Reinforcement Learning for Optimizing a Delivery Path in a Hospital Setting

Aaila Arif

HO 396 Thesis/Capstone Proposal

Honors School

Monmouth University

11 December, 2023

Dr. Jiacun Wang

Abstract:

Reinforcement learning(RL) is a branch of machine learning that facilitates autonomous agents' interaction with their environments. This is done by “teaching” an agent efficient decision-making through iterative processes of exploration and trial-and-error. This project delves into an application of RL within the healthcare industry, where the objective is to construct a model capable of efficient navigation and task execution in a simulated hospital environment. The purpose of this is to optimize the pickup and delivery of essential supplies and medications to rooms within hospital settings. A virtual hospital floor will be created, housing an agent programmed to discover the quickest routes to designated destinations while avoiding obstacles and completing tasks. Python and OpenAI’s Gym Library will be utilized to create a 10x10 grid in GridWorld with customizable obstacles and destinations to reach. The program will employ Q-learning to create a state-transition system, simulating potential agent movements. With each movement, the agent will gain a numerical reward. Throughout training, it will explore, calculate, and store Q-values, which estimate both immediate and long-term action value. This learning process will guide the agent to choose actions with the highest cumulative reward, ensuring efficient task completion. Such an innovation has the potential to enhance the efficiency of hospital operations, contributing to improved patient well-being.

Introduction and Thesis:

The proposed area of research focuses on the application of reinforcement learning (RL) within the healthcare industry. The goal of this research centers around creating a model that enhances autonomous agent decision-making in a simulated hospital environment. In this context, the project addresses the optimization of pickup and delivery processes for essential supplies and medications within hospital settings. The core methodology involves the creation of a virtual hospital floor, featuring a programmable agent navigating through a 10x10 grid, avoiding obstacles, and reaching designated destinations. The implementation will utilize Python and OpenAI's Gym Library, employing Q-learning to establish a state-transition system that simulates potential agent movements. Through the exploration of the agent, it will calculate and store Q-values that estimate both immediate and long-term action values. This learning process will direct the agent to choose actions that yield the greatest cumulative reward, guaranteeing an effective completion of tasks.

The significance of this research lies in its potential to address critical challenges within the healthcare industry. Utilizing reinforcement learning makes it possible to enhance efficiency in hospital operations by developing an autonomous agent capable of optimizing pickup and delivery processes. By automating such processes, a positive impact on both patient and caregiver well-being can be achieved. The research is relevant in the context of applying machine learning techniques to solve real-world problems leading to abundant benefits for both healthcare providers and patients. Therefore, the significance of this research lies in its potential to bring about positive and practical changes in the healthcare sector.

The justification for completing this project comes 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 and Q-learning. The proposed research aims to bridge these gaps through the implementation of efficient machine learning techniques. Challenges faced in hospitals include shortage of staff, workplace hazards, and crowded facilities [6][7][8]. The project addresses these challenges by creating an agent capable of navigating through a hospital and performing tasks that are repetitive or dangerous to do in certain scenarios, such as medicine delivery. The project also addresses gaps in research such as the "reality gap" by introducing randomization of obstacle placement to simulate real dynamic hospital environments and create adaptability in the agent [20].

Review Of Literature

Section 1: Introduction

Subsection 1.1: Background

In order to contextualize the complexities behind the advanced methodology to be employed in this project, a comprehensive review of the literature underpinning 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 in order to improve both patient and healthcare provider well-being will be approached through reinforcement learning, the Q-learning algorithm, and the Python coding language. This review of literature will predominantly center around an explanation of the project, 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 are being addressed with the project.

