aerospace

MDPI

Article

Air Channel Planning Based on Improved Deep Q-Learning and Artificial Potential Fields

Jie Li, Di Shen *, Fuping Yu and Renmeng Zhang

Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China; 18681899056@163.com (J.L.); junjingj@163.com (F.Y.); 18600588220@163.com (R.Z.)
* Correspondence: shendi1103@163.com

Abstract: With the rapid advancement of unmanned aerial vehicle (UAV) technology, the widespread utilization of UAVs poses significant challenges to urban low-altitude safety and airspace management. In the coming future, the quantity of drones is expected to experience a substantial surge. Effectively regulating the flight behavior of UAVs has become an urgent and imperative issue that needs to be addressed. Hence, this paper proposes a standardized approach to UAV flight through the design of an air channel network. The air channel network comprises numerous single air channels, and this study focuses on investigating the characteristics of a single air channel. To achieve optimal outcomes, the concept of the artificial potential field algorithm is integrated into the deep Q-learning algorithm during the establishment of a single air channel. By improving the action space and reward mechanism, the resulting single air channel enables efficient avoidance of various buildings and obstacles. Finally, the algorithm is assessed through comprehensive simulation experiments, demonstrating its effective fulfillment of the aforementioned requirements.

Keywords: air channel network; path planning; urban air traffic control; DQN

check for updates

Citation: Li, J.; Shen, D.; Yu, F.; Zhang, R. Air Channel Planning Based on Improved Deep Q-Learning and Artificial Potential Fields. *Aerospace* **2023**, *10*, 758. https://doi.org/10.3390/aerospace10090758

Academic Editor: Álvaro Rodríguez-Sanz

Received: 13 June 2023
Revised: 23 August 2023
Accepted: 25 August 2023
Published: 27 August 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

In recent years, there has been a growing recognition of the convenience and practicality of unmanned aerial vehicles (UAVs), leading to their widespread applications in various fields such as earth science, traffic management, pollution monitoring, and product delivery [1]. This increasing commercialization of drones, however, does not imply a decline in civilian drone usage. Statistics indicate a steady rise in the frequency and quantity of public drones, predominantly employed for aerial photography and recreational purposes. Looking ahead, civilian drones are anticipated to replace conventional means of transport in sectors like express delivery and logistics. Yet, the extensive use of drones will inevitably present significant challenges in terms of urban drone traffic management, posing potential security threats to cities. Consequently, effectively regulating the flight behavior of urban UAVs in the future, and mitigating in-flight conflicts and collisions with urban infrastructure, has emerged as a pressing issue that requires immediate attention.

In response to these challenges, several scholars have presented their perspectives on potential solutions. Ali [2], for instance, defines the functions of an urban UAV traffic management system and conducts exploratory research to discern the disparities between manned and unmanned driving, thereby identifying the current challenges. Based on the research findings, Ali proposes an architectural framework for urban UAV traffic management. Mohammed Faisal Bin Mohammed Salleh [3], on the other hand, constructs three networks: the AirMatrix route network, the route network above buildings, and the route network on roads. Through an analysis of these networks in terms of capacity and throughput, Salleh evaluates their performance. Additionally, Timothy McCarthy [4] examines the approaches taken by the United States and Europe in unmanned transportation. He introduces a comprehensive method of airspace modeling and a traffic management

platform developed by his research team, offering unique insights for urban UAV traffic management.

To address these challenges, this paper proposes a fundamental concept of establishing an urban low-altitude UAV channel network, designs an air traffic network based on the existing ground road network, formulates the flight methodology for the air channels, and provides a detailed analysis of the utilization of the air channel network. The second chapter of the paper presents the specific details of these approaches. Furthermore, to effectively implement this design concept, this paper focuses on designing a single UAV air channel guided by the aforementioned ideas. This single air channel serves as a foundational element for the subsequent formation of an air channel network comprised of multiple interconnected air channels.

When establishing a single air channel, it is crucial to consider the impact of surrounding high-rise buildings in the city. Entering such areas not only jeopardizes the safety of UAV flights but also poses significant risks to urban structures. Classical methods employed in air channel planning under similar circumstances encompass the genetic algorithm [5], artificial potential field algorithm [6], A* algorithm [7], antcolony algorithm [8], and particle swarm algorithm [9]. In recent years, scholars have made noteworthy advancements in this field. For instance, Seyedali Mirjalili introduced the whale algorithm, which mimics the behavior of whales, offering fresh insights for optimization problems [10]. Abdel-Basset Mohamed analyzed the principles and applications of the flower pollination algorithm, compared its efficacy with other algorithms, and evaluated its performance [11]. Wenguan Luo introduced diverse enhancement strategies to the Cuckoo Search (CS) algorithm and proposed the Novel Enhanced CS Algorithm (NECSA) to improve convergence speed and accuracy [12]. Yinggao Yue provided a comprehensive analysis of the sparrow search algorithm, highlighting its strengths, weaknesses, and influencing factors, while proposing relevant improvement strategies [13].

