j\}, i = 1,2,...,n, j = 1,2,...,n$$

3) Actions

Actions trigger transitions between states, and the agent transfers from one state to the next through the action. Without any restrictions, a robot can enter into any of its four neighboring cells, or stop and drop medications.

$$a \in \{forward, left, right, back, stop\}$$

Aside from $stop$, each action changes the state by incrementing or decrementing either i or j. Each time an action occurs, the agent receives a reward for its decision.

4) Rewards

The reward in RL is generally set according to the target value. To minimize the overall medication delivery time, rewards are set up as follows: -1 for each step taken by the agent, -1000 for hitting an obstacle or attempting to move through walls, 1000 for performing a medication drop-off, and 100 for returning to its starting place. The final rewards metric at the end of an agent's testing indicates how well the agent learns the environment under the applied parameters.

C. DEEP REINFORCEMENT LEARNING

Deep reinforcement learning is a subset of machine learning and reinforcement learning that combines the the principles of reinforcement learning with deep learning techniques [16] [17]. While reinforcement learning algorithms learn to make decisions through interactions with the environment and gain rewards and punishments based on actions taken, deep reinforcement learning builds on this by utilizing deep neural networks to represent complex functions [18] [19]. These deep neural networks convert states into action values. Reinforcement learning then takes these action values and executes an action that relates to it. There are risks associated with the deployment of robots in real-world systems, such as damage to the robot through risky actions. Deep reinforcement learning decreases the amount of deployments needed to establish the policy and train, which is beneficial when it comes to costs associated with training the agent and avoiding accidents.

D. REINFORCEMENT LEARNING VS. ALTERNATIVE MODELS

There are three main branches of machine learning - reinforcement learning, supervised learning, and unsupervised learning [20]. Reinforcement learning maximizes the reward signal by mapping environment states to actions. This type of learning prioritizes a high reward through a trial-and-error approach but is somewhat delayed due to the exploration and training phase. Supervised learning trains a model on a labeled data set consisting of pairs of input and output data [21] [22]. By allowing the model to map known inputs to actions, it aims to train the model to make predictions on unseen data. Unsupervised learning aims to train a model to discover patterns and hidden structures in datasets that are unlabeled [23] [24]. Since there are no explicit labels, the algorithm must create groupings with the input data. For this reason, reinforcement learning was determined to be the optimal model to implement due to various reasons - versatility, adaptability to dynamic environments, and environmental interaction. When reinforcement learning agents learn policies through exploration and trial and error, they can apply these to handle a variety of tasks within an environment [25] [26]. In the case of supervised learning, since the models are given defined training data, they would be unable to adapt to tasks that go far beyond the scope of the data. Unsupervised learning models focus more on uncovering patterns within data rather than task-specific generalizations [27]. In terms of adapting to changing environments, reinforcement learning proves to be the most effective due to the exploratory and continuous nature of the model [28].

Since a reinforcement learning agent continuously learns about the environment and the route that is the most optimized, it would be able to adjust its behavior based on a changing environment. This is not the case for supervised learning, where any environment that differs greatly from the training data would cause the performance of this model to degrade significantly. Similarly, unsupervised models are sensitive to changes in data. In the context of hospital systems that are constantly undergoing changes, or if the robots are to be relocated to a completely different hospital, reinforcement learning would be the most responsive to the changes in environment. Lastly, reinforcement learning agents learn by interacting with the environment to gain either rewards or punishments [13]. During the exploration and training phase, the approach of trial-and-error allows the agent to make decisions and discover the most optimized route. The interactive learning and decision-making aspect of reinforcement learning is lost with both supervised and unsupervised learning [20]. Its ability to be versatile, adaptable to dynamic environments, and ability to continuously learn and improve demonstrates why it was the most effective approach [29].

A relevant approach to finding an optimal solution to a problem is intelligent search, which refers to algorithms that systematically explore a problem space to find optimal or satisfactory solutions, often using heuristics or learned guidance. Many such algorithms have been developed over the past decades [30] [31] [32]. Table 1 provides a comparison between intelligent search and reinforcement learning. In general, reinforcement Learning excels in dynamic, uncertain environments but comes with higher computational complexity due to iterative learning and large state spaces. On the other hand, intelligent Search is more deterministic and controllable, often yielding faster solutions when heuristics are well-designed, but may struggle in environments requiring adaptation or learning. Optimization can also be done through formal modeling and analysis [33] [34].

