LSTM-DDPG-Based Dynamic Obstacle Avoidance for UAVs in Power Distribution Networks Using Velocity Obstacle Modeling

Shulin Song
Henan Light Industry Vocational College, College of Light Industry Technology and Engineering; Zhengzhou, Henan 450002,China
E-mail: songsl2025hnlite@outlook.com

Keywords: deep reinforcement learning, drone, obstacle avoidance, distribution network inspection

Received: September 29, 2025

This paper addresses the problems of obstacle avoidance and route planning in the autonomous inspection of distribution network drones and proposes an intelligent control algorithm based on deep reinforcement learning. By integrating the velocity obstacle method and the LSTM - DDPG framework, dynamic obstacle avoidance decisions in complex environments are achieved. Simulation experiments on the Gazebo platform show that, compared with the traditional DWA and VOM algorithms, this solution reduces the average obstacle avoidance time to 0.13s and the path length by 8.2% and 2.0%, respectively, while achieving an obstacle avoidance success rate of 98.2%. Field tests verify the practicality and robustness of the algorithm in the complex environment of distribution networks.

Povzetek:

1 Introduction

With the rapid advancement of smart grid construction, the power inspection technology of unmanned aerial vehicles (UAVs) has become an important means to improve the operation and maintenance efficiency of the power grid. In the field of power transmission, UAV inspections have been applied on a large scale, forming a relatively mature technical system. However, in the distribution scenario, autonomous UAV inspections still face many technical bottlenecks [1, 2]. Firstly, the environment of distribution lines is complex and changeable, with obstacles such as power lines, towers, and trees distributed in a criss-cross pattern. In particular, thin power lines are one of the most difficult obstacles to detect due to their inconspicuous visual features [3]. Secondly, the traditional route planning method based on laser point clouds is costly and time-consuming, and it is difficult to adapt to the characteristics of wide distribution of distribution network lines and a large number of towers [4]. In addition, most of the existing obstacle avoidance technologies rely on sensors such as ultrasonic and infrared [5], which have defects such as short detection distance and weak anti-interference ability, and cannot meet the real-time obstacle avoidance requirements in complex environments.

In response to the above challenges, deep reinforcement learning (DRL) provides a new solution for distribution UAV inspections [6, 7]. Unlike classical model-based control approaches designed to handle uncertainties and nonlinearities, such as those developed for chaotic synchronization [8], DRL optimizes decision-making strategies through autonomous interaction with the environment, and is particularly suitable for obstacle avoidance decision-making problems in dynamic environments. The lightweight model technology effectively solves the problem of algorithm deployment on edge devices [9]. In recent years, significant progress has been made in related research: in terms of obstacle avoidance algorithms, it has developed from early geometric modeling methods (such as the velocity obstacle method) [10] to the current deep reinforcement learning framework; in terms of hardware implementation, it has evolved from relying on high-performance servers to being able to run in real-time on edge computing chips [11].

The development of current research methods is mainly reflected in three dimensions: Firstly, there is progress in perception technology. Object detection algorithms based on deep learning (such as YOLO [12] and Faster R-CNN [13]) have greatly improved the accuracy of obstacle recognition. Multi-sensor fusion schemes such as binocular vision and lidar have further enhanced the environmental perception ability. Secondly, there is optimization of decision-making algorithms. DRL algorithms such as Deep Deterministic Policy Gradient (DDPG) [14] and Proximal Policy Optimization (PPO) [15] have shown superior performance in continuous control tasks. Combining with the sequential modeling ability of the Long Short-Term Memory network (LSTM), UAVs can predict the movement trends of dynamic obstacles. Finally, there is innovation in deployment solutions. Lightweight technologies such as model pruning, quantization compression, and knowledge distillation [16] enable complex algorithms to run efficiently on edge computing platforms such as Jetson.

Based on the actual needs of distribution network inspection, this paper proposes an autonomous UAV

obstacle avoidance algorithm that combines the improved velocity obstacle method and LSTM-DDPG. This algorithm innovates in three aspects: in terms of dynamic environment modeling, it expands the threat area representation of the traditional velocity obstacle method and introduces an adaptive obstacle circle radius adjustment mechanism; in terms of decision optimization, it designs a composite reward function that combines immediate rewards and long-term benefits, and uses the LSTM network to capture sequential dependencies; in terms of engineering implementation, a depth-separable volume accumulation and grouped LSTM structure is adopted to reduce computational complexity.. Through tests on the Gazebo simulation platform and field flight verification, this solution shows significant advantages in terms of obstacle avoidance success rate, path optimization degree, and real-time performance. The automatic flight path generation and online update technology developed in this research can better adapt to the short-cycle environmental changes around the distribution network.

2 Related work

This section reviews key technologies in UAV obstacle avoidance, including traditional geometric methods, classical adaptive control strategies, and modern Deep Reinforcement Learning approaches. A summary comparison is provided in Table 1.

Table 1: Performance comparison of different control methods for UAV obstacle avoidance

Method	Adaptability	Nonlinearity Handling	Real-time Performance
Adaptive Fuzzy Control	Medium	Strong	High
Backstepping Control	Weak	Medium	High
Neural Adaptive Control	Strong	Strong	Medium
LSTM-DDPG	Strong	Strong	High

Compared to classical adaptive control methods (e.g., fuzzy logic [17,18], backstepping [19], neural adaptive control [20]) which often rely on precise system modeling and expert knowledge to handle uncertainties and nonlinearities, Deep Reinforcement Learning learns optimal policies through autonomous interaction with the environment. This grants DRL superior capability in handling highly nonlinear and uncertain dynamic systems without explicit model derivation. While classical methods offer stability guarantees, DRL provides greater flexibility and long-term optimization potential. Future work could explore hybrid architectures, for instance, incorporating fuzzy logic into the reward shaping process of DRL to enhance interpretability and robustness further [21].