When delineating a single air channel, path planning alone may not suffice to meet the requirements of different stages within the UAV air channel network. Classical path planning algorithms have some effectiveness in addressing these challenges. However, the Deep Q-learning (DQN) algorithm with a reward mechanism can offer an alternative perspective to tackle such problems. By enabling the UAV to continuously interact with its environment, the DQN algorithm helps the UAV discover optimal behaviors [14]. To overcome the limitation of the DQN algorithm's generalization ability, Chunling Liu [15] proposed a multi-controller model combined with fuzzy control. This approach provided ample positive samples for the model at the initial stage of the experiment, which enhance training efficiency. Yunwan Gu [16] resolved the M-DQN algorithm into a value function and an advantage function based on its network structure, leading to accelerated convergence and improved generalization performance. While these approaches improved pathfinding, convergence of the final reward and experimental results varied. Siyu Guo [17] optimized the traditional reward function by incorporating a potential energy reward for the ship at the target point. This approach facilitated the ship's movement towards the target point, significantly reducing calculation time. However, the final test and simulation processes employed relatively simplistic environmental designs, making it challenging to verify the experimental efficacy under more complex environments and constraints. Lei Luo [18] introduced the A* algorithm into the deep reinforcement learning framework, creating the AG-DQN algorithm, which achieved promising results in solving resource management and flight scheduling problems. Therefore, to address the complexities of path planning in continuous action environments, this paper integrates the artificial potential field algorithm into the reward mechanism and action space of deep reinforcement learning and presents a viable approach to achieve the desired outcomes.

2. Design Principle of the Air Channel Network

Numerous studies have been conducted on the establishment of unmanned aerial vehicle (UAV) channels in urban airspace. For instance, Qingyu Tan et al. [19] proposed a

city air route planning system consisting of three components: traffic network construction, route planning, and take-off time scheduling. When a flight request is received, route planning is performed, and takeoff time scheduling is conducted to minimize conflicts, thus presenting a novel approach to urban UAV airspace management.

In this paper, we present the construction of a new aerial corridor structure and flight regulations to address flight conflicts. We engage in discussions aimed at alleviating these conflicts.

When UAVs fly in urban areas, they must adhere to pre-programmed rules based on the established air channel network, as illustrated in Figure 1. The air channel network comprises three components: Class I channel, Class II channel, and Class III channel. The delineation of these channels is organized in accordance with the requirements of UAV air control. This paper assumes the following requirements for UAV air control:

1. Adherence to Air Channel Network: The UAV must strictly follow the designated air channel network. It should not deviate from the pre-defined routes and paths outlined in the network.
2. Diversion Areas: The air channel network should include diversion areas to accommodate the flight requirements of UAVs heading to different destinations.
3. Varying Heights: The air channel network has different height restrictions in different regions. The single air channel planning considers only the two-dimensional plane situation, without accounting for vertical dimensions.
4. Ground Identification Devices: UAVs need to be equipped with the capability to respond to ground identification devices. These devices ensure proper communication and identification between the UAV and ground control systems.
5. Directional Flight: When operating within the air channel network, UAVs must strictly adhere to the prescribed direction of the channel. They should not change their flight direction at will while flying within the designated channels.

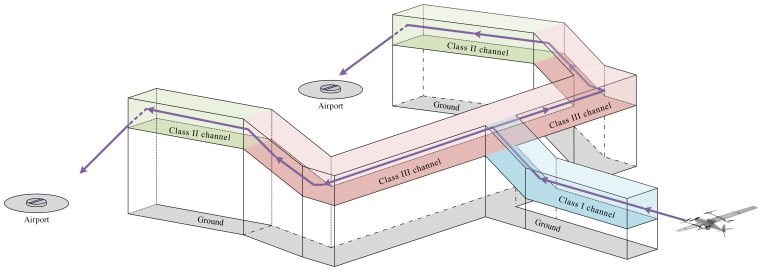


Figure 1. Diagram of air channel.

According to the above requirements, the air channel principle is defined as follows.

The air channel network consists of three main components: Class I channel, Class II channel, and Class III channel, shown in Table 1. Class I Channel is comparable to the arterial roads in a ground road network. It serves as the primary component of the air channel network. UAVs can fly at relatively high speeds within this channel while maintaining a safe distance from other drones. The Class II channel connects the take-off or landing areas to the Class I channel. It functions similar to the connecting roads that link destinations to the arterial roads in a ground road network. In this channel, there is a requirement for information identification of drones to prevent destination errors caused by system errors. Additionally, to enhance safety in the take-off or landing area, drones should operate at lower speeds within this channel. The Class III channel primarily serves as a diversion area within the air channel network. It plays a role similar to intersections and overpasses in a ground road network. UAVs navigate through these areas and adjust their flight direction based on their destination. Class III channels are usually found at intersections and only a few channels connect these intersections. The height changes depicted in the Figure 1 are not significant in practical terms.

Table 1. Definition of air channel.

Classification	Definition	Flight Speed Limit
Class I channel	The main component of the air channel network	high speed
Class II channel	The channel near the take-off and landing area	low speed
Class III channel	Mainly responsible for the shunt function	intermediate speed

To accommodate the needs of UAV direction changes and prevent collisions at intersections, the air channel network is designed with specific height and width parameters. Additionally, regulations regarding UAV steering are specified to ensure safe and efficient operations. This paper focuses on the example of right flight.

To minimize interactions between different aircraft within the same channel, an air channel is designed to take inspiration from ground highways. The configuration of the air channel is depicted in Figure 2.