TABLE 1. Comparison between Reinforcement Learning and Intelligent Search

Aspect	**Reinforcement Learning (RL)**	**Intelligent Search**
Learning	Learns from interaction and feedback	May use heuristics or learned guidance
Goal	Maximize cumulative reward	Find optimal path or solution
Adaptability	Highly adaptive over time	Depends on search strategy
Exploration	Balances exploration vs. exploitation	Explores based on heuristics or rules
Use of Memory	Learns from past experiences	May store past paths or hints
Typical Output	Policy or value function	Solution path or best match

In terms of optimal path optimization, A* search is a powerful and widely used pathfinding algorithm in computer science and artificial intelligence [35]. It is based on graph traversal and especially popular in games, robotics, and navigation systems because it finds the shortest path efficiently by combining the strengths of two other algorithms: Dijkstra's algorithm and Greedy Best-First Search. However, A* search is only applicable to deterministic search problem, in which a full map of the search area is required. On the other hand, reinforcement learning can explore unknown area through interaction with the environment. Considering patients and medical staff are constantly moving here and there, a hospital is an ever-changing environment, and thus A* search is not a good option for solving the delivery path optimization problem under discussion.

III. ROBOTICS IN HEALTHCARE

In recent decades, substantial advances have been made in technology in the fields of computer science, machine learning, and robotics. Robotics, in particular, has been one of the most rapidly evolving technologies that is ushering in a new era of possibilities across numerous industries such as healthcare, military, entertainment, and others [4] [22]. In the healthcare sector, robotics finds applications extending to various roles including caregivers (nurses and physicians), hospital services, delivery, and beyond [3] [36].

A. THE NECESSITY

The necessity of integrating robotics into healthcare has become more apparent in recent years. As challenges and workplace hazards continuously arise in hospitals, the well-being of both caregivers and patients are negatively impacted [3] [4] [37] [38]. This manifests in heightened costs for hospitals, increased stress levels, and diminished quality of life [4]. The global challenges presented by the COVID-19 pandemic further underscored the critical need for technological solutions when issues such as increased risk of contracting the disease, lack of healthcare centers and services and lack of staff were revealed. Healthcare workers are continuously exposed to occupational hazards such as bloodborne pathogens, infections, and other healthcare associated diseases [38]. The persistent issue of staff shortages, particularly in nursing, have exacerbated the problem [4]. This leads to the remaining nurses, doctors, and other caregivers shouldering extended and exhausting work hours, as they must manage more patients while working considerably longer shifts. This scenario directly contributes to fatigue and burnout, which can pose potential repercussions on patient/caregiver health and the quality of care provided. As advancements in technology and medicine continue to progress, an ever-aging population and escalating number of patients imposes an additional burden on hospitals. The increasing patient influx drains hospital resources, leading to crowded facilities, stretched staffing levels, and competition for essential services and equipment [4]. As a consequence, patients must endure longer waiting times, delayed treatments, and a general decrease in quality of care. For healthcare professionals, an increasing patient load causes increased work pressure and prolonged working hours. Continually participating in repetitive tasks, such as the distribution of medications in expansive hospitals, places a physical strain on the staff [5]. As a result, a decline in the mental and physical well-being of the caregiving workforce is inevitable. Patients experience the negative implications of these challenges, and reduced personalized attention, inability to receive resources, and prolonged waiting times can potentially lead to distrust in the healthcare system and long-term consequences in patient health. A common theme established in the reviewed literature demonstrates the necessity of integrating robots into hospital operations. In order for hospitals to address the issues raised by aging populations, increasing patient load, staffing shortages, health risks, and other concerns, hospitals must incorporate robots capable of navigating optimized routes for efficient resource delivery to patients.

B. BENEFITS OF ROBOTIC AUTOMATION

The benefits of robotic automation in hospital settings are numerous. For example, robotic automation can streamline repetitive processes such as patient scheduling, billing, and claim processing, freeing up staff to focus on patient care. It can also help reduce human error in data entry, medication management, and record-keeping, leading to safer and more reliable care. By cutting down on labor-intensive tasks, reducing operational costs and allowing resources to be redirected to critical care areas. Automation will also result in faster service, fewer delays, and more accurate information contribute to higher patient satisfaction.

Hojjat et al. present our recent research on integrating artificial emotional intelligence in a social robot, named Ryan, and studies the robot's effectiveness in engaging older adults [39]. Robot systems are employed in home care. To increase task execution efficiency, Zhu et al. propose a smart home system architecture that integrates a mobile robot with better event perception and task execution performance [40]. The adoption of social robots, an emerging field of significance as these technologies become more ingrained in daily life, was investigated in [41], and revealed that human-like qualities such as appearance and behavior play a vital role in perceived enjoyment and social attraction. Considering the ever-increasing aging population in China, Yu et al. introduced a massage robot, designed to replicate the seated knee adjustment manipulation, a specific traditional Chinese medicine technique [42]. In [43], Chang et al. presented an interactive healthcare question answering system that uses attention-based models to answer healthcare-related questions. The system employs attention-based transformer models to efficiently encode semantic meanings and extract the medical entities inside the user query individually.

