

## Review

# Dynamic Obstacle Avoidance and Path Planning through Reinforcement Learning

Khawla Almazrouei <sup>1,2,\*</sup> , Ibrahim Kamel <sup>1</sup> and Tamer Rabie <sup>1</sup><sup>1</sup> College of Computing and Informatics, University of Sharjah, Sharjah P.O. Box 27272, United Arab Emirates; kamel@sharjah.ac.ae (I.K.); tratie@sharjah.ac.ae (T.R.)<sup>2</sup> Autonomous Robotics Research Center, Technology Innovation Institute, Abu Dhabi P.O. Box 9639, United Arab Emirates

\* Correspondence: khawla.almazrouei@tii.ae

**Abstract:** The use of reinforcement learning (RL) for dynamic obstacle avoidance (DOA) algorithms and path planning (PP) has become increasingly popular in recent years. Despite the importance of RL in this growing technological era, few studies have systematically reviewed this research concept. Therefore, this study provides a comprehensive review of the literature on dynamic reinforcement learning-based path planning and obstacle avoidance. Furthermore, this research reviews publications from the last 5 years (2018–2022) to include 34 studies to evaluate the latest trends in autonomous mobile robot development with RL. In the end, this review shed light on dynamic obstacle avoidance in reinforcement learning. Likewise, the propagation model and performance evaluation metrics and approaches that have been employed in previous research were synthesized by this study. Ultimately, this article's major objective is to aid scholars in their understanding of the present and future applications of deep reinforcement learning for dynamic obstacle avoidance.

**Keywords:** machine learning; deep learning; mobile robot; navigation; collision avoidance



**Citation:** Almazrouei, K.; Kamel, I.; Rabie, T. Dynamic Obstacle Avoidance and Path Planning through Reinforcement Learning. *Appl. Sci.* **2023**, *13*, 8174. <https://doi.org/10.3390/app13148174>

Academic Editors: Renato Vidoni, Andrea Giusti and Lorenzo Scalera

Received: 15 June 2023

Revised: 7 July 2023

Accepted: 10 July 2023

Published: 13 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Despite the development of sensing and computer technology, autonomous driving systems, like mobile robots and autonomous vehicles, have received considerable academic attention and have seen an increase in commercialization. In logistics services, mobile robots are used as logistics robots, which are in charge of moving goods as well as in several other industries, including food delivery, restaurant and cafe service, informational guidance, and airport cleaning [1]. Modern intralogistics systems increasingly use autonomous mobile robots (AMRs) rather [2] than autonomously guided vehicles (AGVs). AMRs, unlike AGVs, may move in open spaces without needing extra floor markers such as painted lines/magnetic tapes, which are most frequently utilized when more flexibility is needed in jobs involving cargo movement. Even though they are free to move around, to improve the safety and efficiency of their operation, it is frequently necessary to restrict the areas in which they are allowed to move [3]. Using DRL for mobile robot indoor path planning has become a prominent trend in recent years, and it has done so with great success. However, given the context of restricted computational resources, the extensive application of DRL-based path planning to mobile robots has been severely constrained by the excessively lengthy training period. Map-free advantages, strong learning capacity, less reliance on sensor accuracy, etc. are all provided by DRL. In order to train DRL-capable robots, a 3D simulation environment with a physics engine is often needed. Physical laws in this case constrain the robot's ability to move. Based on the interactive trial-and-error process used in its execution, training takes a lot longer. As a result, the method struggles to converge to the intended targets in complicated contexts or activities with randomly initialized network parameters. For instance, the weights and biases may worsen the DRL-enabled path planning's performance [3]. Furthermore, most DRL frameworks constantly

execute actions whenever impacts from observation states are deduced. Because those actions are frequently localized, this may restrict and limit the potential of a mobile robot to make an appropriate global judgment.

### *1.1. Contributions of the Manuscript*

This study responds to our responsibility as scholars to develop a solid theoretical foundation for this emerging discipline. In contrast to the previous reviews, this review is a systematic literature review (SLR) of high-quality literature published between 2018 and 2022 from pre-defined resources and based on pre-defined inclusion/exclusion criteria. It focuses on dynamic obstacle avoidance and path planning through reinforcement learning. Studies were classified into different categories based on how they addressed dynamic obstacle avoidance and path design. Out of the 546 papers that were found utilizing the search approach, 34 met the requirements and were finally included. Different RL-based path planning and dynamic obstacle avoidance algorithms were evaluated based on their performance, the environment's complexity, and the type of RL algorithm they employed. Based on how different studies addressed dynamic obstacle avoidance and path planning, several studies were developed.

The process of the research in this area is well discussed. The obstacles in this field are then discussed, and the effectiveness of the method in terms of safety, effectiveness, and scalability is compared with the type of RL algorithm employed, the propagation model used, the complexity of the environment, and those factors.

Finally, the research gaps are thoroughly examined, and recommendations for potential future study topics are provided. By giving a complete picture and highlighting the areas that require development in dynamic obstacle avoidance and reinforcement learning-based path planning, these contributions can progress this field of study. The information on dynamic obstacle avoidance and path planning needed to start this field of study's evolution is all included in this paper.

### *1.2. Whats Is known*

#### *1.2.1. Application of Reinforcement Learning Algorithm in Mobile Robot Application*

In recent years, RL has gained significant attention and is widely used in various domains, including robotics. RL algorithms are used to design autonomous robots that can learn and adapt to new environments and tasks. Mobile robotics refers to the design and development of robots that can move autonomously in the physical world. There are several applications for mobile robots, which are logistics and healthcare, as well as agriculture, to perform navigation, object recognition, and manipulation tasks. They are equipped with various sensors and actuators that enable them to perceive their surroundings and interact with the environment.

Mobile robots with RL algorithms have demonstrated impressive capabilities in navigating complex environments, avoiding obstacles, and performing tasks accurately.

While mobile robot applications have been adopted in most human activities, reinforcement learning algorithms have also been simultaneously utilized to solve several problems and the shortcomings/inefficiency of mobile robot applications. For example, in recent times, mobile robots have found it difficult to navigate autonomously in a dynamic environment since most traditional algorithms are only efficient in a static environment where obstacles are not moving. In such an environment, a mobile robot usually works based on mapped information on the existing obstacles. So, it becomes a major problem for the mobile robot to autonomously navigate in a dynamic environment as it is not equipped to learn to plan its path to avoid dynamic obstacles continuously. Therefore, several techniques have been employed to help mobile robots solve such problems, and reinforcement learning is one of those promising algorithms for mobile robot autonomous navigation.

RL algorithms have been applied to mobile robots, which enhance them to learn and evolve to new spaces and tasks. The RL algorithm learns by interacting with the environment, and the robot's behaviors are modified in response to feedback from its

surrounding environment. This allows the robot to draw learning from its slips to improve over time.

Learning is one of the RL algorithms that is most frequently used in mobile robotics. Mobile robots use the well-known reinforcement learning algorithm Q-learning [4] to discover the best strategies for performing a variety of tasks, including navigation, obstacle avoidance, and object tracking. The goal of Q-learning is to teach the robot the optimum course of action to adopt in a given situation [5]. This method can learn the best policy regardless of the behavior policy (off-policy) and is model-free (it learns the best policy without a model of the environment).

By choosing actions that would maximize the projected future rewards, the robot learns to navigate its environment. The most effective action is the one that, out of all the potential actions, has the highest Q value [6]. Q-learning is therefore a value-based approach. To determine the best course of action, the agent uses the greatest expected Q values during the learning process [7].

Q-learning [8] blends temporal difference (TD) learning with theories like the Bellman equations and the Markov decision process (MDP). Agents can learn how to behave best in regulated Markovian domains. One-step Q-learning, which is its form, is what Equation (1) defines.

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha(r + \gamma \max_{a'}(Q(s', a')))) \quad (1)$$

$Q(s, a)$  estimates the action value following the application of an action in state  $s$ , where  $r$  is the immediate reward received,  $\gamma$  is the discount factor, and  $\alpha$  is the learning rate [9]. The following are the primary components of Q-learning [5]:

1. Agent: the learner who is equipped with a collection of sensors and actuators to interact with the environment;
2. Environment: the world around an agent or everything that interacts with it;
3. Policy: a mapping between the set of seen states,  $S$ , and the set of actions,  $A$ ;
4. The reward function, which maps state–action pairs to a scalar value;
5. Q value: the total payment that an agent might anticipate accumulating over time, starting from that state.