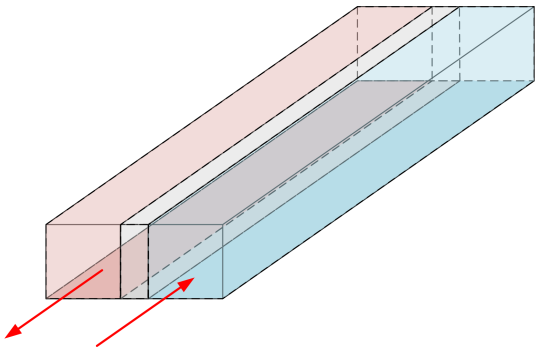


Figure 2. Diagram of the air channel structure.

In Figure 2, the air channel is divided horizontally into three parts: the one-way channels on both sides and the anti-collision channel in the middle. The anti-collision channel is specifically designed to minimize interactions between drones flying in opposite directions. It ensures a sufficient horizontal flight interval and enhances safety.

The air channel network differs from ground roads in that drones cannot wait in the air, making it impossible to install traffic lights at intersections to coordinate steering demands. To meet the requirements of drone flight, this paper proposes the configuration shown in Figure 3.

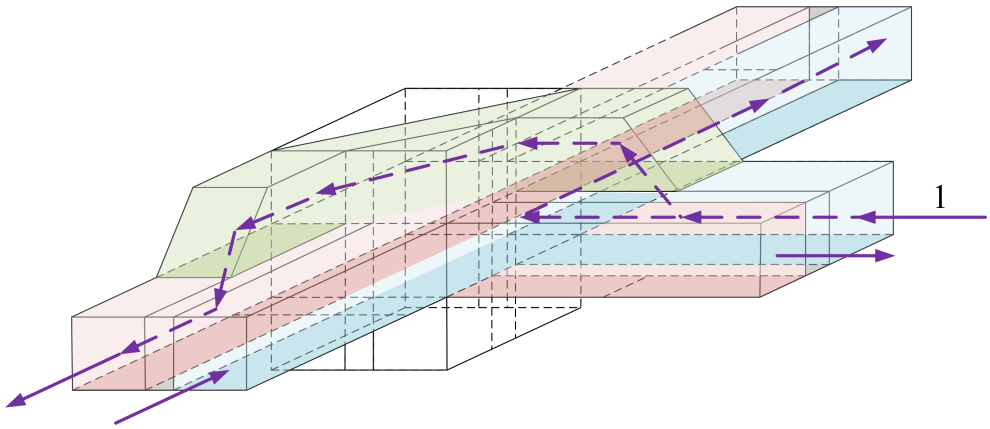


Figure 3. Diagram of flights at intersections.

In the intersection, the flight channel is positioned in the middle, with up and down steering channels on either side. The flight channels for north-south and east-west directions are set at different heights to eliminate potential conflicts. However, conflicts between left

and right directions cannot be resolved through this method. Therefore, the following provisions are proposed:

1. The right turn needs to be made in the flight channel layer;
2. The east-west drone needs to turn left in the upper side turn channel, and the north-south drone needs to turn left in the lower side turn channel.

As shown in Figure 3, when a UAV follows route 1 into the intersection, it lifts height and enters the green steering channel prior to reaching the intersection. It then proceeds into the upper steering channel along the left designated area and merges into the north-south air channel. The same procedure applies for a north-south UAVs entering the lower steering channel.

By designing the aforementioned concept of a UAV air traffic network, flight conflicts can be effectively resolved, and the flight behavior of urban UAVs can be regulated. To design the air traffic channel network, this paper analyzes a single air channel, considering its consistency between the final result and theoretical principles to verify the scientific basis of the theory.

3. Traditional Deep Q-Learning and Artificial Potential Field Algorithm

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation as well as the experimental conclusions that can be drawn.

3.1. DQN

Deep Q-learning (DQN) is an algorithm improved by Q-Learning [20–22]. The expression of the traditional reinforcement Q-learning algorithm is shown in the following formula:

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

In the above equation, $Q(s,a)$ represents the state-action value function, s represents the current state of the agent, and a represents the action taken under the current state. $\max_a Q(s',a')$ represents the maximum state-action value function obtained by taking action a' under the next state s'; α represents the learning rate and γ represents the rate of decay of reward r. The solution of the Q-learning algorithm is obtained according to the final $Q(s,a)$ value table; however, it is easy for this algorithm to produce dimension disaster [23], and so cannot be popularized in practical application.

DQN combines Deep Learning [24–26] and Q-Learning and the specific principle is shown in the Figure 4.

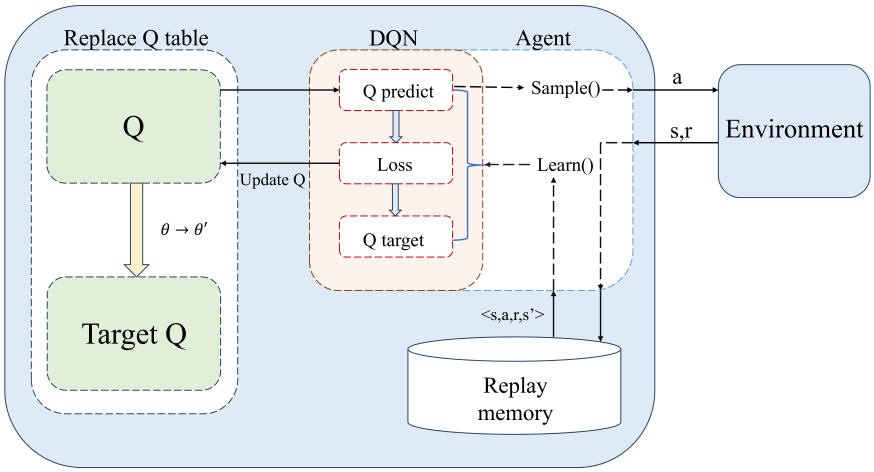


Figure 4. Principles of deep Q-learning.