In terms of robot navigation, Shi et al. [44] proposed an end-to-end navigation planner that translates sparse laser ranging results into movement actions. The agents trained by simulation agents can be extended to the real scene for practical application. A study on human-awareness of social robot navigation is presented in [45]. In [46], Zhu et al. conducted a survey on robot navigation based on reinforcement learning. The hospital is a complex environment with patients and medical staff constantly moving here and there. Zhao et al. [47] present an improved reinforcement learning-based algorithm for local path planning that allows robots to perform well when there are more dynamic obstacles. In Day2024, Day et al. introduce a new human-robot interaction dataset focusing on interactions between humans and small differential drive robots running different types of controllers. In [48], Mulvey et al. propose the novel design of a deformable mobile robot. A deformable mobile robot can adopt a wider or narrower stance to fit the environment.

C. SERVICE ROBOTS IN HOSPITALS

Nowadays, service robots are seen in hospital corridors. Robots in nursing can handle less technical tasks such as patient transport and rehabilitation activities. This support allows caregivers to focus on less strenuous nursing duties and more direct patient care [49]. In [50], the authors review the benefits of the use of social robots to patients, healthcare workers, customers, and organizations during the COVID-19 pandemic, provide a view of the emerging focal issues for healthcare services, such as logistics of patients and supplies. Considering the rapidly aging population and pressure on healthcare services, Broadbent et al. provide review on human responses to healthcare robots and summarizes the variables that have been found to influence responses in [51]. In [52], Wang et al. presented a portable back massage robot which can complete the massage operations such as tapping, kneading and rolling was designed to improve the level of intelligence and massage effect. In the case of infectious diseases, medical staff should refrain from directly contact patients, service robots can play a big role [53], [54]. In [55], Shen et al. provide a survey based on over 200 reports covering robotic systems which have emerged or have been repurposed during the past several months, to provide insights to both academia and industry. A smart sterilization robot system is also developed to spray disinfectants in operating theaters or patients' rooms, designed according to the results of controlled experiments and the requirements for hospital disinfection in [56]. In [57], a review by the authors found that robots have played 10 main roles across a variety of clinical environments, with the two predominant roles being surgical and rehabilitation and mobility.

When considering current delivery robots, such as TUG mobile robots, the advantages of using reinforcement learning become apparent [58] [19]. TUG robots consist of a battery, load and carrying modules, and a control unit with features such as hazard detection, door and elevator opening skills, and "speaking" ability. The applications of this delivery robot include delivering resources such as medications, food service, and loading and unloading carts of medical supplies. Some instances where the TUG delivery robot proved effective in hospital systems include its use in El Camino Hospital in California and Children's Hospital in Boston. In El Camino, once hospital management realized that the distances between departments were large and would cause increased costs and wasted time in delivery of supplies such as food, linens, and medicine, the TUG system was installed. With these robots, 80 percent of the delivery within that hospital system was able to be automated, saving the hospital considerable amounts of money.

D. REINFORCEMENT LEARNING FOR ROBOT NAVIGATION

Reinforcement learning (RL) has become a powerful approach for robot navigation, enabling robots to learn optimal movement strategies through trial and error in complex environments. Key techniques include deep Q-networks (DQN), which learns value functions for discrete actions; proximal policy optimization (PPO), which balances exploration and exploitation for continuous control; and twin delayed deep deterministic policy gradient (TD3), which is used for stable learning in continuous action spaces. Mobile robot navigation normally is normally composed of path planning and obstacle avoidance.

Many research results have been reported in this area in recent years. Zhu and Zhang provide a comprehensive review on deep reinforcement learning-based mobile robot navigation in [59], which discuss four typical application scenarios: local obstacle avoidance, indoor navigation, multi-robot navigation, and social navigation. In [60], a review on bioinspired robots (BIRs), which can learn to locomote or produce natural behaviors similar to animals and humans, is conducted. In [61], a multimodal locomotion framework is presented that is composed of a hand-crafted transition motion with a learning-based bipedal controller. A new a new local navigation method for steering the robot to global goals without relying on an explicit map of the environment is presented in [62], in which the model is trained based on the Advantage Actor–Critic method. Similar works are presented in [63] [64], where a reinforcement learning-based path generation approach is used for mobile robot navigation without a prior exploration of an unknown environment. In [65], autonomous navigation task in large-scale environments with crowded static and dynamic objects using graph relational reinforcement learning.