Subsection 1.2: Brief Overview of Project

Reinforcement learning, a branch of machine learning, is an approach where an entity is trained to complete certain tasks [1]. This entity, called an agent, must interact with its environment and discover the most optimal route to accomplish its tasks [1][2]. The agent(controller) undergoes positive or negative reinforcement in response to an action, accessed through the reward and punishment functions [3]. 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 [4]. The policy, which maps states to actions, is pivotal [4]. Reinforcement learning aims to continuously identify and develop the most efficient policy to maximize cumulative future rewards [2][3][4]. One commonly adopted method for establishing a connection between the actions and rewards involves learning the anticipated quality of actions in a given state, referred to as Q-Learning [5]. The program employs this method to create a state-transition system to simulate potential agent movements [4][5]. Following each movement, the agent accrues a numerical reward, either a positive or negative number [4][5]. During the training phase of the agent, the Q-values are computed and stored as the agent systematically explores the environment, creating estimations of both immediate and long term action value [4]. This learning process directs the agent to select actions which will lead to the highest rewards, and thus, proficient task completion. Based on the research being done on reinforcement learning, its use in the current project will provide an efficient way to solve the problem of optimization of supply delivery in hospitals.

Subsection 1.3: Applications of Project

The research being conducted aims to apply methods such as reinforcement learning and Q-Learning simultaneously in order to enhance the efficiency of hospital operations such as collecting and distributing crucial supplies. This will be done by utilizing Python and OpenAI's Gym Library within the Pycharm IDE to create a 10x10 grid in GridWorld with tasks to complete, obstacles to avoid, and destinations to reach. This grid will simulate a virtual hospital floor, serving as the environment where the agent is tasked with collecting objects and reaching the endpoint using the shortest route.

Section 2: Robotics in Healthcare

Subsection 2.1: Evolving Role of Robotics

There have been substantial advancements in technology in the fields of computer science, machine learning, and robotics in recent decades. Robotics, in particular, has been one of the most rapidly evolving technologies ushering in a new era of possibilities across numerous industries such as healthcare, military, entertainment, and others [6][7]. In the healthcare sector, robotics finds applications extending to various roles including caregivers (nurses and physicians), hospital services, delivery, and beyond [6]. In the current project, the agent that is being trained with reinforcement learning in a virtual hospital floor will simulate a real-life service robot that can deliver necessities such as medicine within a hospital setting.

Subsection 2.2: Challenges in Healthcare

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 [6][7]. This manifests in heightened costs for hospitals, increased stress levels, and diminished quality of life [6][7]. The global challenges presented by the COVID-19 pandemic further underscored the critical need for technological solutions when issues such as an increase in the risk of contracting the illness, lack of healthcare centers and services, and lack of staff were revealed [6][7]. Healthcare workers are continuously exposed to occupational hazards such as bloodborne pathogens, infections, and other healthcare associated diseases [6][7]. The persistent issue of staff shortages, particularly in nursing, have exacerbated the problem [7]. 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 [7]. This scenario directly contributes to fatigue and burnout, which can pose potential repercussions on patient/caregiver health and the quality of care provided [7]. As advancements in technology and medicine continue to progress, an ever-aging population and escalating number of patients imposes an additional burden on hospitals [8]. The increasing patient influx drains hospital resources, leading to crowded facilities, stretched staffing levels, and competition for essential services and equipment [7][8]. As a consequence, patients must endure longer waiting times, delayed treatments, and an overall decrease in quality of care [7][8]. For healthcare professionals, an increasing patient load causes intensified work pressure and prolonged working hours. Constantly engaging in repetitive tasks, such as the distribution of medicines across expansive hospitals, causes a physical strain on staff [9]. As a result, a decline in the mental and physical well-being of the caregiving workforce is inevitable

[7]. Patients experience the negative ramifications of such 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 [7][8]. A common theme established in the reviewed literature demonstrates the necessity of integrating robots into hospital operations. In order for hospitals to address issues raised by aging populations, an 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.