In the Figure 4, DQN consists of two networks with the same structure. The Q network generates experience through interaction with the environment, continuously optimizing the target network. The historical experience generated by the Q network is stored in the reply memory pool. In the running process of the algorithm, the current strategy is evaluated first, and then the strategy is improved. In addition, the reward of the executable action in the next state is calculated, and the action execution with the maximum reward is selected.

When algorithm ε-greedy is used to randomly select an action in the action space N_{action} , when probability $P = \varepsilon$ is adopted, an action is randomly selected from the action space. When the probability is $P = 1 - \varepsilon$, the action corresponding to the maximum state-action value function $\arg\max Q(s, a)$ is selected, as shown in the following formula:

$$a = \begin{cases} \text{random } N_{\text{action},} & P = \varepsilon \\ \arg\max Q(s, a), & P = 1 - \varepsilon \end{cases} \tag{2}$$

In DQN, the convolutional neural network is used to define the optimal state-action value function $Q(s, a)$ as $Q(s, a; \theta)$, and θ represents the parameter of the neural network [27]. The updating rules are as follows:

$$\theta_{t+1} = \theta_t + \alpha \left[r + \gamma \max_{a'} Q\left(s', a', \theta^-\right) - Q(s, a; \theta) \right] \nabla Q(s, a; \theta) \tag{3}$$

Loss function $(L(\theta))$ of DQN refers to the difference between the predicted value of the current network and the real value of the target network, which is used to update the state-action value function. $L(\theta)$ is as follows:

$$L(\theta) = E\left[\left(r + \gamma \max_{a'} Q\left(s', a', \theta^-\right) - Q(s, a; \theta) \right] \right)^2 \tag{4}$$

The experience gained from each interaction between the agent and the environment is stored in the form of $e_t = (s_t, a_t, r_t, s_{t+1})$ in the reply memory pool $M_t = (e_1, e_2, \ldots, e_t)$. During the training of the neural network, a batch is randomly selected to reduce the correlation between the data of each training and enhance the scientific nature and stability of the network.Therefore, the expression for the update iteration of the action-state value function is as follows:

$$Q_{t+1}(s, a) = E_{s'} \left[r + \gamma \max_{a'} Q_t\left(s', a'\right) \right] \tag{5}$$

3.2. APF

APF is a path planning problem that seeks the optimal solution by establishing the potential field distribution of the whole environment. In the original APF algorithm, assuming that the distance between the agent and the target is D_{am}, the distance between the agent and the obstacle is D, the distance threshold of the repulsive force field is D_{rep}, and the gravitational and repulsive coefficients are ξ and η, respectively. Consequently, the gravitational field function U_{att} and the repulsive force field function U_{rep} are respectively

$$\begin{gathered} U_{att} = \frac{1}{2} \xi D_{am}^2 \\ U_{rep} = \begin{cases} \frac{1}{2} \eta \left(\frac{1}{D} - \frac{1}{D_{rep}} \right)^2, & D \leq D_{rep} \\ 0 & D > D_{rep} \end{cases} \end{gathered} \tag{6}$$

Then, the total potential field U and resultant force F are

$$\begin{aligned} U &= U_{att} + U_{rep} \\ F &= -\nabla U \end{aligned} \tag{7}$$

4. Aerial Channel Planning Based on APF-DQN

To effectively avoid tall buildings during flight and reach the destination efficiently, a combination of the artificial potential field (APF) algorithm and deep Q-learning (DQN) algorithm is proposed. The APF algorithm is known for its effectiveness in obstacle avoidance and attraction towards the destination. In this approach, the advantages of APF are integrated into the action space and reward mechanism of the DQN algorithm. Designing a reward that attracts the UAV to the destination continuously aligns the resultant force direction with the direction of the destination. Furthermore, by designing rewards for aircraft undergoing continuous actions and considering the selection of different actions at each step, we address the limitations of DQN when dealing with discrete spaces. Additionally, the repulsion generated by other building elements within the APF has a significant impact on action choice. Thus, the UAV learns to automatically avoid building areas.

In the course of flight return, it is assumed that the safe distance between the aircraft and the airspace is D_1, the navigation and positioning error is D_2, the minimum response distance of the pilot is D_3, and the normal flight between the aircraft and the airspace is $D_{\max}$. When the distance between the aircraft and the airspace is within $D_1 \leq D < D_{\max}$, this constitutes the safe zone (SZ). When the distance between the aircraft and the airspace is $D_3 < D < D_1$, the aircraft is diverted to ensure safety, defining the diverted area. When the distance between the aircraft and the airspace is $D < D_3$, compulsory actions will be taken to avoid the airspace and alter the flight area, as depicted in the Figure 5.

