A Robust UAV Control System Based on Kalman Filter, MLP and DQN

Zhenxiao Jiang, *Tongji University,* Yuxuan Du, *Tongji University*

***Abstract*—Autonomous control systems have been integrated into everyone's life, and UAVs are a popular application area for them. With the continuous development of artificial intelligence and neural networks, the research on UAV control systems in complex environments with noise has been increasing dramatically, to This paper provides an in-depth discussion on the autonomous path planning problem of UAVs in complex environments, and proposes an innovative robust control system for the challenge of system robustness in dynamic and unpredictable environments. The system integrates a Kalman filter and a multilayer perceptron architecture, aiming to improve the stability and safety of UAVs in the face of significant measurement disturbances and potential sensor failures by optimising the data processing and decision making mechanisms. Experimental results show that the system exhibits significantly superior robustness and control performance in a predefined experimental environment by comparing it with the conventional deep Q-network approach and the traditional hidden Markov model. The research in this paper not only provides new ideas and methods for autonomous control of UAVs, but also provides important references and lessons for future development in the field of autonomous driving and intelligent control. Looking ahead, we plan to further explore the scalability and applicability of the system to assess its potential and value in a wider range of control problems.**

***Index Terms*—robust control system, artificial intelligence, machine learning, UAV, Kalman filter, MLP, DQN, DDQN.**

I. INTRODUCTION

AUTONOMOUS control systems have been used in a wide range of fields, including but not limited to manufacturing [1], [2], [3], logistics and transport [4] to energy management [5]. The ability of these systems to sense, analyse and control system operations without human intervention is one of the cornerstones of modern technological advancement. They are profoundly changing the field of unmanned operations, especially in the context of Unmanned Aerial Vehicles (UAVs) [6]. These systems have been seamlessly integrated into every aspect of our lives, from agriculture [7] to disaster response [8], demonstrating efficiency and safety beyond human capabilities. Traditional automated control systems are based on mathematics, physics and engineering and are designed to achieve precise control of system behaviour. Their core lies in describing the system dynamics through mathematical models and physical principles, as well as controller design and optimisation through engineering methods [9], [10]. The rapid development of artificial intelligence in recent years has added wings to these systems [11], and its application in the field of autonomous driving has brought the development of this field to a new stage.

Autonomous control systems rely on a holistic framework of AI, where deep learning provides [12], [13] powerful data processing and decision support capabilities to help the system perform more effectively in complex environments [11], [14], [15], [16], [17], [18]. AI is gradually becoming dominant in driving control [19], [20], [21], path tracking control [22], [23], path planning [20], [24], and multi-intelligence collaboration [25]. However, behind this remarkable progress, a key question remains: how to ensure the reliability and robustness of these systems and AI models in dynamic and unpredictable environments?

In the complex environments in which UAVs operate, the seamless integration of autonomous control systems is critical to achieving precise and safe missions. Imagine a drone tasked with travelling through an intricate cityscape to reach a target location while avoiding obstacles such as tall buildings and intricate power lines. In this challenging scenario, the drone's reliance on its sensors and control systems becomes critical. However, real-world operations are often fraught with challenges, including significant measurement disturbances caused by electromagnetic interference and variable weather conditions, which can severely compromise the integrity of sensor data, and random noise that can significantly degrade the accuracy of path-planning and obstacle-avoidance algorithms. More seriously, severe signal interference can lead to a complete breakdown of sensors, cameras, positioning systems, etc., with fatal effects on autonomous controllers. Our research is dedicated to tackling these challenges head-on by building control frameworks that incorporate robust optimisation techniques.

For traditional automatic control systems, increasing the complexity of the system with the introduction of new mechanisms is a good way to enhance the robustness of the system. In [26], a perturbation observer is proposed to achieve perturbation decay in a specified time and an auxiliary dynamics system with time-varying gain is constructed to handle the input saturation phenomenon to form a robust tracking control scheme for AUVs. In order to improve the robustness of the path tracking control system of an AUV, [27] proposed a robust model predictive control system designed by solving a cluster point model using LMI. [28] established a networked path-tracking control method including feedback linearisation, event-triggered controller, and safe sliding controller to the path-tracking control problem of self-driving vehicles under sensor and actuator attacks. In order to achieve path tracking without lateral velocity information, [29] used a robust H1 static output feedback controller based on a hybrid genetic algorithm and linear matrix inequalities with good results.