Subsection 2.3: Benefits of Robotic Automation

Various benefits can be reaped by automating hospital processes with robots that take optimized routes. By replacing humans with robots when doing repetitive or dangerous tasks- those either undesirable or difficult due to physical limitations- it opens up many opportunities for improvement within hospital systems [9]. When it comes to crowded hallways filled with gurneys, stressed family members, and hospital staff, conducting simple tasks such as the delivery of essential resources becomes challenging [9]. With smaller robots substituting for humans in these scenarios, the timely and accurate distribution of such resources is possible [9]. Currently, industrial and commercial robots have extensive applications with reduced long-term costs and enhanced reliability in fields such as transport, surgery, assembly, and more [9]. In the context of smart hospitals, caretaker robots can lessen the workload of nurses and other caregivers by completing simple tasks [7]. This is especially important in today’s landscape where the decrease in staffing of hospitals continues to be an urgent issue [7]. By taking over menial tasks, doctors and nurses can focus on providing quality care for patients without being under mental and physical strain [7]. Moreover, these robots can improve patient experience and provide quick and reliable delivery of resources, contributing to enhanced patient health and outcomes [7]. Within hospital systems, caregivers are constantly exposed to a plethora of illnesses and diseases [6][7]. The integration of robots, which can reduce direct contact to contagious patients and exposure to harmful substances, serves as a protective measure for caregivers [6][7]. The containment of illnesses can be controlled through utilizing robots in tasks involving patient interactions [6][7]. By leveraging robots in these situations, a more hygienic and safer environment for both patients and caregivers can be fostered [6][7]. Examining potential benefits from the perspective of patient emotions, robots taking the place of humans can alleviate feelings of embarrassment, anxiety, and stress [8]. By eliminating human social presence, receiving service from a robot can decrease the likelihood of patients experiencing feelings of social judgment [8]. If a patient is being delivered an embarrassing product or having issues with hospital staff, robots can make patients feel more relaxed and less judged [8]. Furthermore, patients may feel less hesitant to seek and take medications when the fear of being judged is diminished [8]. This positively impacts health outcomes and leads to a more positive patient experience and perception of healthcare [8]. Both the emotional and physical needs of caregivers and patients are addressed by the addition of robots in hospitals.

Section 3: Program Implementation

Subsection 3.1: Project Explanation

The project will be implemented using the Python coding language in the Pycharm IDE. With the integration of OpenAI's Gym environment, a programming interface can be created for agents trained with reinforcement learning to interact with and navigate through a customizable environment [10]. Through the usage of the GridWorld environment, a 10x10 grid will be assembled with obstacles placed in user-chosen locations. The obstacles simulate objects that one may see in real life hospitals, such as machines, people, and other impediments. There are also two objectives that the agent must reach before going to the endpoint, and they represent resources/supplies that the robot must pick up to deliver to patients. The obstacles will be represented as black boxes, the objectives will be green boxes, the agent will be a blue circle, and the end point is a red box. The inputs from the agent at each step in time provides the agent with a positive or negative reward, and the goal of the agent is to take the route that maximizes the reward while collecting the items and avoiding obstacles [3][4][10]. By calculating and storing these rewards as Q-values, the agent approaches the optimization problem through a trial-and-error approach to find the most optimized route with the greatest rewards [4][10].

Subsection 3.2: Reinforcement Learning Approach

One fundamental aspect of machine learning is the ability to make decisions sequentially [11]. This refers to utilizing past experiences to determine what sequence of actions to perform in an unknown environment to accomplish a task [11]. Such decision-making tasks can be applied to various domains including healthcare, robotics, finance, and more [11]. Reinforcement learning takes inspiration from behavioral psychology, where a living organism can engage with its surroundings and learn about its environment [11]. Similarly, artificial agents aim to learn by interacting with the environment in order to optimize certain tasks [11]. The reward function defines the objective in reinforcement learning [11][12]. Each state and action is paired and corresponds to a number, called the reward [12]. This number can either be positive or negative, and through exploration and training of the agent, the agent aims to maximize this reward [11][12]. The policy maps states in an environment to the actions that can be conducted there [4][12]. It serves as a crucial mathematical function that is acquired by the agent since it contains all the necessary information for the agent to regulate its behavior once it is done training [12]. The agent aims to keep improving the policy in order to achieve the most optimized route [4][11][12]. To do this, the agent estimates values from the environment, and this reveals what actions are the most optimal to take in the long run [12]. In the ongoing project, reinforcement learning will be employed to enhance the efficiency of resource delivery routes within hospital networks. The approach taken aims to optimize the delivery process of supplies such as medicine, contributing to improved effectiveness of hospital operations.