Learning functions by incrementally increasing its assessments of the effectiveness of activities at specific stages [7]. Similar to Sutton's approach of temporal differences, learning involves an agent trying an action at a specific state and evaluating the results in terms of the immediate reward or penalty it receives as well as its estimation of the worth of the state from which it was taken. Then, it discovers which actions—as determined by long-term discounted reward—are generally the best by repeatedly attempting all states and actions [10]. A series of unique stages or incidents make up the agent's experience. The agent in the  $n$ th episode [11]

1. Notices its current state  $s_t$ ;
2. Chooses and takes action  $a_t$ ;
3. Monitors the subsequent state  $s_t + 1$ ;
4. Receives a payout right away  $r_t$ .

Then, in accordance with [9], the agent modifies its Q values using a learning rate  $\alpha$ . Please take note that  $Q_{m \times n}(s_t, a_t)$  is assumed to be represented by a look-up table in this explanation.

Learning and performance characteristics are the two core components of a learning agent [12]. What kind of performance element is utilized, which functional component is to be taught, how that functional component is represented, and the type of feedback available all influence the design of a learning element [5].

In mobile robotics, Q-learning can be applied in various ways depending on the task. For instance, in navigation tasks [6], the robot can learn to select the optimal path to reach a target location by evaluating the expected reward of each possible action (such as moving

forward or turning left/right) in each state (that is, the robot's current location as well as the orientation).

Similarly, in obstacle avoidance tasks [11], the robot may learn to avert obstacles by evaluating the expected reward of each possible action (such as moving forward, turning, or stopping) in each state (such as the robot's proximity to the obstacle and its orientation). To implement Q-learning in mobile robots, one needs to define a suitable state representation that captures relevant information about the robot's environment and a set of possible actions the robot can take [12]. The Q value is then updated based on the rewards obtained by the robot as it interacts with the environment, using the Bellman equation [12]. Another RL-based algorithm used in mobile robotics is deep reinforcement learning (DRL) [9]. DRL uses a deep neural network to approximate the Q values. DRL has been shown to be effective in complex tasks such as object recognition and manipulation [13]. Mobile robots are also trained to carry out tasks like object detection and grasping using DRL-based algorithms [14]. In recent years, deep reinforcement learning (DRL) has grown to be one of the areas of artificial intelligence that has attracted the most attention [15]. Through high-dimensional perceptual input learning, it directly regulates agents' behavior by combining the perception of deep learning (DL) and the capacity for decision-making of reinforcement learning (RL). It offers a fresh concept for resolving robot navigation issues.

One of the main advantages of RL-based algorithms in mobile robotics is their ability to adapt to new environments and tasks [16]. The robot learns by interacting with the environment, and its behavior is adjusted depending on the feedback the environment has provided. As a result, the robot may learn new situations and activities without the need for manual programming. Mobile robots' RL-based algorithms, however, encounter several challenges. The high-dimensional state space is one of the difficulties [17].

Mobile robots have a variety of sensors that let them sense their surroundings. The high-dimensional state space, however, makes it challenging to store and update the Q values. Deep neural networks are used by DRL-based algorithms to approximate the Q values in order to get around this problem [18]. Another challenge is the exploration–exploitation trade-off [19]. The robot must explore the environment to learn about the rewards of different actions. However, exploration can be costly in terms of time and energy. Therefore, the robot must balance exploration and exploitation to achieve optimal performance [20].

### 1.2.2. Description of Terms

- Mobile Robot

Mobile Robot is a unique robot designed to move around and perform tasks in different environments without being physically fixed to one spot [4]. Instead, these robots have a mobility mechanism that enables them to move around on their own or with human control, such as wheels, tracks, or legs. Applications for mobile robots include exploration, transportation, manufacturing, and healthcare [2]. In addition, they can be programmed to carry out activities like object manipulation, delivery, surveillance, and inspection [2].

Khatib [4] discussed that mobile robots can sense and navigate their environments using various sensors, including cameras, lidars, sonars, and microphones. In addition, they use software and control algorithms to decide, plot paths, avoid obstacles, and communicate with people or other robots [9]. Furthermore, most mobile robots can learn from their surroundings and enhance their performance over time owing to artificial intelligence and machine learning [4]. These cutting-edge features make mobile robots incredibly versatile and adaptive to different tasks and environments.

- Reinforcement Learning

Reinforcement learning (RL) [2] reveals how natural and artificial systems can learn to predict the results of their actions, optimize those actions in environments where they might result in rewards or punishments, and move them from one state or situation to another.

Akanksha et al. [12] described RL as a subset of machine learning that focuses on teaching algorithms how to maximize a reward signal to make decisions in dynamic situations. It uses a procedure in which an agent interacts with its surroundings and is given feedback through rewards or punishments for its activities [12]. As a result, the agent learns optimal decision-making through trial and error by exploring the environment and taking actions that lead to the highest possible reward [12].

RL is especially well suited for use on mobile robots. This is because RL enables mobile robots to learn from their environment and adapt to changing conditions, which is necessary for autonomous navigation and decision-making; this is confirmed by [6]. RL algorithms can also enhance the operational efficiency of mobile robots by determining the best course of action in real-time [18]. Another author [9] discussed that RL could be used for various mobile robot applications, such as autonomous navigation, object recognition and manipulation, and task execution.

Complete independence from human labeling is the key advantage of reinforcement learning [2]. Furthermore, reinforcement learning is a promising strategy for creating autonomous systems that can adapt to changing environments and learn from their own experiences because the estimation of the action-value function is self-driven by the trial-and-error interaction with the environment and takes the robot states as the input of the model [6]. However, the training process can be time-consuming and requires careful tuning of hyperparameters to achieve optimal performance.

- Autonomous Navigation

Mobile robots have to go through a complex environment, dodge hazards, and eventually choose the optimal path. DRL has recently made some advancements in the labyrinth and indoor navigation. The A3C algorithm's performance in maze navigation was proposed to be enhanced by employing unsupervised auxiliary tasks by Jaderberg M. The proposed algorithm increased the convergence speed, resilience, and success rate. Li presented a DQN [21] and visual serving-based path-planning technique for mobile robots. To accomplish indoor autonomous navigation, the initial environment and the target image were collected as inputs, and matching relations and control techniques were formed through training. With the use of navigation technology, the development of navigation technology is becoming increasingly important for the development of robot path-planning technology.

- Propagation Models

Regarding mobility, topological variations, propagation, and energy limitations, mobility and propagation models in RL are exceptional [22]. Therefore, the mobility and propagation models in RL, which are also a component of the research topics in this study, must be clarified in order for this discussion to be completed. There are two types of reinforcement learning methods:

1. Positive

Positive reinforcement is defined as an occurrence as a result of a particular action. It has a favorable effect on the action taken by the agent and strengthens and repeats the behavior.

You can improve performance and maintain change for a long time with the assistance of this reinforcement. However, too much reinforcement could over-optimize the state, which might have an impact on the outcomes.

2. Negative

Negative reinforcement is the strengthening of behavior brought on by an adverse situation that ought to have been prevented or stopped. You can use it to specify the performance standard minimum. The disadvantage of this approach is that it just offers enough to satisfy the minimum behavior.

## 2. Materials and Methods

A systematic literature review model was considered to achieve an integrated and well-synthesized summary of previously published literature on this research phenomenon since it used secondary data/information synthesized from existing research in this domain. Similarly, its suitability was leveraged because this research method provided a comprehensive summary of the current findings related to the designed questions [23]. Instead, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocols were used, along with qualitative information from secondary data, as per the recommendations made by [24]. This helped in the collection and synthesis of data from high-quality scientific studies pertinent to the topic. This idea was consistent with the suggested study strategy from [25].

This review's methodology was staged; the protocol formulation came first, then the inclusion and exclusion criteria, and finally a search of an electronic database for pertinent papers. Following data extraction and synthesis analysis came the screening phase, which included the abstract, title, and full text.

### 2.1. Protocol Development (PD)

#### Research Questions (RQs)

This was the first stage of the research. It defined the main research topic, which served as the basis for the article search, paper selection, data sources, search terms, inclusion and exclusion criteria, and findings section [26]. In this process, the review questions were developed. The main questions that this study attempted to answer were as follows:

RQ1: What are the technical background and applications of reinforcement learning?

RQ2: What are the principles of dynamic obstacle avoidance/path planning?

RQ3: What is the propagation model?

RQ4: What are the performance metric and evaluation methods of reinforcement learning in obstacle avoidance algorithms in a dynamic environment?

RQ5: What is the gap in the literature?

The questions were motivated by the following:

RQ1: The scientific involvement in RL;

RQ2: The principles of dynamic obstacle avoidance/path planning;

RQ3: The various propagation models used;

RQ4: Performance metrics and evaluation model;

RQ5: Research gap.

### 2.2. Search Strategy

Considering the domain and focus of this present study, only database queries of the WOS, IEEE Xplore, and Google Scholar were considered for publications on dynamic obstacle avoidance and AMR through reinforcement learning for the past four years (2018–2022). The query of databases was based on formulated keywords, which were developed to focus on studies that discussed dynamic obstacle avoidance, autonomous mobile robots, their contributions to society, and their effects. Various words were combined using the Boolean operator ("AND" and "OR") based on the above scenario.