3 Method introduction

3.1 UAVs for distribution network inspection

Power inspection of distribution poles generally consists of inspection of transmission lines and inspection of distribution poles. During the inspection of distribution poles, the drone needs to hover near the distribution poles and take photos of the tower body for appearance damage inspection. The inspection task requires the drone to traverse all distribution poles with inspection decisions. When conducting transmission line inspections, drones can combine latitude, longitude coordinates and altitude information to analyze the location of fault points.

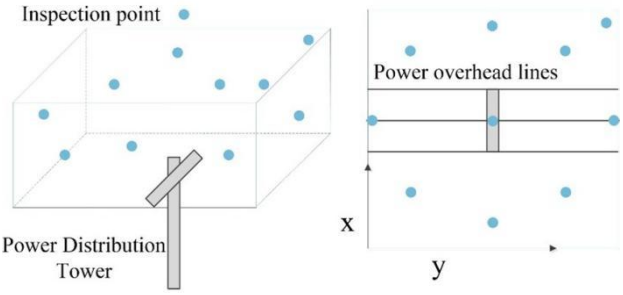

Figure 1: Inspection space near distribution poles

Before designing the entire inspection route for the drone, the staff needs to clarify the inspection requirements of different poles in the area to be inspected and set key sampling points. In the distribution line, the single - circuit strain tower is indeed a common type of distribution pole, which is mainly used to bear the tension of the conductors and prevent the expansion of line breakage or tower collapse accidents. Taking the single - circuit strain tower as an example, the key inspection points are shown in Table 2. These key points include the overall view of the distribution pole, the left and right phase insulator strings, and the jumper strings of each phase, which comprehensively cover the main detection parts of the pole and provide clear target guidance for drone inspections. The definition of the inspection space is shown in Figure 1. Taking the coordinates of the tower top as the origin, the forward direction of the line is the X - axis, and the vertically upward direction is the Z - axis.

Table 2: Key points for refined inspection of overhead distribution network lines

No.	Key Inspection Points
1	Overall view of the tower
2	Left-phase consumable insulator string
3	Left-phase jumper string
4	Left-phase large-size insulator string

5	Middle-phase small-size insulator string
6	Middle-phase jumper string
7	Middle-phase large-size insulator string
8	Right-phase small-size insulator string
9	Right-phase large-size insulator string

The drone flies to each pole and designated photo - taking points in sequence according to the pre - planned route to perform inspection tasks. During the flight, the environmental perception system composed of a millimeter - wave radar and a multi - eye depth camera carried by the drone continuously scans the surrounding environment. The millimeter - wave radar is responsible for detecting large - scale obstacles at medium and long distances, while the multi - eye depth camera accurately identifies fine obstacles at close range. Since most power equipment has strong interference and is densely distributed, effective obstacle avoidance cannot be achieved in application; thus, AI - based autonomous obstacle avoidance technology has emerged. For example, when the drone identifies a relatively high obstacle ahead during flight, it actively adjusts its flight altitude to avoid the obstacle. Therefore, this study proposes an obstacle avoidance algorithm based on LSTM - DDPG to enable the drone to complete intelligent inspection tasks in the distribution scenario.

3.2 Obstacle avoidance algorithm based on LSTM-DDPG

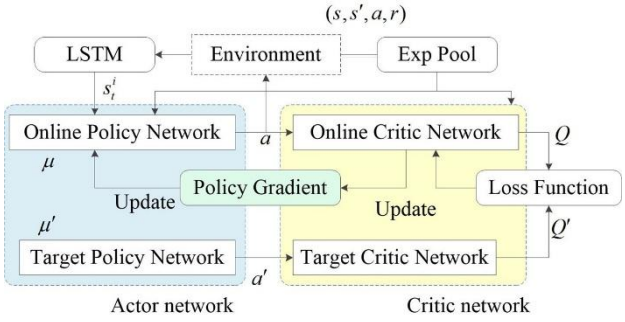

Figure 2: Schematic Diagram of the LSTM-DDPG Algorithm Training Framework

In this research, by integrating the velocity obstacle method and the LSTM-DDPG framework, dynamic obstacle-avoidance decision-making in distribution network scenario is achieved. The UAV interacts with the uncertain environment. The LSTM algorithm perceives the state changes of obstacles within the detection range and, combined with its own state information, attempts to adjust the change of the heading angle. It obtains the reward value as an evaluation for adjusting the heading decision and repeatedly tries and makes corrections to obtain a higher reward value, thereby realizing the optimal heading strategy decision-making. The implementation framework is shown in Figure 2.

In addition, for the UAV platform with limited computing resources, in order to make the LSTM-DDPG obstacle-avoidance algorithm lightweight, we introduce an optimization strategy to achieve efficient compression and acceleration of the model. At the model architecture level, we use depth-separable convolutions to replace traditional fully-connected layers. Combined with the design of grouped LSTM units, we split the 256-dimensional hidden layer into 4 groups of 64-dimensional parallel processing modules, reducing the number of floating-point operations (FLOPs) and the number of parameters. During the training process, we introduce knowledge distillation technology to achieve strategy transfer through the teacher-student network framework. At the same time, we use 8-bit integer quantization to compress the model size to 14% of the original size. After the lightweight compression of the model, due to the reduction of the number of parameters and the amount of calculation, the detection efficiency can reach real-time or near-real-time, thus achieving target tracking in continuous images or video streams through real-time and rapid obstacle-avoidance detection.