Subsection 3.3: Deep Reinforcement Learning

Deep reinforcement learning is a subset of machine learning and reinforcement learning that combines principles of reinforcement learning with deep learning techniques [13]. While reinforcement learning algorithms learn to make decisions through interactions with the environment and gaining rewards and punishments based on actions taken, deep reinforcement learning builds upon this by utilizing deep neural networks to represent complex functions [4][13][14]. These deep neural networks convert states into action values [13]. Reinforcement learning then takes these action values and executes an action that relates to it [13]. There are risks associated with the deployment of robots in real world systems, such as damage to the robot through risky actions [14]. 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 [14]. For the scope of the current project, reinforcement learning will be applied in order to optimize essential supply delivery routes in hospitals. However, deep reinforcement learning can later be applied to enhance the efficiency of the project in complex real life systems.

Subsection 3.3: Q-Learning Approach

Q-Learning stands out as a prevalent strategy in reinforcement learning [15]. Known for its versatility, it has been applied in various artificial intelligence systems to address specific issues [15]. However, there still remains a notable gap in understanding how these algorithms can be integrated into artificial intelligence workflows [15]. The Q-Learning algorithm operates without a predefined model by relying on values [15][16]. The “Q” in Q-Learning means quality, and the function Q*(s,a) signifies anticipated cumulative reward [16]. The input “s” stands for state, and input “a” stands for action. Utilizing the optimal policy and two inputs, the Bellman Equation calculates the expected Q-value [15][16]. Quality values represent the reward gained by an agent when taking an action in a state and following a policy. The Markov Decision Process is a mathematical model used in Q-Learning to formalize sequential decision-making [15][16][17]. S represents the set of states in the environment, A is the set of possible actions in a state, P is the transition probability function, and R is the reward function [15][17]. P is utilized to define the probability of moving from one state to another state when an action is performed, and the reward function gives a reward to the agent when an action is performed in a state [15]. Through exploration and gaining rewards, the agent updates the policy based on actions that yield the best results, and therefore the most optimized path [15][16][17]. The goal of Q-Learning is to balance testing new actions to discover what effects they have with choosing actions known to have high reward values to create the most efficient policy [15]. As the algorithm progresses, the agent tends to follow a more greedy policy, selecting actions known to have high rewards [15]. In this project, Q-learning will be utilized alongside reinforcement learning to train an agent to choose the optimal policy and, consequently, the most efficient route for delivering resources within a hospital setting.

Section 5: Comparative Analysis

Subsection 5.1 Current vs. Previous Implementations

The choice of utilizing reinforcement learning as the underlying methodology of this project becomes relevant when examining alternatives to this approach. When considering current delivery robots, such as TUG mobile robots, the advantages of using reinforcement learning become apparent [18]. 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 [18]. The applications of this delivery robot include delivering resources such as medications, food service, and loading and unloading carts of medical supplies [18]. 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 [18]. 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 [18]. With these robots, 80% of the delivery within that hospital system was able to be automated, saving the hospital considerable amounts of money [18]. In the Children’s Hospital, TUG robots served as a food-delivery service and the sight of robots caused the children joy [18]. This improved the patient experience as well as cutting delivery costs for the hospital [18]. However, if the TUG systems for these hospitals were swapped, the robots would not easily be able to adapt to the changes in environment [18]. With reinforcement learning, the trial-and-error exploration approach can adapt to changing environments and unforeseen circumstances, whereas the TUG robots would need manual reprogramming to be effective [4][18]. With the control module approach of the TUG robots, a map and floor plan of the facility must be programmed when installing the system [18]. The adaptability of the robots would come into question in the case of changes in environment and floor plan, causing increased costs in manual re-programming [18]. Also, when considering optimization of routes and behavior over time, reinforcement learning triumphs over the TUG approach [18]. While the TUG robots are designed for autonomous decision making, they do not have the learning capability to optimize their routes and behavior over time [18]. With the reinforcement learning approach, the calculation and storage of Q-values while exploring the environment through trial-and-error allows the robot to improve its route over time [4][5].