#### 2.2.1. Search Terms or Keywords

To identify all scientific articles related to path planning and dynamic obstacle avoidance in RL, the major search terms identified for this purpose were as follows: "Autonomous Mobile Robot" and its synonyms "AMR", "Path planning," and "Dynamic obstacle avoidance".

The three databases were searched using the following keywords: ("reinforcement learning") AND ("mobile robot" OR "autonomous robots") AND ("dynamic obstacle avoidance" OR "path planning").



### 2.2.2. Resources

As mentioned, this study explored citations indexed in three databases, including WoS, IEEE Xplore, and Google Scholar, due to their significance to the research phenomenon. In addition, bibliographic data retrieved from the database search were analyzed for literature completion.

A search for the aforementioned query was conducted in the PubMed databases, and a filter was used to only provide results from 2018 to 2022. The four databases were queried to select only relevant and high-quality studies. The database search and the link are presented below in Table 1.

**Table 1.** Searched Resources.

Sources	URL
Web of Science	<a href="https://webofscience.com">https://webofscience.com</a> (accessed on 7 December 2022)
Google Scholar	<a href="https://scholar.google.com/">https://scholar.google.com/</a> (accessed on 7 December 2022)
IEEE Xplore	<a href="https://ieeexplore.ieee.org/">https://ieeexplore.ieee.org/</a> (accessed on 7 December 2022)
Science Direct	<a href="https://sciencedirect.com/">https://sciencedirect.com/</a> (accessed on 7 December 2022)

### 2.2.3. Eligibility Criteria

The criteria for the studies' eligibility before inclusion into the synthesis analysis are presented here. Retrieved articles that met all the inclusion criteria were included in this study, while the exclusion of studies from the synthesis was based on the developed exclusion criteria. The inclusion and exclusion criteria for article selection are presented below.

#### Exclusion Criteria (EC):

EC1: All articles published in languages other than English were not considered.

EC2: Studies not focusing on path planning or dynamic obstacle avoidance in reinforcement learning were excluded.

EC3: Papers that did not discuss reinforcement learning were excluded.

EC4: Studies that did not provide an answer to one or more of the review questions were removed.

#### Inclusion Criteria (IC):

IC1: All publications from 2018 to 2022 were considered in the study.

IC2: Only articles published in English and that met other inclusion criteria were considered.

IC3: Studies focusing on path planning or dynamic obstacle avoidance in reinforcement learning were included.

IC4: Articles focusing on the review question(s) were selected.

### 2.2.4. Database Search Outcome

The results retrieved from the four databases are presented in Table 2. However, the database search result clearly shows that most publications on this subject matter were between 2020 and 2022. This research concept began to amass popularity among researchers in recent years.

**Table 2.** Database and Search Results (2018–2022) Before Inclusion/Exclusion Criteria.

Sources	Before
Web of Science	261
Google Scholar	87
IEEE Xplore	100
Science Direct	98
<b>Total</b>	<b>546</b>

### 2.3. Quality Assessment

To further enhance the quality of this review, all eligible papers were subjected to a quality assessment using the Jonnas Briggs checklist for analytical cross-sectional studies. In addition, the study appraisal was carried out to ensure the selection of only the most relevant papers for final synthesis in this survey. “Quality instruments” [27] used eight checklist criteria elaborated based on the author and are highlighted below. It is, however, very important to note that appraisals were based on “yes”, “no”, “unclear”, or “not applicable”.

#### *Quality Checklist criteria (QCs):*

QC1: Were inclusion criteria clearly defined (choice of a robot)?

QC2: Were the study subject (mobile robot) and the settings (environment) clearly defined?

QC3: Was the proposed model clearly defined?

QC4: Did the study use standard criteria and objectives to design the model?

QC5: Did it identify the confounding factors?

QC6: Were strategies for dealing with the confounding factors discussed?

QC7: Were the outcome measured validly and reliably?

QC8: Did the study use appropriate performance metrics?

The risk of bias was termed low if the study had an above 70% assessment score; below 70% but less than 50% were graded moderately low, while below 50% quality scores had a high risk of bias.

### 2.4. Data Extraction

After the criteria in Section 2.2.3 were processed, the findings were refined in order to extract the data from the studies selected that would address the research questions (RQs) posed in Section 2.1. When analyzing state-of-the-art articles, the following traits of the included studies were combined: identification information, publication year, country of the corresponding author, journal of publication, information corresponding to the six research questions, and main findings of the research articles were extracted using a preformed extraction form.

### 2.5. Data Synthesis

Depending on the type of information, the results’ information was extracted, tabulated, and/or visually plotted.

## 3. Results

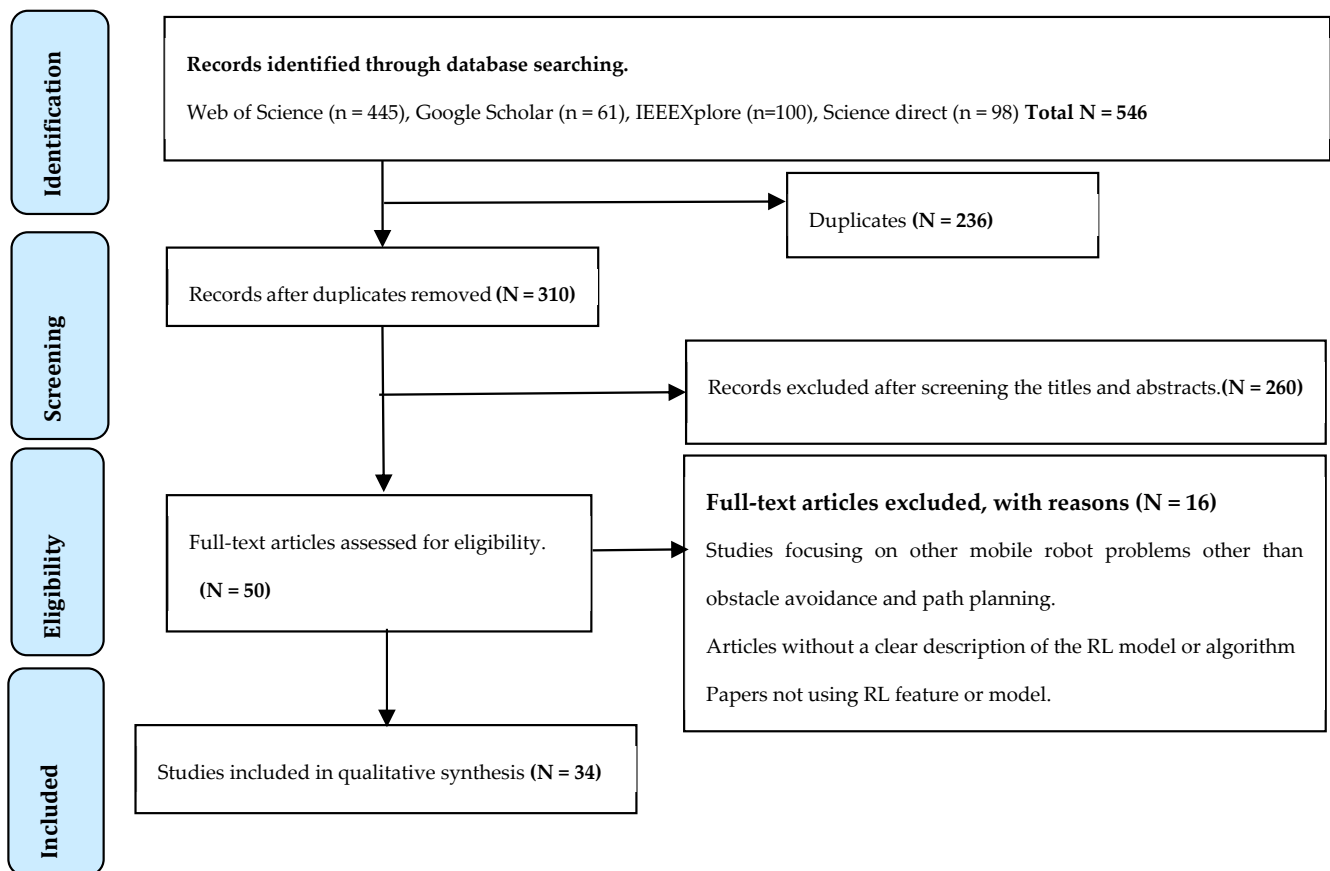
### 3.1. Data Selection

After conducting a keyword search across the three databases, 546 results were collected and imported into Excel 12.0. After the reference had been organized, cleaned up, and classified, 236 duplicates had been eliminated. Then, to find studies that fit with the research objectives, the titles and abstracts of the relevant studies were screened.