IV. DISCUSSIONS

A. THE "REALITY GAP"

Although previous research has been done in the realm of reinforcement learning in robotics, notable gaps still remain. One such issue is the 'reality gap' that exists between the virtual and the real world [66] [67]. Due to the disparities between complex systems in the real world and the virtual world, developing reinforcement agents is a difficult task. The performance of reinforcement learning agents trained in virtual settings decreases as the real-world system with which the robot will interact becomes more complex. Since research on lessening the gap between the virtual real world is still being conducted, a common method is used, called 'domain randomization'. This method aims to introduce changes into the environment by randomizing or altering certain elements rather than training in a static environment. This method addresses the challenge of the agent to adapt to a dynamic environment, which is one of the largest issues pertaining to the "reality gap". The project aimed to further the bridge of the reality gap using both randomization and simulation of the hospital floor with utmost precision. During each run of the program, the user was prompted to enter x and y values for coordinates of obstacles that were placed into the virtual environment. By allowing the user to pick where the obstacles were being placed with each run of the program, the agent was trained in a dynamic environment. An additional strategy that was incorporated into the program was virtual hospital walls to simulate the layout of real hallways in hospital systems. With this, the real world was simulated as closely as possible in order to bridge the gap between both domains.

B. IMPLEMENTATION CHALLENGES IN HOSPITALS

Another gap being addressed by researchers is the failure of most hospitals to incorporate fully functional AI systems [68]. Although service robots have increasingly been researched and incorporated into many areas in recent years due to their various benefits, their implementation and scaling are still difficult obstacles to overcome. Only a small group of hospitals that implemented these robots were able to keep utilizing them. This is in part due to challenges faced when placing the robots into real-world environments that are constantly changing and the expensive programming of such robots. Previously, hospitals would resort to robots that required a map of the environment and large amounts of manual coding to perform tasks such as delivery, food service, and more [69]. In the case of TUG robots, this strategy seemed to work. However, when it comes to environments that are constantly changing or when the robots are moving to a whole different environment, this method becomes less effective. The current project tackled these challenges by implementing a mix of Q-Learning and reinforcement learning.

C. ADDRESSING CONCERNS

Robots are proficient in their abilities to take over simple tasks once handled by humans, such as assembly jobs and delivery services [70]. With recent advancements in technology, the capabilities of robots skyrocketed, and more complex physical and cognitive jobs can be carried out by robots, such as identifying signs of dementia in patients as well as detecting hazards in stores. However, with the increased integration of AI and robotics into various industries, potential concerns have been posed by those working alongside them [69]. According to a report in 2018, 30-65 percent of jobs face the potential of automation, putting them at risk of being replaced by robots [71]. For those whose jobs are included in this statistic, they may experience "service robot risk awareness". This means that job insecurity is instilled into those whose industries introduce and adopt service robots [71]. The workplace is undergoing rapid transformations, and a growing number of employees have expressed concerns with keeping up with robots that could replace them [72]. While the displacement of jobs appears to be a tangible risk with the increase in robotics and AI in a plethora of sectors, an opportunity arises for the creation of new jobs [70]. Jobs such as managing and maintaining such technology as well as developing it can create new positions to be filled while improving efficiency in service roles.

Another issue posed by the adoption of robots into industries is the frustration that is caused by lack of autonomy in jobs that can be automated [71]. However, this issue can be mitigated with collaborative robots that enhance the effectiveness and safety of employees. Robots lack emotional intelligence as well as social skills, and this is where the opportunity for collaboration between robots and humans appears. In jobs characterized by the necessity of a large social presence as well as complex emotional responses, the need for such collaboration becomes obvious [71]. While robots and AI handle more routine tasks, professionals can focus on more engaged and personalized care [72]. A prevalent theme in the literature reviewed in this section is the positive impact that is offered by the integration of robots and AI into industries such as healthcare.

While job displacement and the fears and frustrations of employees are issues that must be addressed, there are numerous benefits that can be reaped through the collaborative effort between robots and humans. Also, the new job opportunities that are created, along with the increased safety and efficiency of workers, demonstrates the positive trajectory of the evolving workplace [72]. The current project aimed to utilize a mix of reinforcement learning with Q-learning to train an agent to navigate through a virtual hospital floor. When the program is implemented in real life, a service robot can use the same programming to carry out various tasks and reap the benefits mentioned.