Subsection 5.2: Reinforcement Learning vs. Alternative Models

There are three main branches of machine learning - reinforcement learning, supervised learning, and unsupervised learning [19]. Reinforcement learning maximizes the reward signal by mapping environment states to actions [19]. 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 [19]. Supervised learning trains a model on a labeled data set consisting of pairs of input and output data [19]. By allowing the model to map known inputs to actions, it aims to train the model to make predictions on unseen data [19]. Unsupervised learning aims to train a model to discover patterns and hidden structures in datasets that are unlabeled [19]. Since there are no explicit labels, the algorithm must create groupings with the input data [19]. For this project, 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 [4][19]. 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 [19]. Unsupervised learning models focus more on uncovering patterns within data rather than task-specific generalizations [19]. 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 [4]. Since the 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 [4][19]. 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 [19]. Similarly, unsupervised models are sensitive to changes in data [19]. 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 [4][5]. During the exploration and training phase, the approach of trial-and-error allows the agent to make decisions and discover the most optimized route [4][5]. The interactive learning and decision-making aspect of reinforcement learning is lost with both supervised and unsupervised learning [19]. Since reinforcement learning aims to constantly develop and improve the policy and choose the most efficient route, it is the best model to use for this project. The recurring theme throughout the literature reviewed in this section highlights that reinforcement learning would be the preferred and most optimal model for this project. Its ability to be versatile, adaptable to dynamic environments, and ability to continuously learn and improve demonstrates why it is the most effective approach.

Section 6: Addressing Research Gaps

Subsection 6.1: The "Reality Gap"

While previous research in the realm of reinforcement learning in robotics has been done, notable gaps still remain. This project seeks to address unexplored areas that continue to pose unresolved issues. One such issue is the "reality gap" that exists between the virtual and real world [20][21]. Due to disparities between complex real world systems and the virtual world, developing reinforcement agents proves to be a difficult task [20][21]. The performance of reinforcement learning agents trained in virtual settings diminishes as the real-world system that the robot will be interacting with becomes more intricate [20][21]. Since research regarding lessening the gap between the virtual real world is still being conducted, a common method being utilized is called "domain randomization" [20]. This method aims to introduce changes into the environment by randomizing or altering certain elements rather than training in a static environment [20]. 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" [20]. The current project aims to further the bridging of the reality gap by both utilizing randomization as well as simulating the hospital floor with utmost accuracy. During each run of the program, the user is prompted to enter x and y values for coordinates of obstacles that will be placed into the virtual environment. By allowing the user to pick where the obstacles are being placed with each run of the program, the agent is trained in a dynamic environment. An additional strategy that is to be incorporated into the program is virtual hospital walls to simulate the layout of real hallways in hospital systems. With this, the real world is being simulated as closely as possible in order to bridge the gap between both domains [20][21].

Subsection 6.2: Implementation Challenges in Hospitals

Another gap that is being addressed by this research is the failure of most hospitals to incorporate fully functional AI systems [22]. 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 [22]. Only a small group of hospitals that implemented these robots were able to keep utilizing them [22]. 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 [18][22]. Previously, hospitals would resort to robots that required a map of the environment as well as large amounts of manual coding to perform tasks such as delivery, food service, and more [18]. In the case of TUG robots, this strategy seemed to work [18]. However, when it comes to environments that are constantly changing or when moving the robots to a whole different environment, this method becomes less effective [18]. The current project tackles these challenges by implementing a mix of Q-Learning and reinforcement learning.