Studies that did not fall inside the scope of the study and at the same time did not concentrate on RL for mobile robot path planning and obstacle avoidance were excluded. A total of 260 papers were excluded from the subsequent screening step after the title and abstract screening.

The remaining 50 acceptable studies, however, underwent full-text screening to assess each study’s research topics, approach and methodology, data analysis, result presentation, and logical conclusion. Each manuscript underwent this examination to determine whether it fit within the parameters of the present study. A total of 16 articles were omitted from this procedure for various reasons. After the full-text screening, only 34 studies were qualified for the final inclusion in the synthesis analysis as shown in Figure 1.





**Figure 1.** PRISMA flowchart shows the screening process.

### 3.2. Quality Assessment

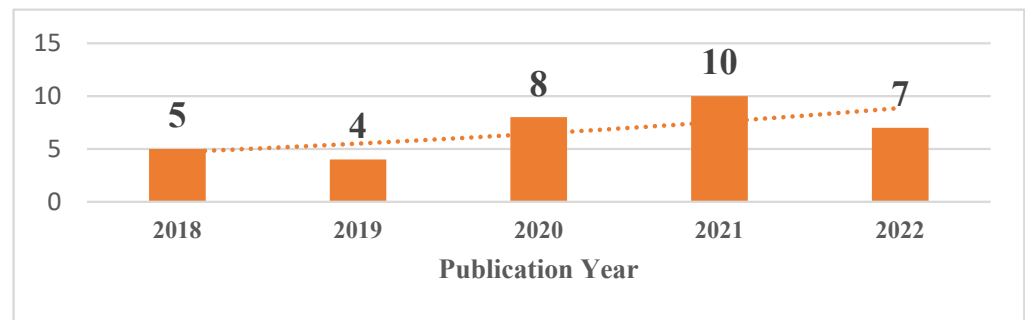
Using the JBI qualitative assessment tool, all 34 studies were appraised to assess the quality of the studies before the final inclusion into the systematic literature review.

Interestingly, the evidence from the quality assessment scores showed that all the included studies were quality studies with a low risk of bias. All the studies had 75% and above quality scores. Of the 34 appraised studies, only 2 had a 75% assessment score, while 52.7%, corresponding to more than half of the total included studies, had an 87.5% score instead of 100% because it was unclear whether a certain robot was used rather than presenting a solution to the mobile robot problem. The remaining studies, 15, had 100% quality scores.

### 3.3. Synthesized Result

34 papers were included in the final synthesis analysis after the quality of the chosen research was evaluated using the JBI checklist. The evidence in Figure 2 demonstrates that during the past five years, research on dynamic obstacle avoidance and path planning through reinforcement learning has increased with the highest number of publications in 2021. Understandably, most of the articles in 2022 may not have been fully published when retrieving the data for this research.

All the included studies applied several principles of reinforcement learning algorithms to solve the problem of mobile robot motion control, navigation, dynamic obstacle avoidance, and path planning in open or closed environments.



**Figure 2.** Publication trend.

#### 4. Discussion

The result of this present study is discussed under subheadings answering each of the formulated research questions.

##### 4.1. Application of Reinforcement Learning Algorithm in the Development of Mobile Robot

The findings of this review show that several theories and algorithms of reinforcement learning have been developed, and the findings of the reviewed experimental research demonstrate the usefulness of reinforcement learning models in mobile obstacle avoidance, path planning and optimization, navigation and learning, motion control, and adaption in both known and unknown environments. The basis of this algorithm is routed in training the robot to learn from the information available in the environment. Evidence from all the review studies shows that mobile robots can successfully transit from the start point to the target point without collision using RL algorithms. RL models have been used in different mobile functions. Their applications are further explained below.

##### 4.1.1. Path Planning, Optimization, and Obstacle Avoidance with RL

Automated mobile robots have widely been employed in the vast majority of working environments, especially in recent times where several work environments have recorded high costs of production. They have varying applications in multiple work environments, including household, agriculture, and industrial operations, for outdoor purposes while exploring resources, and wastewater transportation in the case of indoor surveillance. For the mobile robot to complete a set task, it moves from one point to another while avoiding obstacles and collisions. Path planning is one of the essential technologies of mobile robot functions in obstacle avoidance. Its importance corresponds to its ability to help the robot find the shortest possible route, free of obstacles and collisions. The mobile robot is expected to plan its path to avoid obstacles in outdoor or indoor environments. However, they often encounter problems of obstacles in the process of rotation and transition from the start point to the target endpoint in both known and unknown environments, although the problems are more complex in an unknown environment since there is little or no information available in such an environment. Hence, partial information is sensed using sensors in such an environment, whereas enough information is available for robots to learn about obstacles and plan their path in a known environment.

The reinforcement learning (RL) algorithm has been employed to solve mobile robot path planning, optimization, and obstacle avoidance problems. RL frameworks are employed to design a mobile robot, so it frequently interacts with any environment to enhance its adaptability, which, as a result, informs the robot's decision-making. For instance, Wang et al. [28] developed a double deep Q-network (DDQN) approach of reinforcement learning coupled with prioritized experience replay (PER). This was said to provide superior success in path planning and optimization. Their RL method enabled the robot to sense local information, thus planning its path in an unknown environment [29]; the experimental model consisted of a chassis, wheels, and about three ultrasonic sensors that equipped the robot with information to avoid obstacles in an unknown environment. Huang et al. [30] similarly

modified the deep Q-network (DQN) reinforcement learning theory in a reward-modified approach to achieve an obstacle-free path in a dynamic unknown environment. In their experiment, the DQN reward-modified algorithms enriched the robot to learn an optical decision in an unknown dynamic environment, modifying the abnormal reward function solved to avoid relative movement between the robot and the obstacles leading to abnormal rewards and collision. In this process, two modified reward thresholds were created, and the mobile robot successfully planned its path to avoid a collision. The authors further argued that the reward modified DQN reinforcement learning method offered the best result in mobile robot path planning and optimization of mobile robots in a dynamic unknown environment compared with other DNQ methods. Gao et al. [31] experimented on a known indoor environment. Deep reinforcement learning (DRL) was employed in their experiment to help a mobile robot plan its path in an indoor environment. To better augment the function and address the problems associated with the credibility of the DRL approach, a novel incremental training model was introduced in their experiment in both 2D and 3D environments. The standard global path planning algorithm Probabilistic Roadmap was combined with Twin Delayed Deep Deterministic policy gradients as the novel path planner in the DRL model. And the results of the experiment confirmed that this model of training mobile robots incrementally could enhance the robot's efficiency in planning its path in an indoor environment. Lee and Jeong [32] developed an RL algorithm based on path-planning techniques for usage in a warehouse. In order to minimize information processing delays, the study demonstrated how reinforcement learning algorithms could increase mobile robot path search accuracy and decrease path search time. The authors also highlighted the usage of Q-learning and Dyna-Q-learning reinforcement learning algorithms to create a single agent and optimize its single path. However, the author described how optimizing the warehouse path in an open field may be difficult as it exists in a complex environment, as the dynamic environment changes occasionally. The author also argued that adopting the single-agent and single-path algorithm in an outdoor environment such as a warehouse was insufficient. Several other RL methods have demonstrated success in helping mobile robots avoid obstacles. Xiang et al. [33] mentioned that continuous space path planning using reinforcement learning alone was enough to help mobile robots plan the path and avoid obstacles. Compared with multiple traditional algorithms, this method showed a high success rate. Similarly, Feng et al. [34] highlighted how several deep reinforcement learning algorithms, such as Deep Q-Network and Double Deep Q-Network combined with Prioritized Experience Replay, could successfully achieve a collision-free path. These frameworks sense information from the environment to autonomously help mobile robots plan their path to avoid obstacles while performing activities. This finding was similar to that of [28], which also found RL in obstacle avoidance using the same approach.

Moreover, several studies have described how complex it is for a mobile robot to plan its path in a dynamic environment compared with a static obstacle environment where information about the environment is fed into the mobile robot, just as the traditional algorithms use obstacle maps. Huang et al. [35] suggested that an improved Q-learning algorithm in a modeled environment, a reward function, and policy adaptation helped mobile robots plan their path, avoid obstacles, and hit the target goal via the fastest possible route in a dynamic environment where no fixed information about obstacles existed. The findings were congruent with those of Choi et al. [36], who also claimed mobile robots could be trained to learn information on collision-free paths and obstacle avoidance in an environment with several non-static dynamic obstacles. Most of this RL framework often integrates a path planner into the model to help mobile robots solve the difficulties of finding their path, which they may encounter in a difficult situation, especially in a dynamic obstacle environment. Reinforcement learning, however, is essential in understanding obstacle intent in a dynamic environment. Robots sometimes may plan their path without understanding the specific purpose of the obstacle, which may reduce mobile robot path planning efficiency. Likewise, interactions between obstacles are neglected, leading to avoidable trajectories. However, studies have demonstrated the importance of

reinforcement learning in understanding obstacle intent. Evidence from a recent study conducted by Xiaoxian et al. [37] affirmed through their experimental research that the mobile robot could learn collision-free policies through a deep reinforcement learning algorithm. Mobile robot obstacle interactions and obstacle-to-obstacle interactions can be properly understood using deep learning techniques of the reinforcement learning method. Meanwhile, previous scholars have suggested an angled pedestrian grid can help understand obstacle interaction in a dynamic environment. This is because it allows mobile robots to autonomously extract enough information about the dynamic environment, making them reach their goal without colliding with obstacles.