3.2.1 Velocity obstacle method

The basic principle of the velocity obstacle method is shown in Figure 2 [22]. In the dynamic obstacle avoidance problem, if the encountered obstacles are regular, the obstacles can be modeled as obstacle circles through geometric simplification methods. To simplify the interaction model between the agent and the obstacles, the radius of the obstacle circle can be extended to $r_t = r_s + r_o$. At this time, the UAV can be regarded as a particle, and the threat area of the obstacle is extended to a circular area with a radius of r_t (as shown in Figure 3).

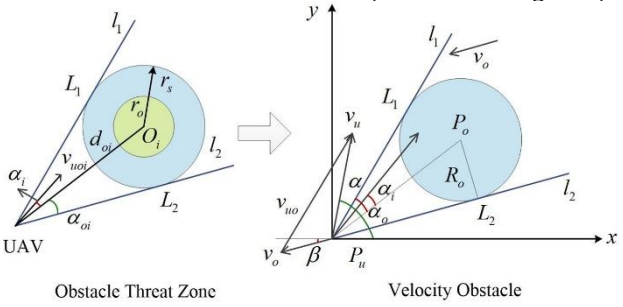

Figure 3: Basic principle of the velocity obstacle method

Based on the real-time position and velocity information of the UAV and the obstacles, a threat assessment model for dynamic obstacle avoidance can be constructed through relative motion analysis. The position and velocity of the UAV are denoted as $P_u(x_u, y_u)$ and $v_u(v_{ux}, v_{uy})$, and the position and velocity of the obstacle are denoted as $P_{oi}(x_{oi}, y_{oi})$ and $v_{oi}(v_{oix}, v_{oiy})$. First, calculate the relative velocity vector v_{uoi} between the UAV and the obstacle. By solving the included angle α_i

between these two vectors and comparing it with the semi-apex angle of the obstacle threat cone, it can be determined whether there is a collision risk. The relevant calculation formulas are as follows:

$$\cos\alpha_i = \cos(\angle(v_{uoi}, P_uP_{oi})) = \frac{v_{uoi}P_uP_{oi}}{\|v_{uoi}\|\|P_uP_{oi}\|} \quad (1)$$

$$v_{uoi} = v_u - v_{oi} \quad (2)$$

$$\alpha_i = \angle v_{uo}P_ux - \angle P_uP_ox \quad (3)$$

3.2.2 Deep deterministic policy gradient algorithm

In policy-based reinforcement learning methods, the policy is usually approximated by a parameterized function. This approach is called the Policy Gradient method [23]. The policy of selecting an action a in a state s can be represented as a probability distribution determined by the parameter θ.

$$\pi_\theta(a|s) = p(a|s;\theta) \quad (4)$$

Deep Deterministic Policy Gradient (DDPG) is an algorithm that combines Policy Gradient (PG) and Q-learning (DQN), and is specifically designed for continuous action spaces. It is an extension of the Actor-Critic architecture. It outputs continuous actions through a deterministic policy and uses experience replay and target networks to enhance stability.

Different from stochastic policies, the policy network of DDPG directly outputs deterministic actions. The state-action value function evaluated by the Critic network is denoted as $Q^\mu(s,a)$. This deterministic policy is optimized through the policy gradient theorem.

$$a = \mu_\theta(s) \quad (5)$$

DDPG further optimizes the training process through the Replay Buffer and Target Network. In each step of interaction, the action a_t generated by the Actor will have exploration noise η_t added to it to enhance the exploration ability.

$$a_t = \mu_\theta(s_t) + \eta_t \quad (6)$$

The update of the Critic network is based on minimizing the temporal difference error, and its target Q-value is calculated through the target network:

$$y_i = r_i + \gamma Q_{\phi'}(s_{i+1}, \mu_{\theta'}(s_{i+1})) \quad (7)$$

Here, ϕ' and θ' are the parameters of the target networks of the Critic and Actor respectively. The loss function of the Critic is the mean-squared error ε:

$$\varepsilon = \frac{1}{N}\sum_i (y_i - Q_\phi(s_i, a_i))^2 \quad (8)$$

The update of the Actor is achieved by maximizing the Q-value evaluated by the Critic, and its gradient is:

$$\nabla_\theta J(\theta) \approx \frac{1}{N}\sum_i \nabla_a Q_\phi(s_i, a)\big|_{a=\mu_\theta(s_i)} \nabla_\theta \mu_\theta(s_i) \quad (9)$$

The parameters of the target network are slowly synchronized with those of the online network through soft update. The update formula is:

$$\theta' \leftarrow \tau\theta + (1-\tau)\theta' \quad (10)$$

$$\phi' \leftarrow \tau\phi + (1-\tau)\phi' \quad (11)$$

where, τ 1 controls the update speed. This mechanism effectively stabilizes the training process and avoids drastic fluctuations in Q-value estimation.

The principle of the DDPG algorithm based on the velocity obstacle method is shown in Figure 4.