V. CONCLUSIONS

This review of the literature explores the conceptual foundation, methodology, and applications of reinforcement learning in the domain of healthcare robotics. It demonstrates the importance of utilizing machine learning algorithms to streamline and automate operations in hospital settings, such as essential resource delivery. The choice of applying reinforcement learning to approach the route optimization problem in hospitals was justified through a comparison with alternative methods. The adaptability, versatility, and continuous learning capabilities of reinforcement learning are highlighted in this comparison, making it apparent that this solution is the most effective. The evolving role of robots is also discussed in the review, as the many challenges faced by hospitals can be mitigated through the use of service robots trained with reinforcement learning. Research gaps, such as the "reality gap" are addressed and solved through domain randomization and non-static programmable environments. Additionally, ethical concerns such as job displacement and frustration are prioritized, and solutions to these issues are presented.

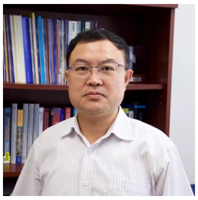
REFERENCES

[1] H. Han, W. Li, J. Wang, G. Qin, and X. Qin, "Enhance explainability of manifold learning," Neurocomputing, vol. 500, pp. 877–895, 2022.
[2] L. Zheng, J. Wang, A. Sheriff, and X. Chen, "Hospital length of stay prediction with ensemble methods in machine learning," in 2021 International Conference on Cyber-Physical Social Intelligence (ICCSI), 2021, pp. 1–5.
[3] I. R. Defi et al., "Healthcare workers' point of view on medical robotics during covid19 pandemic: A scoping review," International Journal of General Medicine, vol. 15, pp. 3767–3777, 2022.
[4] N. K. Lakshmi et al., "Fuzzy guided autonomous nursing robot through wireless beacon network," Multimedia Tools and Applications, vol. 81, no. 3, pp. 3297–3325, 2022.
[5] J. Holthöfer and J. van Doorn, "Robots do not judge: Service robots can alleviate embarrassment in service encounters," Journal of the Academy of Marketing Science, vol. 51, pp. 767–784, 2023.
[6] C. Maslach and M. P. Leiter, "Understanding the burnout experience: recent research and its implications for psychiatry," World Psychiatry, vol. 15, no. 2, pp. 103–111, 2016.
[7] M. K. Shah, N. G. J. P. Cimiotti, N. Ghose, M. Moore, and M. K. Ali, "Prevalence of and factors associated with nurse burnout in the us," JAMA Network Open, vol. 4, no. 2, p. e2036469, 2022.
[8] S. Spanó et al., "An efficient hardware implementation of reinforcement