Further, non-permanent obstacle characteristics can be adapted via an attention mechanism, thus allowing joint obstacle impact on the obstacle avoidance strategies. While studies have tested the importance of the RL method in an outdoor environment, another recent study by Song et al. [38] considered a multimodal deep reinforcement learning technique in a complex indoor environment. Like other scholars, the study demonstrated DRL's significance in learning multiple control policies using sensors. However, their reliable and robust multimodal DRL model overcame the difficulties posed by the heterogeneity of sensor modalities and the complicated indoor environment. It was further argued that the mobile robot model based on DRL had an associated high-performance rate in terms of convergence speed, average accumulated rewards, and success rate.

#### 4.1.2. Navigation and Learning with RL

One of a mobile robot's fundamental features is autonomous navigation. It is without a doubt essential to robotic operations. It tends to navigate in an environment with a static or dynamic obstacle to hit its target through a thorough comprehension of its navigating environment in order to move from its set point to its goal.

However, the efficiency and accuracy of the robot depend on a good perception of its environment and learning to navigate within the shortest possible time. This seems to be one of the leading problems in mobile robot development, especially with a less efficient traditional algorithm in mobile robot navigation. This traditional algorithm works based on a previous obstacle map in a navigation environment. Despite the high accuracy and short path of the algorithm, establishing the obstacle map is lengthy and, in some conditions, may not be available. Also, the accuracy of the algorithm depends on the precision of the map range. However, the emergence of the reinforcement learning algorithm has helped solve several problems that limited mobile robot navigation. RL offers a proper understanding of known and unknown environments through learning easy robot navigation without encountering obstacles. Several experimental interventions adopting different reinforcement learning algorithms have been developed and tested for their possible success in navigation. As opposed to the traditional navigation algorithm, Taghavifar et al. [39] develop and test a path-planning for nonholonomic robots that takes into account a moving/stationary obstacle avoidance approach which is achieved by integrating a chaotic metaheuristic optimization technique with a single-time velocity estimator-based reinforcement learning RL algorithm. The study further contributes by incorporating the impacts of tire sinkage into the deformable terrain on the dynamics of the robot and applying the principles of Terr mechanics to determine the ideal compensatory force/torque magnitude to support a stable and fluid motion. Through continuous interaction with the navigating environment, reinforcement learning models enable mobile robots to learn and navigate safely in any environment compared with supervised learning.

Another reinforcement learning model was tested by Chew and Kumar [40]. The experiment found that the Q-learning algorithm of the RL method successfully helped the mobile robot avoid an obstacle in a navigation environment through alternative path planning different from the prior established optimal path used in most traditional algorithms. The study posited that using reinforcement learning techniques or methods made it easy for mobile robots to navigate dynamic obstacle environments. Several scholars also intervened in the problem of autonomous learning of conventional or traditional algorithms used for

mobile robot development using deep reinforcement learning (DRL). This was confirmed by Ruan et al.'s [41] experiment reviewed by this paper. Similar to [38], they used an end-to-end deep reinforcement learning algorithm to solve the problem of autonomous mobile robot navigation in an unknown environment. A combination of dueling network architecture for RL (Dueling DQN) plus DRL with double q learning (DDQN) to form a D3QN algorithm allowed the mobile robot to autonomously learn about the environment in a gradual process while transiting and learning to navigate its path to the target goal with an RGB-D camera in the environment. Like previous authors, the study also demonstrated that adopting the RL model in mobile robot development allowed the robot to reach its goal without colliding with obstacles. To address the problems that are instances of local stable points, an improved black-hole potential field and reinforcement learning can be employed [42]. In a reinforcement learning algorithm, the environment is the black-hole potential field. Agents automatically adjust to their surroundings and learn to use the most basic environmental data to locate targets.

Also, in creating an end-to-end model, a mobile robot navigated a mapless environment using a critical soft actor of the deep learning algorithm. In Ref. [40], compared with the traditional approach, the simulation could help mobile robots autonomously navigate to accomplish their task based on end-to-end autonomous navigation learning. Deep reinforcement learning techniques have proven their worth in terms of performance in mobile robot navigation. Equipping mobile robots with information about the environment beforehand is characteristic of traditional approaches, which become inefficient, especially in a dynamic environment and a scenario where information about the environment is unavailable. The introduction of DRL algorithms has solved this. The mobile robot now autonomously navigates in a dynamic obstacle environment without collision (Han et al. [43]).

#### 4.1.3. Motion Control and Adaptation with RL

Reinforcement learning has also been used to solve the motion control problem in mobile robot applications. Often, point stabilization control is a major shortcoming of traditional algorithms. This is because it usually encounters several constraints, including non-linear factors, non-holonomic conditions, and under-actuated constraints, among other constraints of mobile robots. However, the identification of the control action through end-to-end training can be achieved using DRL to solve these constraints of mobile robots. Two reviewed studies confirmed RL's importance in mobile robot motion control. For example, Gao et al. [44] utilized deep reinforcement learning to solve the motion control problem encountered by non-holonomic constrained mobile robots. To build the DRL, a point stabilization kinematic control law was used. The experiment's result confirmed the success of RL in motion control as it effectively identified the stabilization control point of the non-holonomic mobile robot. Furthermore, Duguleana and Mogan, [45] experimented with neural network-based reinforcement learning to overcome collision encounters by mobile robots in both dynamic and static environments. Network planning and Q-learning were interestingly used to solve the path-planning problem of the mobile robot. Safe and swift navigation was achieved in a scenario with the availability of global information. The speed of the robot can be set prior to the computation of the trajectory, which provides a great advantage in time-constrained applications. Zhang et al. [46] adopted trajectories from a practical life motion planner and previously executed trajectories as feedback observations. This reinforcement algorithm model helped the mobile robot with motion adaptation behavior concerning environmental changes.

#### 4.2. Principles of Dynamic Obstacle Avoidance and Path Planning

Several principles have been tested in developing a mobile robot capable of planning its paths and avoiding obstacles while performing the set task in a static, dynamic, or unknown environment. However, evidence from the review of 19 experiential studies on dynamic obstacle avoidance and path planning shows that all the studies rely on the principles



of reinforcement learning techniques in training mobile robots to move in an obstacle- and collision-free path. However, these principles of reinforcement learning techniques rally round training the robot to learn and extract information about the obstacles in the navigating environment [47]. In addition, the reinforcement learning model has employed deep learning and continuous learning principles to allow mobile robots to plan their path in a dynamic obstacle environment [48]. This allows a mobile robot to decide on motion control autonomously and how to plan its path and avoid obstacles without being pre-informed about the environment. Likewise, it is a safe time for a self-monitored approach.

To overcome path planning issues for mobile robots, [49] conducted an experiment that focused on creating a novel, robust, RL-based algorithm using Q-learning and the Artificial Potential Field (APF) approach. Q-learning has become more widely used recently to plan the path of mobile robots. It does, however, show slow convergence to the ideal answer. The APF method improves the conventional Q-learning approach in order to get around this restriction. In this experiment, suitable paths for mobile robots in both known and new locations were calculated using a QAPF learning algorithm. The length of the path, how smooth the path was, and the learning time were all utilized to determine how effective planning was. Path planning issues were successfully resolved in offline and online modes for both known and unknowable situations by the QAPF learning algorithm. The experiment's simulation findings demonstrated that the algorithm may accelerate learning and improve path planning in terms of path length and path smoothness. The algorithm successfully satisfied the three criteria for the path planning problem: smoothness, length, and safety. It was very dependable for mobile robot route planning since it generated better results across all test scenarios and had a stable training duration.

In order to enhance the performance of mobile robots, Quan et al. [50] presented an experimental navigation system based on DRL that integrated a recurrent neural network with DDQN. It practiced with physical robots, three-dimensional TurtleBot simulation environments, and two-dimensional simulation environments. Obstacles, prizes on the map, and discount variables were adjusted in accordance with the environment to increase the viability of the navigation system. The strategy operated well, as demonstrated by the experiments, making the outcomes of path planning more practical. Additionally, the experimental results demonstrated that the suggested algorithm increased the path length and path-finding efficiency.