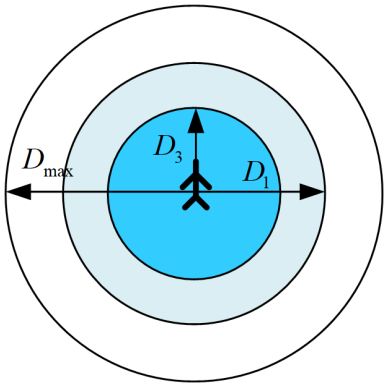


Figure 5. Diagram of aircraft diversion area.

4.1. Improve Action Space

The original DQN algorithm excels in handling discrete spaces within a raster map environment, but it falls short when it comes to continuous action spaces [28]. Consequently, the action selection mechanism of the DQN algorithm is modified to utilize a unit vector that represents the desired flight direction of the aircraft. The resultant force F of APF is put into the activation function sigmoid(x), and the resultant force F is mapped to the value τ between $(0, 1)$.τ is multiplied by the basic action step v_{base} to obtain the final action step, namely the distance of each action. The activation functions sigmoid(x) and τ are as follows:

$$\begin{aligned} \text{sigmoid}(x) &= \frac{1}{1+e^{-x}} \\ \tau &= \text{sigmoid}(F) \end{aligned} \tag{8}$$

When the aircraft is in the safe area, it is only affected by the target gravity, and its motion direction can be obtained by adding the output direction $\overrightarrow{A_q}$ of DQN and the direction $\overrightarrow{F_a}$ of gravity. Thus, the motion of the aircraft can be written as follows:

$$A = \left(\overrightarrow{A_q} + \overrightarrow{F_a}\right) v_{\text{base}} \tag{9}$$

When the aircraft approaches the airspace and is in the rerouting area, the movement of the aircraft can be expressed as follows:

$$A = \tau\left(\vec{F} + \overrightarrow{A_q}\right) v_{base} \tag{10}$$

As evident from the equation above, the closer the aircraft is to the airspace, the longer the step length of the action becomes, which is more advantageous for quick avoidance. When the aircraft approaches the airspace and is in the mandatory course change area, assuming that $\overrightarrow{A_C}$ is the direction of avoiding collision, the action of the aircraft is as follows:

$$A = \left(\frac{1}{2}\overrightarrow{A_q} + \overrightarrow{A_C}\right) v_{base} \tag{11}$$

4.2. Improve the Reward Mechanism

In the traditional DQN algorithm, only when the agent contacts the obstacle or reaches the end point can there be a reward, and other actions do not have any effective feedback. Therefore, to solve this problem, a new reward mechanism is proposed, as follows:

$$R = \begin{cases} R_a, & D \geq D_1 \\ R_b, & 0 < D < D_1 \\ R_{\text{end}}, & \text{others} \end{cases} \tag{12}$$

In order to make the aircraft reach the end point faster, the reward mechanism is set according to the distance and Angle between the end point and the aircraft. Set the distance from the starting point to the end point as $d_{\max}$, the distance between the current position and the end point as d_{goal}, the direction Angle of gravity in the artificial potential field as ϕ_{att}, and the direction Angle of the current aircraft as ϕ_a. In accordance with the premise of the question, the closer the aircraft is to the end point, the higher the reward value. Thus, the reward calculation is expressed as follows:

$$R_a = \text{sigmoid}\left(\left|\frac{\phi_a - \phi_{att}}{\phi_{att}} \bullet \frac{d_{\max} - d_{\text{goal}}}{d_{\max}}\right|\right) \tag{13}$$

When the aircraft is located in the rerouted area, the reward value should decrease as the distance between the aircraft and the airspace decreases. In this region, the aircraft receives the combined action of the gravitational and repulsive forces in the artificial potential field, the resultant force is F, and the resultant direction Angle is ϕ_F. Thus, the reward function is as follows:

$$\begin{gathered} R_b = \begin{cases} \frac{1}{2}(\lambda_1 + \lambda_2)\frac{D}{D_1}, & D_2 \leq D < D_1 \\ \frac{1}{2}\lambda_2\frac{D}{D_2}, & 0 \leq D < D_2 \end{cases} \\ \lambda_1 = \text{sigmoid}\left(\left|\frac{\phi_a - \phi_{att}}{\phi_{att}}\right|\right) \\ \lambda_2 = \text{sigmoid}\left(\left|\frac{\phi_a - \phi_F}{\phi_F}\right|\right) \end{gathered} \tag{14}$$

5. Simulation Verification

During the flight of UAVs, the buildings significantly affect the airspace, requiring the avoidance of such areas when designing the layout of air channels. According to potential situations during return flights, higher building areas should be added into the environment as constraining factors. This enables the creation of an urban environment model, as shown in Figure 6.

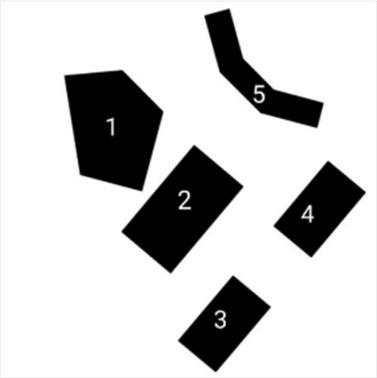


Figure 6. Flight environment.