Section 7: Addressing Concerns

Robots are proficient in their abilities to take over simple tasks once handled by humans, such as assembly jobs and delivery services [23]. With recent advancements in technology, the capabilities of robots skyrocketed, and more complex physical and cognitive roles can be carried out [23]. For example, roles such as identifying signs of dementia in patients as well as detecting hazards in stores [23]. However, with the increased integration of AI and robotics into various industries, potential concerns have been posed by those working alongside them [23][24][25]. According to a report in 2018, 30%-65% of jobs face the potential of automation, putting them at risk of being replaced by robots [24]. For those whose jobs are included in this statistic, they may experience “service robot risk awareness” [24]. This means that job insecurity is instilled into those whose industries introduce and adopt service robots [24]. The workplace is undergoing rapid transformations, and a growing number of employees have expressed concerns with keeping up with robots that could replace them [25]. 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 [23]. 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 [23]. 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 [23][24]. However, this issue can be mitigated with collaborative robots that enhance the effectiveness and safety of employees [25]. Robots lack emotional intelligence as well as social skills, and this is where the opportunity for collaboration between robots and humans appears [24]. In jobs characterized by the necessity of a large social presence as well as complex emotional responses, the need for such collaboration becomes obvious [23][24]. While robots and AI handle more routine tasks, professionals can focus on more engaged and personalized care [23][24][25]. 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 [23][24][25]. 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 [23][24][25]. The current project aims 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.

Section 8: Conclusion

This literature review explores the conceptual foundation, methodology, and applications of reinforcement learning alongside Q-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 project that is detailed in the review will employ reinforcement learning and a Q-learning algorithm to train an agent within a virtual dynamic hospital setting. The choice of applying reinforcement learning to approach the route optimization problem in hospitals is justified through a comparison of other 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. A comparative analysis to previous implementations of service robots is also conducted, demonstrating the ability of reinforcement models to adapt to changing environments as well as learn when faced with unforeseen circumstances, which are common when it comes to hospital settings. The evolving role of robots is also mentioned in the review, as the many challenges faced by hospitals can be mitigated through the use of service robots trained with the methods that will be used in this project. Research gaps, such as the “reality gap” are addressed and solved through domain randomization and non-static programmable environments [20]. Additionally, ethical concerns such as job displacement and frustration are prioritized, and solutions to these issues are presented [23][24][25]. For example, collaborative efforts between robots and humans as well as new job positions based on managing the technology are a few solutions [23][24]. This literature review provides a strong foundation for the proposed project, presenting it as a significant enhancement to the domain of healthcare robotics.

Methods of Problem Solving

Subsection 1.1: Overview

This section outlines the methodology to be utilized for the implementation of the project. The project’s goal is to optimize hospital operations, such as the delivery of essential resources, through the application of reinforcement learning coupled with Q-learning in the context of healthcare robotics. A virtual environment will be created to simulate a hospital floor using the Python coding language, OpenAI’s Gym library, and Pycharm IDE. An agent will be trained in a dynamic and realistic environment to replicate real-life hospital systems.

Subsection 1.2: Data Collection

The selection of resources will be 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 will be utilized to select sources based on the topics of reinforcement learning, Q-learning, robotics and AI in healthcare, and gaps that exist in the research being done. In the second step, the papers will be read through entirely to discover common themes that can be established between multiple papers. In the last step, a final selection of papers will be conducted based upon sources that provide data pertaining to the project being done as well as containing lines of reasoning that can be linked to multiple other articles.

Subsection 1.3: Implementation Environment

The Python coding language along with the Pycharm IDE will be employed to program this project. The OpenAI Gym Library will be utilized to create a 10x10 grid in GridWorld to mimic a real-life hospital floor. Obstacles will be placed based on user input to simulate obstacles that may be seen in hospitals, such as people, manchines, etc. The objectives that the agent must pick up and deliver will be represented as green boxes, and the endpoint that indicates the delivery point will be represented as a red box.

Subsection 1.4: Agent Training

Reinforcement learning will be utilized to train the agent based on positive and negative reinforcement. The rewards gained by the agent will guide its actions, as it aims to gain the highest cumulative reward. Q-learning will map states to actions, creating the most optimal policy while the agent explores the virtual environment. This will help the agent regulate its behavior and choose the most efficient path for picking up and delivering the resources in the grid. The agent will be represented by a blue circle.