As a framework to describe complex relationships and collaboration, Everett et al. [51] present deep reinforcement learning in avoiding pedestrian collisions, safe and smooth transitions. However, when the number of agents in the environment increase, they are built using fundamental presumptions about the behavior of other agents that are false. In order to efficiently avoid obstacles and arrive at a specific point, mobile robot agents were taught. To address the issue of not being able to discover a path in a particular circumstance, a path planner was combined with reinforcement learning-based obstacle avoidance.

For the differential drive wheel robot, the training policy was implemented in the robot operating system (ROS) and evaluated in both virtual and physical situations. In order to learn policies that would lessen the differences between the training and real-world driving settings, a soft actor-critic (SAC) reinforcement learning method was utilized. Driving to the goal point while maintaining forward and rotational velocity are all components of the reinforcement learning-based obstacle avoidance driving strategy. The goal of developing Mobile Robot Collision Avoidance Learning with Path (MCAL\_P) was to address the issue of driving in a particular environment.

It was integrated with a path planner, and the trained policy's input was changed with the help of lidar data, the robot's speed, and the separation from the look-ahead point. Through simulations and actual tests, it was demonstrated that the dynamic object avoidance performance was outstanding, and the obstacle avoidance performance was excellent.

We confirmed that robots using MCAL\_P can also avoid things that arise quickly at quick response rates in real-world environmental testing. We also discovered that because



robots employing MCAL\_P avoid one another according to the same rules, avoidance in multi-agent environments is more efficient and secure. However, several drawbacks also need to be addressed, such as the inability to drive straight, the inability to create efficient courses around static objects, and the lack of acceleration or deceleration when driving in addition to obstacle avoidance.

Deep Reinforcement learning's application technique was explored by Niroui, et al. [52] to aid the efficiency of rescue robots in urban search and rescue in an unknown clustered environment. The two most fundamental algorithms, Q-Learning and Dyna-Q, were put to the test, and the experimental outcomes were contrasted in terms of path search times and ultimate path lengths. Although the Q-Learning algorithm was found to be quicker, it struggled to locate search paths in a challenging setting with numerous barriers. With "dynamic reward adjustment based on action to improve the results," the MR was improved. Robots on wheels typically seek areas with high reward potential. The experiment confirmed that a straightforward method to increase path search accuracy and keep path search duration could reduce ineffective exploration.

The final path length was 71% less than the Q-learning algorithm, while the path search time was 5.5 times faster than the Dyna-Q methods. The warehouse environment in the field is significantly more complicated and constantly shifting, however. To tackle more complicated and realistic situations, a thorough analysis of the enhanced reinforcement learning algorithm is therefore required.

MDRLAT, a brand-new multimodal DRL technique, was put forth in [53] for the obstacle avoidance of mobile indoor robots. The 2D laser range discoveries and depth photos were combined using the bilinear fusion (BF) module, and the resulting multimodal representation was supplied into the D3QN to produce control commands for indoor mobile robots. The experimental outcomes demonstrated the method's exceptional performance in terms of the average accumulated reward, convergence speed, and success rate.

#### 4.3. Propagation Model of RL

Scholars have employed several propagation models of reinforcement learning. However, Q-network learning, continuous learning, 2D and 3D sensors, and soft actor-critics were the most adopted propagation models of reinforcement learning in obstacle avoidance and path planning algorithms [47,53,54]. Meanwhile, different improved propagation techniques of Q-network learning based on deep reinforcement learning have been the most prominent approaches adopted across different experimental studies. In general, acceptability and adoption may be traced to the ability to accurately extract information about the obstacle in a dynamic environment, which is used to create a navigation map for the robot to move in a collision-free path, as the real-world environment is heterogeneous, containing both static and dynamic environments. This propagation model, therefore, helps mobile robot algorithms consider unforeseen obstacles and dynamic objects. Ref. [54] uses the idea of partially directed Q-learning and the artificial potential field (APF) method to enhance the traditional Q-learning strategy. Because Q-learning and the APF approach are combined, the proposed QAPF learning algorithm [49] for path planning can speed up learning and improve performance overall.

A propagation model of recurrent neural network with DDQN was used in [55,56], aiming at navigation and path planning problems. In [57], the authors used a Mobile Robot Collision Avoidance Learning (MCAL) Soft Actor-Critic (SAC) based on RL. For RL as a path optimization technique, the propagation model of Q-learning and Dyna-Q algorithms was employed. A multimodal DRL with Auxiliary Task (MDRLAT) was utilized by [53] to help indoor mobile robots avoid obstacles. In [57,58], a Deep Deterministic Policy Gradient (DDPG) propagation was employed. Refs. [59,60] employed the Deep Neural Network (DNN) propagation technique to design a mobile robot's path autonomously.

#### *4.4. Performance Metric and Evaluation Method of RL Propagation Model in Autonomous Mobile Robot*

Different ways of evaluating the performance metrics of a reinforcement learning algorithm in a mobile robot path planning and obstacle avoidance were synthesized by this review. However, the Gazebo simulation remains the most explored simulation environment. It represents a virtual environment that permits the training of the mobile robot using the RL techniques. However, the dimension and the nature of the training environment vary across studies, some with static obstacles and others with dynamic or complex obstacles. A mapping gripper and turtle bot are often equipped in the simulated environment created with Gazebo, and the propagation model's efficiency is evaluated [61]. The turtle bot is the experimental agent in the simulated environment, whose success rate is evaluated against the target goal in the simulated simple/complex environment. Some authors prefer to use a test map [49] that evaluates the success rate of the proposed model in a 3D simulated virtual environment [50]. Different types of sensors are applied in each model in the test environment. Other notable means are 2D range sensors [53], kinetic sensors, and simple test scenarios. However, the performance of the RL algorithm, a reward function value, a punishment value, a state space, or a Q value are evaluated, and finally, the accumulated reward value is considered [54,59]. The reward function is essential in reinforcement algorithms, especially in training RL results [55]. The RL model is considered good if its reward value is evaluated as a positive reward value [56]. And for this to occur, the distance between the goal position and the mobile robot agent is less than a set value, making it effectively complete the target navigation [57]. However, when the minimum distance between the mobile robot and dynamic obstacles is less than a safe distance, like a point where collision is possible, the mobile robot agents oftentimes are punished by allocating a penalty value like a negative reward value [58]. To this end, an accumulated reward is calculated. In a negative reward function scenario, the training episode is ended. In this way, the RL model success rate is evaluated.

#### *4.5. Research Gap in Reinforcement Learning Application in Mobile Robot Obstacle Avoidance*

Environmental complexity remains the major issue in mobile robot obstacle avoidance. Most research on reinforcement learning algorithms sometimes fails in a dead zone, such as a complex experimental environment, perhaps due to some inappropriate reward condition. This, however, makes it difficult for the robot to search the solution space; consequently, learning along the path of the wall and escaping the dead zone becomes difficult. Creating appropriate reward functions in the dynamic environment remains the main gap in research. However, the review of the included studies also shows that while authors are trying to help mobile robots plan their path and avoid obstacles in dynamic and static environments using the RL algorithm, little is known about the interaction of multiple mobile robots within a single static or dynamic environment. Few or no studies have demonstrated how multiple mobile robots can co-exist within a single environment, carrying out different tasks without clashing with one another while still avoiding obstacles. There is also the issue of robot cooperation, and more issues will need to be resolved. Hence, there is a need for more studies to investigate multiple-robot cooperation using the RL algorithm in a different environment. Another important gap considered by this current study is that most review studies on this phenomenon of obstacle avoidance use virtual and simulated robots and test environments. In contrast, few studies investigated the efficacy of the RL algorithm in a real environment. The autonomous navigation of a robot in dynamic surroundings could still use a predictive navigation framework that can forecast the robot's long- and short-term trajectories. However, this could enhance the robot's capacity for learning.