In Figure 6, areas 1, 2, 3, 4 and 5 represent the higher buildings in the city. In order to ensure flight safety and not affect the normal operation of other tasks, the UAV should pay attention to avoid these areas when returning.

The deep neural network diagram in this method is illustrated in Figure 7. The input information comprises the UAV's position, the distance between the endpoint and the UAV, angle and the perceived obstacle information surrounding the UAV. The network processes this information to derive the UAV's current status, reward, and ultimately output action.

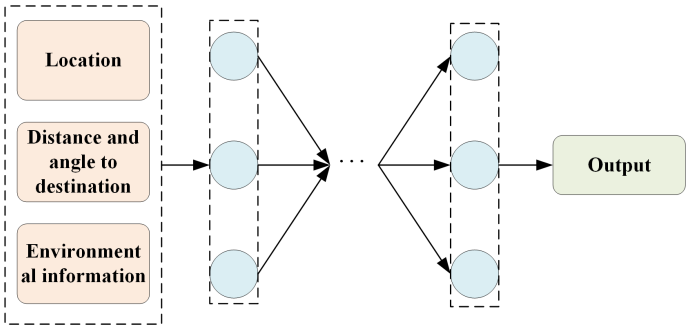


Figure 7. Neural network structure.

To ensure the UAV can locate the destination and minimize the path length, the method set specific rewards and penalties . Returning to the endpoint is rewarded with a value of 200, while mistakenly entering another building area carries a penalty of −200. Moreover, each action undertaken by the UAV incurs a reduction in the final reward value. The parameters for the method are determined in Table 2.

Table 2. Argument of DQN.

Argument	Symbol	Numerical Value
Reward decay rate	$\gamma_{\circ}$	0.9
Learning rate	lr	0.01
ε-greedy	ε	0.1
Experience replay memory	$M_{\max}$	5000
Replay batch size	b	64
Target net update interval	c	50

5.1. DQN Training

Based on the given parameters and environment, the simulation results depict the number of steps taken in each episode, as depicted in Figure 8. Initially, the first 200 flights fail to reach the destination stabilized, resulting in big fluctuations in the number of steps. Throughout this learning process, the aircraft continually learns from the experience repository and explores different available air channels. After undergoing over 400 training iterations, the number of steps gradually stabilizes, eventually converging within a consistent range. This indicates that the aircraft has achieved stable navigation towards the destination, with relatively optimal channel selection.

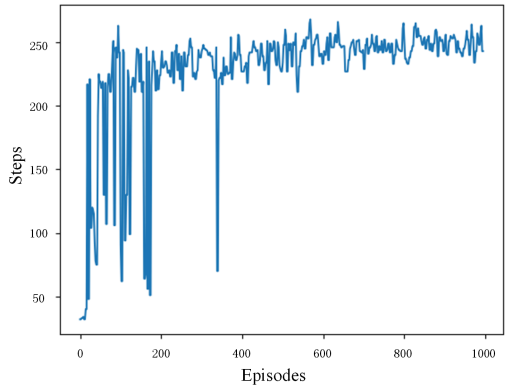


Figure 8. The numbers of steps per training.

As the Figure 9 shows, in the first image, which represents the initial stage, the aircraft struggles to arrive at the endpoint accurately. Consequently, it mistakenly enters the building area, resulting in a penalty. However, as the training progresses, the aircraft gradually improves its ability to arrive at the endpoint and begins to receive rewards for successfully reaching the destination, as depicted in the second image. While the reward value changes based on the length of the channel, the optimal channel is yet to be discovered at this stage. The training process focuses on learning and exploring various routes, gradually refining the aircraft's navigation strategy.

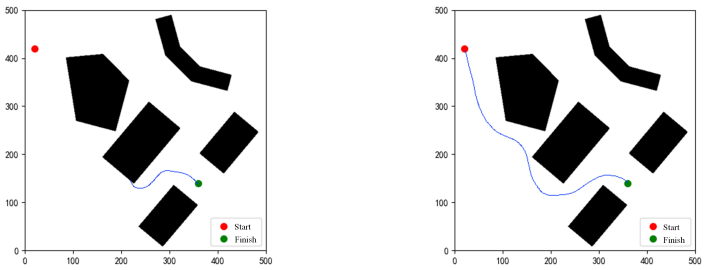


Figure 9. Total distance and convergence rate.

Figure 10 presents the evolution of reward values as training progresses. Analysis of the figure reveals several observations: In the initial 100 training episodes, the reward experiences big fluctuations. This is attributed to the aircraft frequently straying into other building areas and failing to locate the optimal air channel. At this stage, the aircraft is still exploring different areas and struggling to find the most efficient route. Between 100 and 500 training iterations, the aircraft gradually improves its ability to arrive at the endpoint and receives rewards accordingly. However, there are still instances where the aircraft enters the building areas. These results indicate that the process still needs to be optimized. After 500 training iterations, the aircraft successfully identifies the optimal solution, resulting in a stable reward.

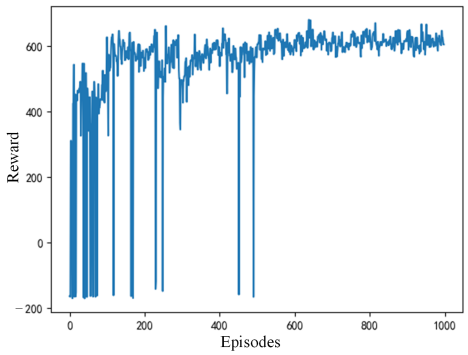


Figure 10. Reward for each training episode.