learning: The q-learning algorithm," IEEE Access, vol. 7, pp. 186 340–186 351, 2019.
[9] L. M. D. Da Silva, M. F. Torquato, and M. A. C. Fernandes, "Parallel implementation of reinforcement learning q-learning technique for fpga," IEEE Access, vol. 7, pp. 2782–2798, 2019.
[10] S. G. Khan, G. Herrmann, F. L. Lewis, T. Pipe, and C. Melhuish, "Reinforcement learning and optimal adaptive control: An overview and implementation examples," Annual Reviews in Control, vol. 36, no. 1, pp. 42–59, 2012.
[11] J. Gu, J. Wang, X. Guo, G. Liu, S. Qin, and Z. Bi, "A metaverse-based teaching building evacuation training system with deep reinforcement learning," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 53, no. 4, pp. 2209–2219, 2023.
[12] H. Shi, G. Liu, K. Zhang, Z. Zhou, and J. Wang, "Marl sim2real transfer: Merging physical reality with digital virtuality in metaverse," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 53, no. 4, pp. 2107–2117, 2023.
[13] M. R. Bonyadi, R. Wang, and M. Ziaei, "Self punishment and reward backfill for deep q-learning," arXiv.org, 2022.
[14] S. Cayci, S. Satpathi, N. He, and R. Srikant, "Sample complexity and overparameterization bounds for temporal-difference learning with neural network approximation," IEEE Transactions on Automatic Control, vol. 68, no. 5, pp. 2891–2905, 2023.
[15] X. Guo, Z. Bi, J. Wang, S. Qin, S. Liu, and L. Qi, "Reinforcement learning for disassembly system optimization problems: A survey," International Journal of Network Dynamics and Intelligence, 2023.
[16] Y. Matsuo, Y. LeCun, M. Sahani, D. Precup, D. Silver, M. Sugiyama, E. Uchibe, and J. Morimoto, "Deep learning, reinforcement learning, and world models," Neural Networks, vol. 152, pp. 267–275, 2022.
[17] M. Yang, G. Liu, Z. Zhou, and J. Wang, "Partially observable mean field multi-agent reinforcement learning based on graph attention network for uav swarms," Drones, vol. 7, no. 7, 2023.
[18] Z. Bi, X. Guo, J. Wang, S. Qin, L. Qi, and J. Zhao, "A q-learning-based selective disassembly sequence planning method," in 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2022, pp. 3216–3221.
[19] O. Gottesman, F. Johansson, and M. Komorowski, "Guidelines for reinforcement learning in healthcare," Nat Med, vol. 25, p. Nat Med, 2019.
[20] J. M. D. Delgado and L. Oyedele, "Robotics in construction: A critical review of the reinforcement learning and imitation learning paradigms," Advanced Engineering Informatics, vol. 54, 2022.
[21] B. Hu and J. Wang, "Deep learning based hand gesture recognition and uav flight controls," International Journal of Automation and Computing, vol. 17, no. 1, pp. 17–29, 2020.
[22] W. Zhang, J. Wang, and F. Lan, "Dynamic hand gesture recognition based on short-term sampling neural networks," IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 1, pp. 110–120, 2020.
[23] X. Heng, Z. Wang, and J. Wang, "Human activity recognition based on transformed accelerometer data from a mobile phone," International Journal of Communication Systems, vol. 29, no. 13, pp. 1981–1991, 2016.
[24] Z. Zhang, J. Liu, G. Liu, J. Wang, and J. Zhang, "Robustness verification of swish neural networks embedded in autonomous driving systems," IEEE Transactions on Computational Social Systems, vol. 10, no. 4, pp. 2041–2050, 2023.
[25] K. Zhang, W. Lou, J. Wang, R. Zhou, X. Guo, Y. Xiao, C. Shi, and Z. Zhao, "Pa-detr: End-to-end visually indistinguishable bolt defects detection method based on transmission line knowledge reasoning," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1–14, 2023.
[26] H. Yang, J. Wang, and J. Wang, "Efficient detection of forest fire smoke in uav aerial imagery based on an improved yolov5 model and transfer learning," Remote Sensing, vol. 15, no. 23, 2023.
[27] H. Han, Y. Wu, J. Wang, and A. Han, "Interpretable machine learning assessment," Neurocomputing, vol. 561, p. 126891, 2023.
[28] J. Xue, Z. Liu, G. Liu, Z. Zhou, K. Zhang, Y. Tang, and J. Wang, "Robust wind-resistant hovering control of quadrotor uavs using deep reinforcement learning," IEEE Transactions on Intelligent Vehicles, pp. 1–10, 2023.
[29] Z. Bi, X. Guo, J. Wang, S. Qin, and G. Liu, "Deep reinforcement learning for truck-drone delivery problem," Drones, vol. 7, no. 7, 2023.
[30] B. Fei, W. Bao, X. Zhu, D. Liu, T. Men, and Z. Xiao, "Autonomous cooperative search model for multi-uav with limited communication network," IEEE Internet of Things Journal, vol. 9, no. 19, pp. 19 346–19 361, 2022.