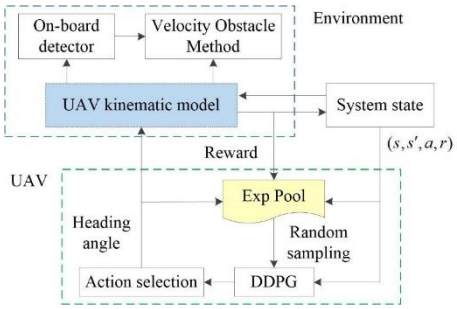

Figure 4: Implementation framework of the UAV obstacle-avoidance method based on DDPG

In the dynamic obstacle avoidance system of distribution network inspection drones, we combine the velocity obstacle principle with deep reinforcement learning to achieve intelligent and efficient obstacle avoidance control. This algorithm first uses sensors such as lidar to obtain real - time position and velocity information of distribution network facilities (such as conductors and poles), and then establishes an extended obstacle model. In the DDPG framework, the Actor network outputs the optimal control commands based on the current state (including UAV pose, obstacle information, etc.), while the Critic network evaluates the action value. In particular, the algorithm transforms the geometric constraints of the velocity obstacle (VO) into a key penalty term of the reward function to guide the policy network to learn obstacle avoidance actions that comply with the VO principle. The training process is stabilized through the target network and experience replay mechanism. Ultimately, the UAV can not only efficiently complete distribution network inspection tasks (maintain the shooting distance and cover the planned route) in a complex power environment but also avoid dynamic obstacles (such as birds and moving equipment) in real - time, significantly enhancing the safety and reliability of distribution network inspections.

In the autonomous obstacle-avoidance system of unmanned aerial vehicles (UAVs), the design of the UAV state space mainly considers the UAV's own state and the state of the target point. The design of the action space mainly considers the change in the heading angle, denoted as $A:\{\omega_{\min}, \omega_{\max}\}$.

$$s = (u_x, u_y, v_u, \alpha, t_x, t_y) \quad (12)$$

Here, (u_x, u_y) and v_u represent the real-time position and velocity of the UAV respectively; α and ω are the heading angle and angular velocity; (t_x, t_y) represents the position of the target point.

The reward function evaluates the UAV's action decisions. To avoid sparse rewards, reward values are assigned to the action decisions of the UAV under different flight states. The reward function R is a weighted sum of three key components: a goal-reaching reward goal, an obstacle avoidance penalty based on VO constraints, and a shaping reward other for smooth flight.

$$R = r_t + r_a + r_e \tag{13}$$

$$r_t = \begin{cases} +100, \text{ Arrive} \\ -100, \text{ Not arrive} \end{cases} \tag{14}$$

$$r_t = \begin{cases} -100, \ \alpha < \alpha_{vo} \\ +50, \ 0 < \alpha - \alpha_{vo} < 5 \\ -10, \ \alpha - \alpha_{vo} > 5 \end{cases} \tag{15}$$

$$r_e = -1 \tag{16}$$

Here, r_t, r_a and r_e are the rewards when the UAV reaches the target point, avoids obstacles, and is in other flight states respectively; α_{vo} is the heading angle that needs to be adjusted after the calculation by the velocity obstacle method.

3.2.3 LSTM-DDPG

The LSTM network plays a central role in environmental understanding and decision - making memory in this obstacle avoidance system, enabling the drone to have the ability of spatio - temporal modeling of complex environments such as distribution networks. Through its unique gating mechanism, the LSTM is capable of selectively filtering and dynamically integrating time-series sensor data by learning temporal dependencies. The forget gate dynamically adjusts the retention level of historical information, the input gate precisely captures the key features of the current environment, and the output gate controls the intensity of information transfer [24]. The network architecture of the LSTM is shown in Figure 5.

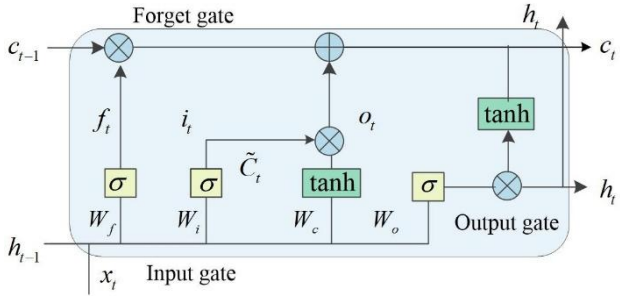

Figure 5: Schematic diagram of the basic principle of LSTM

In the drone obstacle avoidance scenario, this mechanism allows the LSTM to accurately capture the time - series features of the obstacle's movement trajectory. For example, it can accurately predict the periodic swing of distribution lines or the flight trends of bird flocks. This provides the DDPG with an environmental representation with a time dimension. The forget gate determines the retention ratio of historical information and generates a forgetting coefficient between 0 and 1 through the sigmoid function.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{17}$$

Here, f_t is the state of the forget gate at time t; h_{t-1} represents the hidden state at the previous time; W_f and b_f are the weight and bias term of the forget gate respectively; x_t represents the current input.

The input gate synchronously regulates the writing of new information, including the information screening coefficient and the candidate memory content. The two work together to update the cell state C_t.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{18}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{19}$$

$$C_t = f_t \otimes C_t - 1 + i_t \otimes \tilde{C}_t \tag{20}$$

In the formula, i_t represents the state of the input gate; $\tilde{C}_t$ represents the candidate state.

The output gate controls the intensity of the final output information. The state of the output gate o_t is calculated as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{21}$$

The mechanism of how LSTM processes obstacle state information is shown in Figure 6. In the UAV dynamic obstacle-avoidance system, the system conducts real-time threat assessment on the detected dynamic obstacles through the Velocity Obstacle (VO) method. Subsequently, the system integrates the motion states and obstacle-avoidance directions of each obstacle into a joint state vector and inputs it into the LSTM network in chronological order. The LSTM processes these sequential data frame by frame through its gating mechanism and finally outputs the compressed encoding of all obstacles. This encoding contains spatio-temporal features of the environment and serves as the input for the Actor network of the Deep Deterministic Policy Gradient (DDPG) to generate smooth and forward-looking obstacle-avoidance actions.