Subsection 1.5: Program Execution

When the program is executed, the user will be prompted to enter x and y coordinates for the placement of two obstacles. This creates a sense of randomness for the agent, which is crucial when it comes to replicating real-life dynamic environments. The obstacles will be represented as black boxes, and this inclusion aligns with the principles of reinforcement learning. With the randomization of obstacles, adaptability of the agent is created. The agent will then undergo a training/exploration phase and demonstrate the most efficient route to pick up the two objectives while avoiding the obstacles and reaching the endpoint.

References

1.S. Spanò et al., "An Efficient Hardware Implementation of Reinforcement Learning: The Q-Learning Algorithm," in IEEE Access, vol. 7, pp. 186340-186351, 2019, doi: 10.1109/ACCESS.2019.2961174.

2.L. M. D. Da Silva, M. F. Torquato and M. A. C. Fernandes, "Parallel Implementation of Reinforcement Learning Q-Learning Technique for FPGA," in IEEE Access, vol. 7, pp. 2782-2798, 2019, doi: 10.1109/ACCESS.2018.2885950.

3.Said G. Khan, Guido Herrmann, Frank L. Lewis, Tony Pipe, Chris Melhuish, Reinforcement learning and optimal adaptive control: An overview and implementation examples, Annual Reviews in Control, Volume 36, Issue 1, 2012, Pages 42-59, ISSN 1367-5788, https://doi.org/10.1016/j.arcontrol.2012.03.004.

4.Ohnishi S, Uchibe E, Yamaguchi Y, Nakanishi K, Yasui Y, Ishii S. Constrained Deep Q-Learning Gradually Approaching Ordinary Q-Learning. Front Neuro Robot. 2019 Dec 10;13:103. doi: 10.3389/fnbot.2019.00103.

5.Bonyadi, Mohammad R., Rui Wang, and Maryam Ziaei. "Self Punishment and Reward Backfill for Deep Q-Learning." *arXiv.org*Cornell University Library, arXiv.org, 2022. *ProQuest*,

https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/working-papers/self-punishment-reward-backfill-deep-q-learning/docview/2389081296/se-2, doi:https://doi.org/10.1109/TNNLS.2021.3140042.

6.Defi, Irma Ruslina et al. “Healthcare Workers' Point of View on Medical Robotics During COVID-19 Pandemic - A Scoping Review.” *International journal of general medicine* vol. 15 3767-3777. 6 Apr. 2022, doi:10.2147/IJGM.S355734

7.Lakshmi, Narayanan K., et al. "Fuzzy Guided Autonomous Nursing Robot through Wireless Beacon Network." *Multimedia Tools and Applications*, vol. 81, no. 3, 2022, pp. 3297-3325. *ProQuest*, https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/scholarly-journals/fuzzy-guided-autonomous-nursing-robot-through/docview/2631383048/se-2, doi:https://doi.org/10.1007/s11042-021-11264-6.

8.Holthöfer, J., van Doorn, J. Robots do not judge: service robots can alleviate embarrassment in service encounters. *J. of the Acad. Mark. Sci.* 51, 767–784 (2023). https://doi-org.ezproxy.monmouth.edu/10.1007/s11747-022-00862-x

9.Liu, Rongrong, et al. “Deep Reinforcement Learning for the Control of Robotic Manipulation: A Focussed Mini-Review.” *arXiv.Org*, 8 Feb. 2021, arxiv.org/abs/2102.04148.

10.Schuderer, Andreas, Stefano Bromuri, and van E. Marko. "Sim-Env: Decoupling OpenAI Gym Environments from Simulation Models." *arXiv.org*Cornell University Library, arXiv.org, 2021. *ProQuest*, https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/working-papers/sim-env-decoupling-openai-gym-environments/docview/2492485503/se-2, doi:https://doi.org/10.1007/978-3-030-85739-4_39.

11. Francois-Lavet, Vincent, et al. “An Introduction to Deep Reinforcement Learning.” *arXiv.Org*, 3 Dec. 2018, arxiv.org/abs/1811.12560.