### **5. Conclusions**

This systematic review has provided a comprehensive understanding of path planning and obstacle avoidance using different techniques of the reinforcement learning algorithm. The review of nineteen experimental studies in this sphere affirmed the significance of

reinforcement learning techniques in solving the challenges of mobile robot obstacle avoidance and path planning in a dynamic environment, which most traditional algorithms often find difficult to navigate in a complex obstacle environment. Through its interactions with the environment, the robot's policy is improved using the robot motion-planning approach based on DRL. Robots can develop strong automated learning and decision-making abilities by adopting this RL or DRL strategy. For unstructured contexts like partially mapped and dynamically changing surroundings, this capability is essential. Additionally, DRL techniques of RL reduce programming complexity and do away with the need for prior environmental knowledge. Likewise, RL is demonstrated to provide the robot with the ability to assess its surroundings and improve its ability to control movement, plan its path in a collision freeway, and avoid obstacles not only in a simulated environment but also in real-time to search objects and make decisions. Similar to Zhang et al.'s [62] comprehensive review and open issue on robot reinforcement learning, it is worthwhile to assert that reinforcement learning techniques promote the development of intelligent robots. Ref. [62] similarly confirmed the significance of reinforcement learning algorithms under uncertainties and how they can assume sequential decisions under such situations via end-to-end continuous and autonomous learning. In addition, the review and open issue similarly identified that a large amount of research investigations on reinforcement learning in robotic applications are still in a simulated environment and therefore may not perform well to solve and tackle real work problems, most especially in a dynamic real-world environment. Hence, Refs. [7,62] opined that having an efficient benchmark and standard environment for a real-world environment is important to achieve an applicable system; otherwise, the feasibility of reinforcement learning algorithm generalization testing and evaluation may become difficult. The argument remains that there is a large difference between simulated environment data and real environment data, even though a simulation environment can facilitate the process of robot learning and provide a reliable evaluation of physical behavior and control. However, the cost of obtaining a real-world experience is on the high side, and real-world experience is often difficult to recreate, which explains the reason why numerous studies on mobile robots based on RL are still in the simulation stage. In spite of these, this study, in line with other comprehensive reviews, presents the state of the art in mobile robot research based on reinforcement learning. As a result, navigation and path planning for mobile robot obstacle avoidance are well underpinned by this review.

**Author Contributions:** Conceptualization, K.A. and I.K.; methodology, K.A., T.R. and I.K.; software, K.A.; validation, T.R. and I.K.; formal analysis, K.A.; investigation, K.A.; resources, K.A.; data curation, K.A.; writing—original draft preparation, K.A.; writing—review and editing, T.R., K.A. and I.K.; visualization, K.A.; supervision, T.R. and I.K.; project administration, T.R. and I.K.; funding acquisition, K.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hadi, B.; Khosravi, A.; Sarhadi, P. Deep reinforcement learning for adaptive path planning and control of an autonomous underwater vehicle. *Appl. Ocean. Res.* **2022**, *129*, 103326. [\[CrossRef\]](#)
2. Kozjek, D.; Malus, A.; Vrabič, R. Reinforcement-Learning-Based Route Generation for Heavy-Traffic Autonomous Mobile Robot Systems. *Sensors* **2021**, *21*, 4809. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Long, P.; Fan, T.; Liao, X.; Liu, W.; Zhang, H.; Pan, J. Towards optimally decentralized multi-robot collision avoidance via deep reinforcement learning. In Proceedings of the 2018 IEEE International Conference on Robotics And Automation (ICRA), Brisbane, QLD, Australia, 21–25 May 2018.

4. Khatib, O. Real-time obstacle avoidance for manipulators and mobile robots. In Proceedings of the 1985 IEEE International Conference on Robotics and Automation, St. Louis, MO, USA, 25–28 March 1985. [\[CrossRef\]](#)
5. Bin Issa, R.; Das, M.; Rahman, S.; Barua, M.; Rhaman, K.; Ripon, K.S.N.; Alam, G.R. Double Deep Q-Learning and Faster R-CNN-Based Autonomous Vehicle Navigation and Obstacle Avoidance in Dynamic Environment. *Sensors* **2021**, *21*, 1468. [\[CrossRef\]](#) [\[PubMed\]](#)
6. López-Lozada, E.; Rubio-Espino, E.; Sossa-Azuela, J.H.; Ponce-Ponce, V.H. Reactive navigation under a fuzzy rules-based scheme and reinforcement learning for mobile robots. *PeerJ Comput. Sci.* **2021**, *7*, e556. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Kober, J.; Bagnell, J.A.; Peters, J. Reinforcement learning in robotics: A survey. *Int. J. Robot. Res.* **2013**, *32*, 1238–1274. [\[CrossRef\]](#)
8. Van Hasselt, H.; Guez, A.; Silver, D. Deep Reinforcement Learning with Double Q-Learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; Volume 30.
9. Song, H.; Li, A.; Wang, T.; Wang, M. Multimodal Deep Reinforcement Learning with Auxiliary Task for Obstacle Avoidance of Indoor Mobile Robot. *Sensors* **2021**, *21*, 1363. [\[CrossRef\]](#)
10. Masadeh, A.; Wang, Z.; Kamal, A. Convergence-Based Exploration Algorithm for Reinforcement Learning Recommended Citation. *Electr. Comput. Eng. Tech. Rep. White Pap.* **2018**, *1*, 1–11.
11. Xie, L.; Wang, S.; Markham, A.; Trigoni, N. Towards Monocular Vision based Obstacle Avoidance through Deep Reinforcement Learning. *arXiv* **2017**, arXiv:1706.09829.
12. Akanksha, E.; Jyoti; Sharma, N.; Gulati, K. Review on Reinforcement Learning, Research Evolution and Scope of Application. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021; pp. 1416–1423. [\[CrossRef\]](#)
13. Lemley, J.; Bazrafkan, S.; Corcoran, P. Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision. *IEEE Consum. Electron. Mag.* **2017**, *6*, 48–56. [\[CrossRef\]](#)
14. Kitchenham, B.; Brereton, O.P.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic literature reviews in software engineering—A systematic literature review. *Inf. Softw. Technol.* **2009**, *51*, 7–15. [\[CrossRef\]](#)
15. Zhu, Y.; Mottaghi, R.; Kolve, E.; Lim, J.J.; Gupta, A.; Fei-Fei, L.; Farhadi, A. Target-driven visual navigation in indoor scenes using deep reinforcement learning. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 3357–3364.
16. Liu, W.; Dong, L.; Niu, D.; Sun, C. Efficient Exploration for Multi-Agent Reinforcement Learning via Transferable Successor Features. *IEEE/CAA J. Autom. Sin.* **2022**, *9*, 1673–1686. [\[CrossRef\]](#)
17. Sengupta, A.; Chaurasia, R. Secured Convolutional Layer IP Core in Convolutional Neural Network Using Facial Biometric. *IEEE Trans. Consum. Electron.* **2022**, *68*, 291–306. [\[CrossRef\]](#)
18. Xiao, X.; Liu, B.; Warnell, G.; Stone, P. Motion planning and control for mobile robot navigation using machine learning: A survey. *Auton. Robot.* **2022**, *46*, 569–597. [\[CrossRef\]](#)
19. Fiorini, P.; Shiller, Z. Motion Planning in Dynamic Environments Using Velocity Obstacles. *Int. J. Robot. Res.* **1998**, *17*, 760–772. [\[CrossRef\]](#)
20. Havenstrøm, S.T.; Rasheed, A.; San, O. Deep Reinforcement Learning Controller for 3D Path Following and Collision Avoidance by Autonomous Underwater Vehicles. *Front. Robot. AI* **2021**, *7*, 211. [\[CrossRef\]](#)
21. Zelek, R.; Jeon, H. Characterization of Semantic Segmentation Models on Mobile Platforms for Self-Navigation in Disaster-Struck Zones. *IEEE Access* **2022**, *10*, 73388–73402. [\[CrossRef\]](#)
22. Joseph, L.M.I.L.; Goel, P.; Jain, A.; Rajyalakshmi, K.; Gulati, K.; Singh, P. A Novel Hybrid Deep Learning Algorithm for Smart City Traffic Congestion Predictions. In Proceedings of the 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 7–9 October 2021; pp. 561–565. [\[CrossRef\]](#)
23. Armstrong, R.; Hall, B.J.; Waters, E. Cochrane update. ‘Scoping the scope’ of a cochrane review. *J. Public Health* **2011**, *33*, 147–150. [\[CrossRef\]](#)
24. Petersen, K.; Vakkalanka, S.; Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Inf. Softw. Technol.* **2015**, *64*, 1–18. [\[CrossRef\]](#)
25. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Higgins, J.P.; Green, S. (Eds.) *Cochrane Handbook for Systematic Reviews of Interventions*; Wiley-Blackwell: Chichester, UK, 2008; Volume 5.
27. Moola, S.; Munn, Z.; Tufanaru, C.; Aromataris, E.; Sears, K.; Sfetcu, R.; Currie, M.; Qureshi, R.; Mattis, P.; Lisy, K.; et al. Chapter 7: Systematic Reviews of Etiology and Risk. In *JBI Manual for Evidence Synthesis*; Aromataris, E., Munn, Z., Eds.; JBI: Adelaide, Australia, 2020.
28. Wang, N.; Zhang, D.; Wang, Y. Learning to Navigate for Mobile Robot with Continual Reinforcement Learning. In Proceedings of the 39th Chinese Control Conference, Shenyang, China, 27–29 July 2020. [\[CrossRef\]](#)
29. Chen, C.; Jeng, S.; Lin, C. Mobile Robot Wall-Following Control Using Fuzzy Logic Controller with Improved Differential Search and Reinforcement Learning. *Mathematics* **2020**, *8*, 1254. [\[CrossRef\]](#)
30. Huang, R.; Qin, C.; Ling, J.; Lan, L. Path planning of mobile robot in unknown dynamic continuous environment using reward-modified deep Q-network. *Optim. Control Appl. Methods* **2021**, *44*, 1570–1587. [\[CrossRef\]](#)