5.2. Comparative Analysis

The APF-DQN algorithm offers several advantages compared to simple APF, DQN, and Q-Learning algorithms. It effectively addresses the limitations of traditional DQN and Q-Learning algorithms in continuous space problems, while also overcoming issues related to weak generalizability and slow solving speed. The three algorithms can be compared based on their convergence speed and total displacement degree:

Figure 11 demonstrates the significant progress made by the improved algorithm in terms of the total distance traveled. It effectively reduces the redundant steering observed in the APF algorithm's path selection, while also avoiding the issue of traditional machine learning algorithms taking longer routes. Consequently, the final number of steps in the improved algorithm is considerably lower than the other two algorithms, indicating improved efficiency. Figure 12 provides further insights into the improved algorithm's performance. The line color in the figure is the same as in Figure 11. The first image illustrates that the improved algorithm reaches the final goal at a much faster speed compared to the other two algorithms. Its convergence speed is also notably faster. The second image showcases the superior performance of the improved algorithm, as it achieves the least number of steps and the smallest total distance traveled.

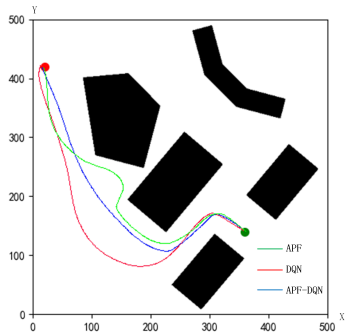


Figure 11. Comparison of three algorithm paths.

Figure 11 showcases that the aircraft has successfully navigated through the urban environment while avoiding buildings, and has made a shorter overall air channel length. This aligns with the analysis presented in Chapter 2, which explores the various components of the air channel network, and it also can be observed and analyzed in Figure 11. For example, in Figure 11 during the turning process, if there are multiple endpoints in the environment, forming a network of air channels composed of a single channel, the turning segment in the figure belongs to an intersection, which can be classified into the Class III channel. The channel used for progressing towards the endpoint after completing the turn belongs to the Class I channel. The channel near the endpoint belongs to the

Class II channel. The design demonstrates reasonable characteristics and exhibits strong adaptability in guiding the aircraft through the urban environment.

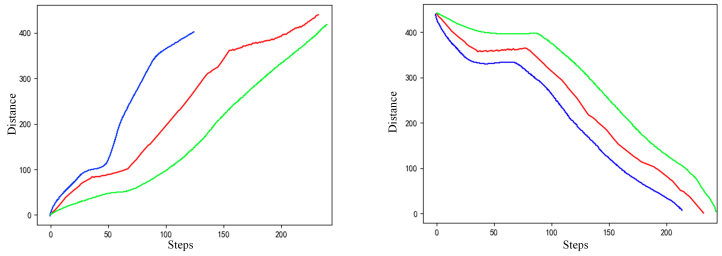


Figure 12. Total distance and convergence rate.

6. Summarize

This paper focuses on regulating the air specifications for urban UAVs and proposes the fundamental theory of UAV air channel networks. It analyzes the structure of the air channel and flight rules at intersections, contributing valuable insights to the development of urban UAV flights. Furthermore, the paper addresses the issue of delimiting a single air channel by integrating the artificial potential field algorithm into deep Q-learning. This integration enhances the reward mechanism and action space, resulting in significant advancements compared to traditional deep Q-learning and artificial potential field algorithms. The improved algorithm effectively meets the requirements of different components within the air channel network and successfully achieves the objective of air channel delimitation. These findings serve as a reference for future research in this domain.

We acknowledge the limitations of this paper, which is confined to the establishment of a single air channel in a two-dimensional space. The current work does not consider the complexities arising from height regulations and requirements in the air channel design. Future research will focus on addressing these shortcomings and expanding the scope of the study. The next steps will involve extending the analysis to three-dimensional space to incorporate height regulations and requirements. This will enhance the practicality and realism of the proposed air channel network. By addressing these deficiencies and advancing the research in these aspects, we aim to contribute to the development of a more robust and adaptable theory for UAV flights in urban areas.

Author Contributions: Conceptualization, J.L. and D.S.; Methodology, J.L.; Software; J.L., R.Z. and F.Y.; Writing—original draft, J.L.; Writing—review & editing, D.S., F.Y. and R.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors are thankful to the anonymous reviewers for their instructive reviewing of the manuscript.

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

References

1. Poudel, S.; Arafat, M.Y.; Moh, S. Bio-Inspired Optimization-Based Path Planning Algorithms in Unmanned Aerial Vehicles: A Survey. *Sensors* **2023**, *23*, 3051. [CrossRef] [PubMed]
2. Ali, B.S. Traffic Management for Drones Flying in the City. *Int. J. Crit. Infrastruct. Prot.* **2019**, *26*, 100310. [CrossRef]
3. Mohamed Salleh, M.F.B.; Wanchao, C.; Wang, Z.; Huang, S.; Tan, D.Y.; Huang, T.; Low, K.H. Preliminary Concept of Adaptive Urban Airspace Management for Unmanned Aircraft Operations. In Proceedings of the 2018 AIAA Information Systems-AIAA Infotech@ Aerospace, Kissimmee, FL, USA, 8–12 January 2018.
4. McCarthy, T.; Pforte, L.; Burke, R. Fundamental Elements of an Urban UTM. *Aerospace* **2020**, *7*, 85. [CrossRef]