12. Uc-Cetina, Victor, et al. “Survey on Reinforcement Learning for Language Processing.” *arXiv.Org*, 15 Mar. 2022, arxiv.org/abs/2104.05565.

13. Mohammed, Marwan Qaid et al. “Review of Learning-Based Robotic Manipulation in Cluttered Environments.” *Sensors (Basel, Switzerland)* vol. 22,20 7938. 18 Oct. 2022, doi:10.3390/s22207938

14. Yutaka Matsuo, Yann LeCun, Maneesh Sahani, Doina Precup, David Silver, Masashi Sugiyama, Eiji Uchibe, Jun Morimoto, Deep learning, reinforcement learning, and world models, Neural Networks, Volume 152, 2022, Pages 267-275, ISSN 0893-6080, https://doi.org/10.1016/j.neunet.2022.03.037.

15. B. Jang, M. Kim, G. Harerimana and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications," in *IEEE Access*, vol. 7, pp. 133653-133667, 2019, doi: 10.1109/ACCESS.2019.2941229.

16.Kavitha, M., et al. "Robot Navigation using Reinforcement Learning." *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 9, 2021, pp. 1862-1867. *ProQuest*, https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/scholarly-journals/robot-navigation-using-reinforcement-learning/docview/2623457334/se-2.

17.Yue Wang, Yuting Liu, Wei Chen, Zhi-Ming Ma, Tie-Yan Liu, Target transfer Q-learning and its convergence analysis, Neurocomputing, Volume 392, 2020, Pages 11-22, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2020.02.117.

18.Bloss, Richard. "Mobile hospital robots cure numerous logistic needs." *The Industrial Robot*, vol. 38, no. 6, 2011, pp. 567-571. *ProQuest*, https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/scholarly-journals/mobile-hospital-robots-cure-numerous-logistic/docview/900913685/se-2, doi:https://doi.org/10.1108/01439911111179075.

19.Juan Manuel Davila Delgado, Lukumon Oyedele, Robotics in construction: A critical review of the reinforcement learning and imitation learning paradigms, Advanced Engineering Informatics, Volume 54, 2022, 101787, ISSN 1474-0346, https://doi.org/10.1016/j.aei.2022.101787.

20. M. Ranaweera and Q. H. Mahmoud, "Bridging the Reality Gap Between Virtual and Physical Environments Through Reinforcement Learning," in *IEEE Access*, vol. 11, pp. 19914-19927, 2023, doi: 10.1109/ACCESS.2023.3249572.

21.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," in *IEEE Access*, vol. 9, pp. 153171-153187, 2021, doi: 10.1109/ACCESS.2021.3126658.

22.Xiaohui Liu, Xiaoyu He, Mengmeng Wang, Huizhang Shen, What influences patients' continuance intention to use AI-powered service robots at hospitals? The role of individual characteristics, Technology in Society, Volume 70, 2022, 101996, ISSN 0160-791X, https://doi.org/10.1016/j.techsoc.2022.101996.

23.Lu, Vinh N., et al. "Service robots, customers and service employees: what can we learn from the academic literature and where are the gaps?" *Journal of Service Theory and Practice*, vol. 30, no. 3, 2020, pp. 361-391. *ProQuest*, https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/scholarly-journals/service-robots-customers-employees-what-can-we/docview/2499028170/se-2, doi:https://doi.org/10.1108/JSTP-04-2019-0088.

24.Yu, Heyao, 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, 2022, pp. 2971-2988. *ProQuest*, https://ezproxy.monmouth.edu/login?url=https://www.proquest.com/scholarly-journals/robots-can-t-take-my-job-antecedents-outcomes-gen/docview/2694403000/se-2, doi:https://doi.org/10.1108/IJCHM-10-2021-1312.

25.Giselle Rampersad, Robot will take your job: Innovation for an era of artificial intelligence, Journal of Business Research, Volume 116, 2020, Pages 68-74, ISSN 0148-2963, https://doi.org/10.1016/j.jbusres.2020.05.019.