31. Gao, R.; Gao, X.; Liang, P.; Han, F.; Lan, B.; Li, J.; Li, J.; Li, S. Motion Control of Non-Holonomic Constrained Mobile Robot Using Deep Reinforcement Learning. In Proceedings of the 2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM), Toyonaka, Japan, 3–5 July 2019; pp. 348–353.
32. Lee, H.; Jeong, J. Mobile Robot Path Optimization Technique Based on Reinforcement Learning Algorithm in Warehouse Environment. *Appl. Sci.* **2021**, *11*, 1209. [\[CrossRef\]](#)
33. Xiang, J.; Li, G.; Dong, X.; Ren, Z. Continuous Control with Deep Reinforcement Learning for Mobile Robot Navigation. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 1501–1506.
34. Fen, S.; Ren, H.; Wang, X.; Ben-Tzvi, P. Mobile Robot Obstacle Avoidance Base on Deep Reinforcement Learning. *Proc. ASME* **2019**, *5A*, 1–8.
35. Huang, L.; Qu, H.; Fu, M.; Deng, W. Reinforcement Learning for Mobile Robot Obstacle Avoidance under Dynamic Environments. In *PRICAI 2018: Trends in Artificial Intelligence*; Geng, X., Kang, B.H., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11012, pp. 441–453. [\[CrossRef\]](#)
36. Choi, J.; Lee, G.; Lee, C. Reinforcement learning-based dynamic obstacle avoidance and integration of path planning. *Intell. Serv. Robot.* **2021**, *14*, 663–677. [\[CrossRef\]](#)
37. Xiaoxian, S.; Chenpeng, Y.; Haoran, Z.; Chengju, L.; Qijun, C. Obstacle Avoidance Algorithm for Mobile Robot Based on Deep Reinforcement Learning in Dynamic Environments. In Proceedings of the 2020 IEEE 16th International Conference on Control & Automation (ICCA), Singapore, 9–11 October 2020; pp. 366–372. [\[CrossRef\]](#)
38. Song, S.; Kidziński, Ł.; Peng, X.B.; Ong, C.; Hicks, J.; Levine, S.; Delp, S.L. Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation. *J. Neuroeng. Rehabil.* **2021**, *18*, 1–17. [\[CrossRef\]](#)
39. Taghavifar, H.; Xu, B.; Taghavifar, L.; Qin, Y. Optimal path-planning of nonholonomic terrain robots for dynamic obstacle avoidance using single-time velocity estimator and reinforcement learning approach. *IEEE Access* **2019**, *7*, 159347–159356. [\[CrossRef\]](#)
40. Chewu, C.C.E.; Kumar, V.M. Autonomous navigation of a mobile robot in dynamic indoor environments using SLAM and reinforcement learning. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *402*, 012022. [\[CrossRef\]](#)
41. Ruan, X.; Ren, D.; Zhu, X.; Huang, J. Mobile Robot Navigation based on Deep Reinforcement Learning. In Proceedings of the 2019 Chinese Control and Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 6174–6178. [\[CrossRef\]](#)
42. Yao, Q.; Zheng, Z.; Qi, L.; Yuan, H.; Guo, X.; Zhao, M.; Yang, T. Path planning method with improved artificial potential field—A reinforcement learning perspective. *IEEE Access* **2020**, *8*, 135513–135523. [\[CrossRef\]](#)
43. Han, S.; Choi, H.; Benz, P.; Loaigiga, J. Sensor-Based Mobile Robot Navigation via Deep Reinforcement Learning. In Proceedings of the IEEE International Conference on Big Data and Smart Computing, Shanghai, China, 15–17 January 2018; pp. 147–154. [\[CrossRef\]](#)
44. Gao, J.; Ye, W.; Guo, J.; Li, Z. Deep Reinforcement Learning for Indoor Mobile Robot Path Planning. *Sensors* **2020**, *20*, 5493. [\[CrossRef\]](#)
45. Duguleana, M.; Mogan, G. Neural networks based reinforcement learning for mobile robots obstacle avoidance. *Expert Syst. Appl.* **2016**, *62*, 104–115. [\[CrossRef\]](#)
46. Zhang, W.; Wang, G. Reinforcement Learning-Based Continuous Action Space Path Planning Method for Mobile Robots. *J. Robot.* **2022**, *2022*, 9069283. [\[CrossRef\]](#)
47. Wang, Y.; Fang, Y.; Lou, P.; Yan, J.; Liu, N. Deep Reinforcement Learning based Path Planning for Mobile Robot in Unknown Environment. *J. Phys. Conf. Ser.* **2020**, *1576*, 012009. [\[CrossRef\]](#)
48. Lin, T.-C.; Chen, C.-C.; Lin, C.-J. Wall-following and Navigation Control of Mobile Robot Using Reinforcement Learning Based on Dynamic Group Artificial Bee Colony. *J. Intell. Robot. Syst.* **2018**, *92*, 343–357. [\[CrossRef\]](#)
49. Orozco-Rosas, U.; Picos, K.; Pantrigo, J.J.; Montemayor, A.S.; Cuesta-Infante, A. Mobile Robot Path Planning Using a QAPF Learning Algorithm for Known and Unknown Environments. *IEEE Access* **2022**, *10*, 84648–84663. [\[CrossRef\]](#)
50. Quan, H.; Li, Y.; Zhang, Y. A novel mobile robot navigation method based on deep reinforcement learning. *Int. J. Adv. Robot. Syst.* **2020**, *17*, 1729881420921672. [\[CrossRef\]](#)
51. Everett, M.; Chen, Y.F.; How, J.P. Motion planning among dynamic, decision-making agents with deep reinforcement learning. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Madrid, Spain, 1–5 October 2018; pp. 3052–3059.
52. Niroui, F.; Zhang, K.; Kashino, Z.; Nejat, G. Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments. *IEEE Robot. Autom. Lett.* **2019**, *4*, 610–617. [\[CrossRef\]](#)
53. Mehmood, A.; Shaikh, I.U.; Ali, A. Application of Deep Reinforcement Learning Tracking Control of 3WD Omnidirectional Mobile Robot. *Inf. Technol. Control* **2021**, *50*, 507–521. [\[CrossRef\]](#)
54. Chen, P.Z.; Pei, J.A.; Lu, W.Q.; Li, M.Z. A deep reinforcement learning based method for real-time path planning and dynamic obstacle avoidance. *Neurocomputing* **2022**, *497*, 64–75. [\[CrossRef\]](#)
55. Bernat, J.; Czopek, P.; Bartosik, S. Analysis of Mobile Robot Control by Reinforcement Learning Algorithm. *Electronics* **2022**, *11*, 1754. [\[CrossRef\]](#)
56. Wen, S.; Wen, Z.; Zhang, D.; Zhang, H.; Wang, T. A multi-robot path-planning algorithm for autonomous navigation using meta-reinforcement learning based on transfer learning. *Appl. Soft Comput.* **2021**, *110*, 107605. [\[CrossRef\]](#)

57. Baek, D.; Hwang, M.; Kim, H.; Kwon, D.S. Path planning for automation of surgery robot based on probabilistic roadmap and reinforcement learning. In Proceedings of the 2018 15th International Conference on Ubiquitous Robots (UR), Honolulu, HI, USA, 26–30 June 2018; pp. 342–347.
58. Yu, J.; Su, Y.; Liao, Y. The Path Planning of Mobile Robot by Neural Networks and Hierarchical Reinforcement Learning. *Front. Neurobotics* **2020**, *14*, 63. [[CrossRef](#)] [[PubMed](#)]
59. Zhu, W.; Guo, X.; Fang, Y.; Zhang, X. A Path-Integral-Based Reinforcement Learning Algorithm for Path Following of an Autoassembly Mobile Robot. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 4487–4499. [[CrossRef](#)] [[PubMed](#)]
60. Bai, W.; Li, T.; Tong, S. NN Reinforcement Learning Adaptive Control for a Class of Nonstrict-Feedback Discrete-Time Systems. *IEEE Trans. Cybern.* **2020**, *50*, 4573–4584. [[CrossRef](#)] [[PubMed](#)]
61. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 3104–3112.
62. Zhang, T.; Mo, H. Reinforcement learning for robot research: A comprehensive review and open issues. *Int. J. Adv. Robot. Systems* **2021**, *18*, 1–22. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.