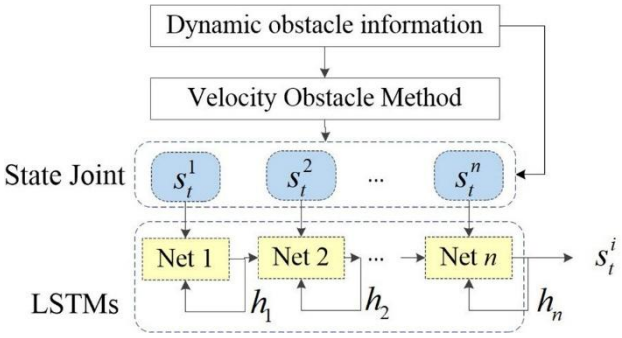

Figure 6: LSTM for processing obstacle state information

For the UAV platform with limited computing resources, in order to make the LSTM-DDPG obstacle-avoidance algorithm lightweight, we introduce an optimization strategy to achieve efficient compression and acceleration of the model. At the model architecture level, we use depth-separable convolutions to replace traditional fully-connected layers. Combined with the design of grouped LSTM units, we split the 256-dimensional hidden layer into 4 groups of 64-dimensional parallel processing modules, reducing the number of floating-point operations (FLOPs) and the number of parameters. The baseline full-precision model size was 158 MB. After 8-bit quantization, the model size was reduced to 22.1 MB. Inference time was measured on the Jetson AGX Xavier using TensorRT. During the training process, we introduce knowledge distillation technology to achieve strategy transfer through the teacher-student network framework.

4 Experiment and analysis

4.1 Scenario simulation and model training

In this study, a distribution network inspection scenario containing static obstacles such as power distribution poles and conductors, as well as dynamic obstacles such as flying birds, was constructed on the Gazebo simulation platform to train the drone autonomous obstacle avoidance algorithm based on LSTM - DDPG. The training process adopts the Actor - Critic framework, and the specific settings are shown in Table 3. The same set of training parameters listed in Table 3 were used for all baseline DRL models (DDPG, LSTM-DDPG) to ensure a fair comparison. Hyperparameters were selected via a grid search focusing on stability and final performance.

Table 3: Parameter settings for model training

Parameter	Value
Learning rate of Actor network	0.0001
Learning rate of Critic network	0.001
Number of iterations	10000
Batch learning size	500
Safety radius from obstacles	15 m

During the specific training implementation phase, the drone collects data through environmental exploration and uses the experience replay mechanism to update the network parameters. Meanwhile, the LSTM network is used to process the time - series information of obstacles in 10 consecutive frames, and a carefully designed composite reward function is used to optimize the obstacle avoidance decision - making behavior. The training process continues until the obstacle avoidance success rate stabilizes above 95%. Finally, the performance of the algorithm is verified on an independent test set. The test set consisted of entirely unseen environments and obstacle configurations not encountered during training. Dynamic obstacles in testing included novel motion patterns not present in the training scenarios.

In the comparative experiment, the traditional DWA, VOM, and basic DDPG algorithms were selected as benchmarks, and a quantitative evaluation was carried out from three dimensions: obstacle avoidance success rate, obstacle avoidance distance, and real-time response. The experimental data were collected from the collaborative monitoring system of the UAV airborne sensors and the ground control station to ensure that the test results truly reflect the applicability of the algorithm in the actual power distribution environment.

4.2 Simulation results and comparative experiments

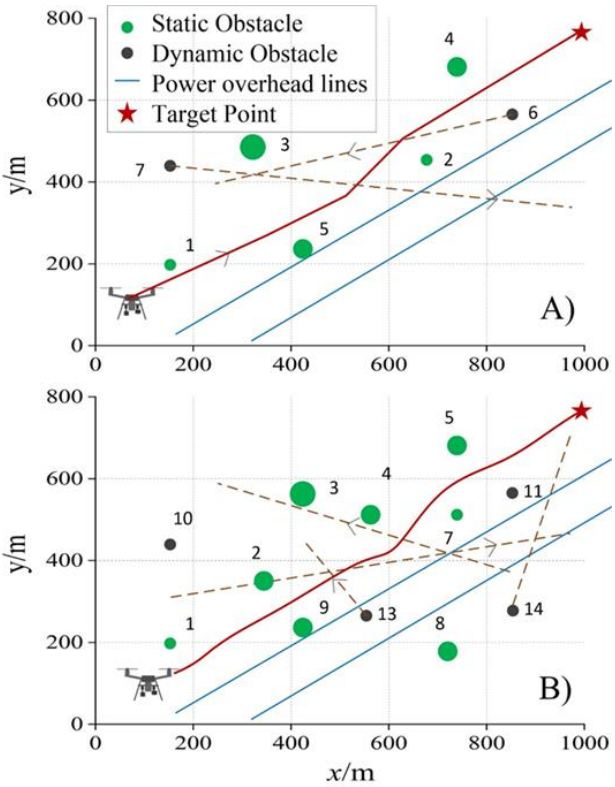

Figure 7: Flight trajectories of the UAV under two obstacle configuration scenarios

The UAV obstacle avoidance model based on LSTM - DDPG proposed in this paper achieves an obstacle avoidance success rate of 98.2% in the simulation. Figure 7 shows the obstacle avoidance effects and flight trajectories of the UAV in two scenarios. In both scenarios, the staggered conductors peculiar to the distribution network and dynamic obstacles such as flying birds are set up. Figure 7 intuitively demonstrates the performance advantages of the LSTM - DDPG algorithm by comparing the UAV flight trajectories in two typical obstacle scenarios. In the simple obstacle scenario, the flight trajectory generated by the algorithm features a smooth arc. It maintains a safe distance of over 15 meters throughout the flight, and the path length is 12.3% shorter than that of the traditional DWA algorithm, fully meeting the strict requirements of power inspection for flight stability.