[31] H.-M. Yoo, J.-M. Moon, J. Na, and E.-K. Hong, "User association and load balancing based on monte carlo tree search," IEEE Access, vol. 11, pp. 126 087–126 097, 2023.
[32] X. Cui, X. Guo, M. Zhou, J. Wang, S. Qin, and L. Qi, "Discrete whale optimization algorithm for disassembly line balancing with carbon emission constraint," IEEE Robotics and Automation Letters, vol. 8, no. 5, pp. 3055–3061, 2023.
[33] J. Wang and J. Wang, "Real-time adaptive allocation of emergency department resources and performance simulation based on stochastic timed petri nets," IEEE Transactions on Computational Social Systems, vol. 10, no. 4, pp. 1986–1996, 2023.
[34] J. Wang, "Patient flow modeling and optimal staffing for emergency departments: A petri net approach," IEEE Transactions on Computational Social Systems, vol. 10, no. 4, pp. 2022–2032, 2023.
[35] D. Delling, P. Sanders, D. Schultes, and D. Wagner, "Engineering route planning algorithms," in Algorithmics of Large and Complex Networks: Design, Analysis, and Simulation, ser. Lecture Notes in Computer Science. Springer, 2009, vol. 5515, pp. 117–139.
[36] M. Q. Mohammed et al., "Review of learning-based robotic manipulation in cluttered environments," Sensors (Basel, Switzerland), vol. 22, no. 20, 2022.
[37] L. A. Kelly, P. M. Gee, and R. J. Butler, "Impact of nurse burnout on organizational and position turnover," Nursing Outlook, vol. 69, p. 96–102, 2021.
[38] M. Schüle, J. M. Kraus, F. Babel, and N. Reißner, "Patients' trust in hospital transport robots: Evaluation of the role of user dispositions, anxiety, and robot characteristics," in Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI '22). Sapporo, Japan: IEEE Press, 2022, pp. 246–255.
[39] H. Abdollahi, M. H. Mahoor, R. Zandie, J. Siewierski, and S. H. Qualls, "Artificial emotional intelligence in socially assistive robots for older adults: A pilot study," IEEE Transactions on Affective Computing, vol. 14, no. 3, pp. 2020–2032, 2023.
[40] Z. Zhu, Y. Fu, W. Shen, A. Mihailidis, S. Liu, W. Zhou, and Z. Huang, "Cbash: A carebot-assisted smart home system architecture to support aging-in-place," IEEE Access, vol. 11, pp. 33 542–33 553, 2023.
[41] J. Edwards, C. Edwards, D. Westerman, and P. R. Spence, "Initial expectations, interactions, and beyond with social robots," Computers in Human Behavior, vol. 90, pp. 308–314, January 2019.
[42] K. Yu et al., "Design and research of a new type of massage hand based on chinese massage," Machine Design & Research, vol. 37, no. 3, pp. 22–28, 2021.
[43] Y.-H. Chang, Y.-T. Guo, L.-C. Fu, M.-J. Chiu, H.-M. Chiu, and H.-J. Lin, "Interactive healthcare robot using attention-based question-answer retrieval and medical entity extraction models," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 12, pp. 6039–6050, 2023.
[44] H. Shi et al., "End-to-end navigation strategy with deep reinforcement learning for mobile robots," IEEE Transactions on Industrial Informatics, vol. 16, no. 4, pp. 2393–2402, 2020.
[45] J. Karwowski and W. Szynkiewicz, "Quantitative metrics for benchmarking human-aware robot navigation," IEEE Access, vol. 11, pp. 79 941–79 953, 2023.
[46] K. Zhu and T. Zhang, "Deep reinforcement learning-based mobile robot navigation: A review," Tsinghua Science and Technology, vol. 26, no. 5, pp. 674–691, October 2021.
[47] K. Zhao and L. Ning, "Hybrid navigation method for multiple robots facing dynamic obstacles," Tsinghua Science and Technology, vol. 27, no. 6, pp. 894–901, December 2022.
[48] B. W. Mulvey, T. D. Lalitharatne, and T. Nanayakkara, "Deformobot: A bio-inspired deformable mobile robot for navigation among obstacles," IEEE Robotics and Automation Letters, vol. 8, no. 6, pp. 3828–3835, 2023.
[49] E. S. Shourmasti, R. Colomo-Palacios, H. Holone, and S. Demi, "User experience in social robots," Sensors, vol. 21, no. 15, p. 5052, July 2021.
[50] J. Holland, L. M. Kingston, C. McCarthy, and M. McConnell, "Service robots in the healthcare sector," Robotics, vol. 10, no. 1, p. 47, March 2021.
[51] E. Broadbent, R. Stafford, and B. MacDonald, "Acceptance of healthcare robots for the older population: Review and future directions," International Journal of Social Robotics, vol. 1, pp. 319–330, 2009.
[52] W. Wang, P. Zhang, Y. Shi et al., "A portable back massage robot based on traditional chinese medicine," Technology and Health Care, vol. 26, no. 4, 2023.