5. Katoch, S.; Chauhan, S.S.; Kumar, V. A review on genetic algorithm: Past, present, and future. *Multimed. Tools Appl.* **2021**, *80*, 8091–8126. [CrossRef] [PubMed]
6. Jiang, Q.; Cai, K.; Xu, F. Obstacle-avoidance path planning based on the improved artificial potential field for a 5 degrees of freedom bending robot. *Mech. Sci.* **2023**, *14*, 87–97. [CrossRef]
7. Persson, S.M.; Sharf, I. Sampling-based A* algorithm for robot path-planning. *Int. J. Robot. Res.* **2014**, *33*, 1683–1708. [CrossRef]
8. Yang, X.S. Nature-Inspired Optimization Algorithms. 2014. Available online: https://www.researchgate.net/publication/263171713_Nature-Inspired_Optimization_Algorithms (accessed on 12 June 2023).
9. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948. [CrossRef]
10. Mirjalili, S.; Lewis, A. The Whale Optimization Algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]
11. Abdel-Basset, M.; Shawky, L.A. Flower pollination algorithm: A comprehensive review. *Artif. Intell. Rev.* **2019**, *52*, 2533–2557. [CrossRef]
12. Wenguan, L.; Xiaobing, Y. A Novel Enhanced Cuckoo Search Algorithm for Global Optimization. *Expert Syst. Appl.* **2022**, *43*, 2945–2962.
13. Yue, Y.; Cao, L.; Lu, D.; Hu, Z.; Xu, M.; Wang, S.; Li, B.; Ding, H. Review and empirical analysis of sparrow search algorithm. *Artif. Intell. Rev.* **2023**, *56*, 10867–10919. [CrossRef]
14. Sivaranjani, A.; Vinod, B. Artificial potential field incorporated deep-q-network algorithm for mobile robot path prediction. *Intell. Autom. Soft Comput.* **2023**, *35*, 1135–1150. [CrossRef]
15. Liu, C.; Xu, J.; Guo, K. Path Planning for Mobile Robot Based on Deep Reinforcement Learning and Fuzzy Control. In Proceedings of the 2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), Xi'an, China, 3–5 November 2022; pp. 533–537. [CrossRef]
16. Gu, Y.; Zhu, Z.; Lv, J.; Shi, L.; Hou, Z.; Xu, S. DM-DQN: Dueling Munchausen deep Q network for robot path planning. *Complex Intell. Syst.* **2022**, *9*, 4287–4300. [CrossRef]
17. Guo, S.; Zhang, X.; Du, Y.; Zheng, Y.; Cao, Z. Path Planning of Coastal Ships Based on Optimized DQN Reward Function. *J. Mar. Sci. Eng.* **2021**, *9*, 210. [CrossRef]
18. Luo, L.; Zhao, N.; Zhu, Y.; Sun, Y. A* guiding DQN algorithm for automated guided vehicle pathfinding roblem of robotic mobile fulfillment systems. *Comput. Ind. Eng.* **2023**, *178*, 109112. [CrossRef]
19. Tan, Q.; Wang, Z.; Ong, Y.-S.; Low, K.H. Evolutionary Optimization-based Mission Planning for UAS Traffic Management (UTM). In Proceedings of the International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, 11–14 June 2019; pp. 952–958.
20. Wang, Y.; Li, T.S.; Lin, C. Backward Q-learning: The combination of Sarsa algorithm and Q-learning. *Eng. Appl. Artif. Intell.* **2013**, *26*, 2184–2193. [CrossRef]
21. Jang, B.; Kim, M.; Harerimana, G.; Kim, J.W. Q-Learning Algorithms: A Comprehensive Classification and Applications. *IEEE Access* **2019**, *7*, 133653–133667. [CrossRef]
22. Zhang, N.; Cai, W.; Pang, L. Predator-Prey Reward Based Q-Learning Coverage Path Planning for Mobile Robot. *IEEE Access* **2023**, *11*, 29673–29683. [CrossRef]
23. Yuan, J.; Wang, H.; Zhang, H.; Lin, C.; Yu, D.; Li, C. AUV Obstacle Avoidance Planning Based on Deep Reinforcement Learning. *J. Mar. Sci. Eng.* **2021**, *9*, 1166. [CrossRef]
24. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [CrossRef]
25. Sünderhauf, N.; Brock, O.; Scheirer, W.; Hadsell, R.; Fox, D.; Leitner, J.; Upcroft, B.; Abbeel, P.; Burgard, W.; Milford, M.; et al. The limits and potentials of deep learning for robotics. *Int. J. Robot. Res.* **2018**, *37*, 405–420. [CrossRef]
26. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef] [PubMed]
27. Cong, S.; Zhou, Y. A review of convolutional neural network architectures and their optimizations. *Artif. Intell. Rev.* **2023**, *56*, 1905–1969. [CrossRef]
28. Han, D.; Mulyana, B.; Stankovic, V.; Cheng, S. A Survey on Deep Reinforcement Learning Algorithms for Robotic Manipulation. *Sensors* **2023**, *23*, 3762. [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.