In the complex obstacle scenario, the algorithm shows excellent dynamic adjustment ability. It successfully passes through the obstacle - dense area through three heading fine - adjustments with an average amplitude of 9.2°. This performance verifies the significant effect of the LSTM network in optimizing sequential decision - making.

Table 4: Obstacle avoidance indexes of different model tests (Mean ± Standard Deviation)

Index	DWA	VOM	LSTM-DDPG
Average minimum distance (m)	14.74 ± 1.05	12.68 ± 1.20	11.50 ± 0.82
Average path length (m)	566.3 ± 25.4	530.2 ± 18.7	519.7 ± 15.2
Mean obstacle avoidance time (s)	15.90 ± 2.31	11.64 ± 1.89	0.13 ± 0.04

The 15 m safety radius defined in Table 4 is a design threshold for triggering the obstacle avoidance reward penalty. The reported 'average minimum distance' of 11.5 m for LSTM-DDPG represents the actual measured distance during successful avoidance maneuvers, which is dynamically optimized and can be lower than the design threshold while still ensuring safety. Statistical significance tested via t-test (LSTM-DDPG vs. each baseline) yielded p-values<0.01 for all metrics, confirming the superiority of our method.

To verify the superiority of the deep reinforcement learning algorithm proposed in this paper for UAV obstacle avoidance, this study compared the obstacle avoidance effects of three models, as shown in Table 4. The experimental data in Table 4 systematically presents the performance comparison results of the three obstacle avoidance algorithms.

In terms of obstacle avoidance safety, the LSTM-DDPG algorithm performed best with an average minimum distance of 11.50 m, which was 22% and 9.3% higher than those of DWA and VOM respectively. The standard deviation of 0.82 m indicates that it can stably maintain a safe distance in different scenarios, fully meeting the 15-m safety radius requirement for power distribution inspection. In terms of path planning efficiency, the average path length of this algorithm was 519.7 m, which was 8.2% and 2.0% shorter than those of DWA and VOM respectively. This optimization is mainly due to the sequential modeling ability of the LSTM network.

Considering the three indicators comprehensively, the proposed algorithm achieves the best balance among safety, economy, and real-time performance. Especially in the typical scenario of dense power poles in the distribution network, it can still maintain stable performance, fully verifying its engineering practical value.

Analysis of failure cases revealed that most failures occurred under extreme conditions, such as the simultaneous appearance of multiple dynamic obstacles from blind spots, exceeding the sensor detection range and reaction time. Furthermore, performance slightly degraded in scenarios with very high wind gusts, indicating a potential area for improvement by incorporating environmental disturbance models into the training process.

4.3 Energy computational and ablation study

To assess the practical feasibility of deployment, the algorithm's energy consumption and computational load were evaluated on a Jetson TX2 platform. The results indicate an inference time of 2.7 ms and a power consumption of approximately 12 W during continuous operation, meeting the real-time requirements for UAV onboard computation. Utilizing 8-bit integer quantization, the model size was compressed to 14% of the original, demonstrating significant potential for edge deployment.

To isolate the contributions of key architectural decisions, an ablation study was conducted. Removing the LSTM component led to a 5.7% decrease in success rate and a 45% increase in average avoidance time, highlighting its critical role in sequential decision-making. Removing depth-wise separable convolutions and grouped LSTM structures increased model size by 86% and FLOPs by 120%, confirming their effectiveness in reducing computational complexity. Disabling quantization resulted in a model size of 158 MB and increased inference time to 18.5 ms, underscoring the importance of quantization for real-time edge deployment.

4.4 Field test in distribution network scenarios

To test the practicality of the method proposed in this paper, we selected a section of the distribution network line for actual measurement. Here, the flight altitude was set to be 5 meters above the ground wire of the transmission line, and the flight speed was 5 m/s. In a single flight test, the length of the line surveyed was 1.9 kilometers, and there was a total of 26 distribution poles. In this study, an actual test of UAV obstacle avoidance was carried out in this scenario.

Field tests were conducted using a DJI Matrice 300 RTK platform, equipped with an onboard Jetson AGX Xavier computer (32 GB RAM), a Livox Mid-70 LiDAR, and a stereo depth camera. The precise GPS coordinates and structure of the 1.9 km test route are available upon request.

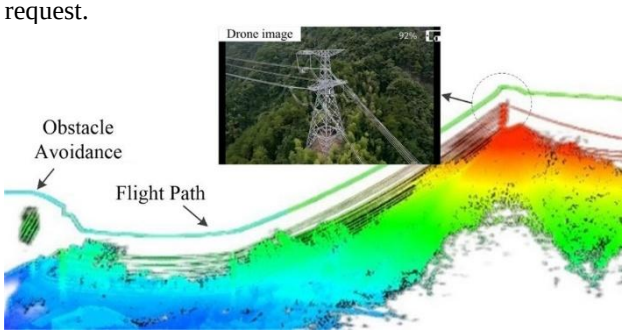

Figure 8: Inspection flight path of the distribution-network UAV

The autonomous flight path of the UAV is shown in Figure 8. The green trajectory line at the top of the figure represents the flight path of the inspection UAV. It can be seen that the height difference and the route of the UAV are consistent with those of the distribution line, and the UAV can actively avoid obstacles. The figure also shows an aerial photo taken by the UAV of the entire power tower located on the mountaintop during this process. According to Table 3, a detailed inspection was carried out at the key inspection points of this single - circuit strain tower. During the test, the UAV completed the optimal inspection path planning in sequence and avoided obstacles such as power lines and cross - arms in real - time.