[53] M. Tsunoda and C. Premachandra, "Remote control of a wheeled robot by visible light for support in infectious disease hospitals," IEEE Access, vol. 9, pp. 124 165–124 175, 2021.
[54] F. Wu, T. Wu, D. C. Zarate, R. Morfuni, B. Kerley, J. Hinds, D. Taniar, M. Armstrong, and M. R. Yuce, "An autonomous hand hygiene tracking sensor system for prevention of hospital associated infections," IEEE Sensors Journal, vol. 21, no. 13, pp. 14 308–14 319, 2021.
[55] Y. Shen, D. Guo, F. Long, L. A. Mateos, H. Ding, Z. Xiu, R. B. Hellman, A. King, S. Chen, C. Zhang, and H. Tan, "Robots under covid-19 pandemic: A comprehensive survey," IEEE Access, vol. 9, pp. 1590–1615, 2021.
[56] Y.-L. Zhao, H.-P. Huang, T.-L. Chen, P.-C. Chiang, Y.-H. Chen, J.-H. Yeh, C.-H. Huang, J.-F. Lin, and W.-T. Weng, "A smart sterilization robot system with chlorine dioxide for spray disinfection," IEEE Sensors Journal, vol. 21, no. 19, pp. 22 047–22 057, 2021.
[57] A. A. Morgan, J. Abdi, M. A. Q. Syed, G. El Kohen, P. Barlow, and M. P. Vizcaychipi, "Robots in healthcare: a scoping review," Current Robotics Reports, vol. 3, no. 4, pp. 271–280, October 2022.
[58] M. A. Goodrich and A. C. Schultz, Human-Robot Interaction: A Survey. Now Foundations and Trends, 2008.
[59] K. Zhu and T. Zhang, "Deep reinforcement learning based mobile robot navigation: A review," Tsinghua Science and Technology, vol. 26, no. 5, pp. 674–691, 2021.
[60] W. Zhu, X. Guo, D. Owaki, K. Kutsuzawa, and M. Hayashibe, "A survey of sim-to-real transfer techniques applied to reinforcement learning for bioinspired robots," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 7, pp. 3444–3459, 2023.
[61] C. Yu and A. Rosendo, "Multi-modal legged locomotion framework with automated residual reinforcement learning," IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 10 312–10 319, 2022.
[62] M. Dobrevski and D. Skočaj, "Deep reinforcement learning for map-less goal-driven robot navigation," International Journal of Advanced Robotic Systems, vol. 18, no. 1, 2021.
[63] L. Zhang, Z. Hou, J. Wang, Z. Liu, and W. Li, "Robot navigation with reinforcement learned path generation and fine-tuned motion control," IEEE Robotics and Automation Letters, vol. 8, no. 8, pp. 4489–4496, 2023.
[64] M. Maroto-Gómez, R. González, Castro-González, M. Malfaz, and M. Salichs, "Speeding-up action learning in a social robot with dyna-q+: A bioinspired probabilistic model approach," IEEE Access, vol. 9, pp. 98 381–98 397, 2021.
[65] Z. Liu, Y. Zhai, J. Li, G. Wang, Y. Miao, and H. Wang, "Graph relational reinforcement learning for mobile robot navigation in large-scale crowded environments," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 8, pp. 8776–8787, 2023.
[66] M. Ranaweera and Q. H. Mahmoud, "Bridging the reality gap between virtual and physical environments through reinforcement learning," IEEE Access, vol. 11, pp. 19 914–19 927, 2023.
[67] E. Salvato, G. Fenu, E. Medvet, and F. A. Pellegrino, "Crossing the reality gap: A survey on sim-to-real transferability of robot controllers in reinforcement learning," IEEE Access, vol. 9, pp. 153 171–153 187, 2021.
[68] X. Liu, X. He, M. Wang, and H. Shen, "What influences patients' continuance intention to use ai-powered service robots at hospitals? the role of individual characteristics," Technology in Society, vol. 70, p. 101996, 2022.
[69] R. Bloss, "Mobile hospital robots cure numerous logistic needs," The Industrial Robot, vol. 38, pp. 567–571, 2011.
[70] V. Lu, "Service robots, customers and service employees: what can we learn from the academic literature and where are the gaps?" ournal of Service Theory and Practice, vol. 38, pp. 361–391, 2020.
[71] H. Yu et al., "Robots can't take my job: antecedents and outcomes of gen z employees' service robot risk awareness," International Journal of Contemporary Hospitality Management, vol. 34, no. 8, pp. 2971–2988, 2022.
[72] G. Rampersad, "Robot will take your job: Innovation for an era of artificial intelligence," Journal of Business Research, vol. 116, pp. 68–74, 2020.

AAILA ARIF graduated from the Honors School of Monmouth University in 2025. She is a master degree student in computer science now. Her research interests include in machine learning and data science.



JIACUN WANG joined Monmouth University in 2004 and is currently Professor of Computer Science. He served as the computer science and software engineering department chair during 2009 – 2015. He has been the graduate program director of the department since 2016. His research interests include machine learning, formal methods, discrete event systems, software engineering, and real-time distributed systems. He published four books and more than 300 papers. He is the founding EiC of International Journal of Artificial Intelligence and Green Manufacturing, as well as an Associated Editor of IEEE Transactions on Systems, Man, and Cybernetics: Systems, IEEE/CAA Journal of Automatica Sinica, IEEE SMC Magazine, and International Journal of Network Dynamics and Intelligence. He has served as general chair and program chair in several international conferences. He is recipient of 2025 Monmouth University "Distinguished Scholar" Award.