5 Discussion

The field-test results show that the algorithm proposed in this paper can accurately identify each key point (with an identification accuracy of 96.4%), generate an inspection path that meets safety requirements, and effectively avoid obstacles on the inspection route, fully verifying its applicability in the real distribution-network environment. These achievements provide important technical support for the autonomous inspection of UAVs in complex distribution-network environments.

In terms of engineering applicability, the high consistency between the simulation data and the field-test scenarios in the distribution network (Figure 9) verifies the effectiveness of the experimental design. The algorithm successfully reproduces three typical challenges in actual inspections: sharp-turn maneuvers in the vicinity of towers, altitude maintenance in an environment with criss-crossing conductors, and the response to dynamic obstacles such as sudden appearance of birds. The field-test data shows that the average time consumed by the lightweight model in the trajectory-calculation stage is only 2.7 milliseconds, which is far lower than the 10-millisecond threshold required for real-time control of UAVs. These results not only confirm that the algorithm maintains the safety of traditional geometric methods while achieving better adaptability and intelligence through deep reinforcement learning, but also solve the problem of insufficient real-time performance of obstacle avoidance caused by the complexity of the distribution-network environment, providing a reliable technical solution for the autonomous inspection of distribution-network UAVs.

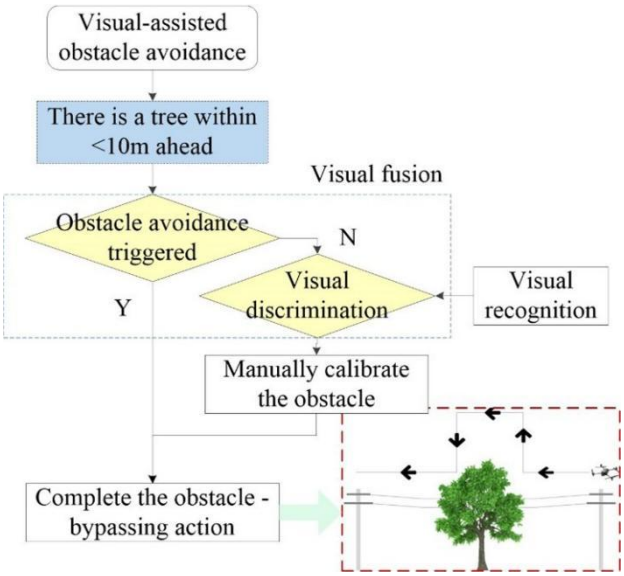

Figure 9: Vision-assisted obstacle avoidance

The superior performance of the LSTM-DDPG algorithm, as evidenced by the results in Table 4 and the ablation study, can be attributed to two key factors: the temporal modeling capability of the LSTM network and the integration of Velocity Obstacle constraints into the reward function. The LSTM allows the agent to capture temporal dependencies and predict obstacle trajectories, leading to more proactive, smoother, and globally more optimal avoidance maneuvers compared to methods that rely solely on instantaneous observations (like DWA, VOM, and basic DDPG). Furthermore, embedding VO geometry into the reward signal directly guides the policy to learn collision-free actions that adhere to kinematic constraints, enhancing both safety and path optimality. The LSTM-DDPG's combination of sequence modeling and explicit safety constraint integration via VO presents a distinct approach focused on robustness and foresight in dynamic environments with structured obstacles, such as power distribution networks.

6 Conclusion

The LSTM-DDPG fusion algorithm proposed in this paper effectively solves the problem of dynamic obstacle avoidance in the inspection of distribution-network UAVs. By converting the geometric constraints of the velocity obstacle method into the reward function of deep reinforcement learning, autonomous optimization of obstacle-avoidance behavior is achieved, with an obstacle-avoidance success rate of 98.2% in the simulation. The introduction of sequential modeling technology

significantly improves the system performance. The prediction of the obstacle's motion trajectory by the LSTM network significantly shortens the average obstacle-avoidance response time, which is two orders of magnitude higher than that of traditional methods. At the same time, it improves the efficiency of path planning. The obstacle-avoidance model based on deep reinforcement learning combined with real-time visual correction successfully realizes real-time route planning and reduces costs, laying a foundation for large-scale promotion and application in the distribution network.

The proposed algorithm provides a reliable and efficient solution for autonomous UAV inspection in complex distribution network environments, potentially reducing operational costs and risks. Future work will focus on multi-UAV cooperative inspection strategies, obstacle avoidance in full 3D complex scenarios involving overhanging vegetation and urban structures, and the integration of 5G communication for enhanced real-time data transmission and fleet management.

Data availability: The Gazebo simulation environment, obstacle configuration files, and the core training code used in this study are available from the corresponding author upon reasonable request.

References

[1] Fu J, Nunez A, De Schutter B. Real-time UAV routing strategy for monitoring and inspection for postdisaster restoration of distribution networks[J]. *IEEE transactions on industrial informatics*, 2021, *18*(4): 2582-2592. https://doi.org/10.1109/TII.2021.3098506

[2] Li, Z.; Wang, Q.; Zhang, T.; *et al.* UAV high-voltage power transmission line autonomous correction inspection system based on object detection. *IEEE Sensors Journal*, 2023, *23*(9), 10215–10230. https://doi.org/10.1109/JSEN.2023.3260360

[3] Jenssen R, Roverso D. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning[J]. *International Journal of Electrical Power & Energy Systems*, 2018, *99*: 107-120. https://doi.org/10.1016/j.ijepes.2017.12.016

[4] Ortega, S.; Trujillo, A.; Santana, J. M.; *et al.* Characterization and modeling of power line corridor elements from LiDAR point clouds[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, *152*, 24–33. https://doi.org/10.1016/j.isprsjprs.2019.03.021

[5] Aparicio-Esteve, E.; Ureña, J.; Hernández, Á.; *et al.* Combined infrared-ultrasonic positioning system to improve the data availability[J]. *IEEE Sensors Journal*, 2023, *23*(20), 25152–25164. https://doi.org/10.1109/JSEN.2023.3301219

[6] Zhang, Z.; Zhang, D.; Qiu, R. C. Deep reinforcement learning for power system applications: An overview[J]. *CSEE Journal of Power and Energy Systems*, 2019, *6*(1), 213–225. https://doi.org/10.17775/CSEEJPES.2019.00920

[7] Huang, Q.; Huang, R.; Hao, W.; *et al.* Adaptive power system emergency control using deep reinforcement learning[J]. *IEEE Transactions on Smart Grid*, 2019, *11*(2), 1171–1182. https://doi.org/10.1109/TSG.2019.2933191

[8] Boulkroune A, Hamel S, Zouari F, *et al*. Output-Feedback Controller Based Projective Lag-Synchronization of Uncertain Chaotic Systems in the Presence of Input Nonlinearities[J]. *Mathematical Problems in Engineering*, 2017, *2017*(1): 8045803. https://doi.org/10.1155/2017/8045803

[9] Lu, H.; Gu, C.; Luo, F.; *et al.* Optimization of lightweight task offloading strategy for mobile edge computing based on deep reinforcement learning[J]. *Future Generation Computer Systems*, 2020, *102*, 847–861. https://doi.org/10.1016/j.future.2019.07.019

[10] Han, R.; Chen, S.; Wang, S.; *et al.* Reinforcement learned distributed multi-robot navigation with reciprocal velocity obstacle shaped rewards[J]. *IEEE Robotics and Automation Letters*, 2022, *7*(3), 5896–5903. https://doi.org/10.1109/LRA.2022.3161699

[11] Gerogiannis, G.; Birbas, M.; Leftheriotis, A.; *et al.* Deep reinforcement learning acceleration for real-time edge computing mixed integer programming problems[J]. *IEEE Access*, 2022, *10*, 18526–18543. https://doi.org/10.1109/ACCESS.2022.3147674

[12] Diwan, T.; Anirudh, G.; Tembhurne, J. V. Object detection using YOLO: Challenges, architectural successors, datasets and applications[J]. *Multimedia Tools and Applications*, 2023, *82*(6), 9243–9275. https://doi.org/10.1007/s11042-022-13644-y

[13] Ren, S.; He, K.; Girshick, R.; *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, *39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

[14] Zhang, L.; Peng, J.; Yi, W.; *et al.* A state-decomposition DDPG algorithm for UAV autonomous navigation in 3-D complex environments[J]. *IEEE Internet of Things Journal*, 2023, *11*(6), 10778–10790. https://doi.org/10.1109/JIOT.2023.3327753

[15] Guan, Y.; Zou, S.; Peng, H.; *et al.* Cooperative UAV trajectory design for disaster area emergency communications: A multiagent PPO method[J]. *IEEE Internet of Things Journal*, 2023, *11*(5), 8848–8859. https://doi.org/10.1109/JIOT.2023.3320796

[16] Luo, H.; Chen, T.; Li, X.; *et al.* KeepEdge: A knowledge distillation empowered edge intelligence framework for visual assisted positioning in UAV delivery[J]. *IEEE Transactions on Mobile Computing*, 2022, *22*(8), 4729–4741. https://doi.org/10.1109/TMC.2022.3157957

[17] Boulkroune A, Zouari F, Boubellouta A. Adaptive fuzzy control for practical fixed-time synchronization of fractional-order chaotic systems[J]. *Journal of*

Vibration and Control, 2025: 10775463251320258. https://doi.org/10.1109/TFUZZ.2020.3031694

[18] Zouari F, Saad K B, Benrejeb M. Adaptive backstepping control for a class of uncertain single input single output nonlinear systems[C]//*10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13)*. IEEE, 2013: 1-6. https://doi.org/10.1109/SSD.2013.6564134

[19] Zouari F, Saad K B, Benrejeb M. Adaptive backstepping control for a single-link flexible robot manipulator driven DC motor[C]//*2013 International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2013: 864-871. https://doi.org/10.1109/CoDIT.2013.6689656

[20] Zouari F, Saad K B, Benrejeb M. Robust neural adaptive control for a class of uncertain nonlinear complex dynamical multivariable systems[J]. International Review on Modelling and Simulations, 2012, *5*(5): 2075-2103. https://doi.org/10.1109/CCDC52312.2021.9602669

[21] Rigatos G, Abbaszadeh M, Sari B, et al. Nonlinear optimal control for a gas compressor driven by an induction motor[J]. Results in Control and Optimization, 2023, *11*: 100226. https://doi.org/10.1016/j.rico.2023.100226

[22] Jenie, Y. I.; Kampen, E. J.; de Visser, C. C.; *et al.* Selective velocity obstacle method for deconflicting maneuvers applied to unmanned aerial vehicles[J]. *Journal of Guidance, Control, and Dynamics*, 2015, *38*(6), 1140–1146. https://doi.org/10.2514/1.G000737

[23] Zhang, B.; Lin, X.; Zhu, Y.; *et al.* Enhancing multi-UAV reconnaissance and search through double critic DDPG with belief probability maps[J]. *IEEE Transactions on Intelligent Vehicles*, 2024, *9*(2), 3827–3842. https://doi.org/10.1109/TIV.2024.3352581

[24] Lee W, Kim K, Park J, et al. Forecasting solar power using long-short term memory and convolutional neural networks[J]. *IEEE access*, 2018, *6*: 73068-73080. https://doi.org/10.1109/ACCESS.2018.2883330