This strategy also has good results for those automatic control systems that incorporate artificial intelligence and deep learning. Further, the model of the neural network can be modified by applying, for example, an event-triggered mechanism to enhance the robustness of the model. [30], [31] studied event-triggered deep reinforcement learning based on event triggering and proposed an event-triggered deep Q-network (ETDQN) suitable for autonomous driving decision making. In order to solve directional planning problems such as turning at rational intersections, [32] established a learning autonomous driving motion planning method based on conditional deep Q networks and fuzzy logic. In order to solve the problem of illegal eavesdropping and interference in autonomous vehicle networks, [33] proposed a solution based on distributed Kalman filtering and deep reinforcement learning techniques. In addition to the above strategies, adopting more excellent and appropriate algorithms or strategies on the underlying algorithms can also enhance the robustness of the model. For the problem of autonomous flight of a tilt-rotor aircraft in a complex and dense environment, [34] proposed a motion planning algorithm using the principle of minimum energy, the A* algorithm and an optimised cost function. In order to solve the problem that traditional reinforcement learning and DQN cannot handle the continuous action space, [35] used the DDPG algorithm for optimisation to output continuous and smooth control values through the policy gradient method, which improves the accuracy of autonomous control of UAVs. In order to enable robust UAV flight in dynamic uncertain environments, [36] created a new deep reinforcement learning method, Robust-DDPG, which introduces three key learning techniques: delayed learning technique, adversarial attack technique and hybrid exploration technique.

DQN was applied to train UAVs to perform autonomous control tasks, enabling UAVs to plan safe flight paths, avoid obstacles and reach their destinations in complex environments, and achieved good results in tests [37], [38], [39], [40]. We plan to use a dual Deep Q-Network (DDQN), a reinforcement learning algorithm that combines deep learning and Q-learning, to solve the pathfinding problem and complete the assigned tasks in complex road conditions [41]. At the same time, we expect to ensure that the control system operates based on reliable information by constructing an estimation and filtering mechanism for different types of disturbances, which can effectively deal with noisy and possibly corrupted sensor data. For the specific problem of this paper, we constructed a set of filters based on mathematical methods and neural network models, including a Kalman filter (KF) [42] and a fault detector. In this paper, we filter the information received by the sensor through the filter consisting of the KF to obtain more accurate information, and use the fault detector and DQN to introduce the approximate severity of the interference and make optimal decisions accordingly.

Finally, we will validate our approach through fine-grained simulations and rigorous experiments, scrutinise how random noise and sensor faults affect decision-making separately and jointly, and evaluate the system's ability to reduce interference and effectively detect potential faults.

II. Problem Formulation

This paper aims to improve the robustness of autonomous UAV control under significant interference by proposing a new systematic framework based on machine learning.

A. Task Description: Autonomous Robust Navigation under Significant Interference

To evaluate and visualize the performance of the new system under extreme conditions, we have deliberately constructed a virtual environment simulating post-disaster UAV rescue operations.

The visualization of the environment is shown in Fig. 1.

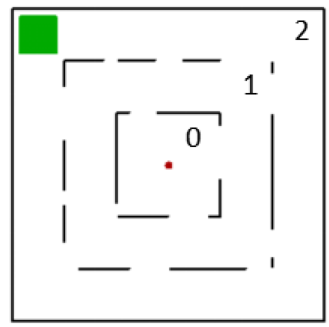

Fig. 1. Virtual Environment Visualization

At the center of this virtual environment lies a single-story building, which consists of two perimeter walls (represented by black lines in Fig. 1), dividing the entire area into three distinct zones labeled Zone 0, Zone 1, and Zone 2. These zones are interconnected by gates of different sizes on the walls. An exit is located in Zone 2 (marked by the green area in the Fig. 1).

A UAV agent (represented by the red point in the figure) initially positioned within the inner area (a random location within Zone 0) must, at each time step, choose to move in one of the four cardinal directions (north, south, east, or west) or to permanently halt its movement and waiting for further rescue operations. The velocity of the agent suffers a tiny random interference at each move instead of being fixed, making it impossible for any system to precisely predict the position after any move.

At each time step, the agent receives positional information from GPS, including the precise location of the exit and its current position. However, due to secondary disasters, the latter is subject to the two following types of interference:

A persistent random perturbation which can be mathematically described as follows:

$$x_{observation} = x_{real} + \varepsilon$$

where

$$\varepsilon = \sum \alpha_i \varepsilon_i$$

where ε_i obey different types of probability distributions, including Gaussian distribution, uniform distribution, etc. α_i are coefficients.

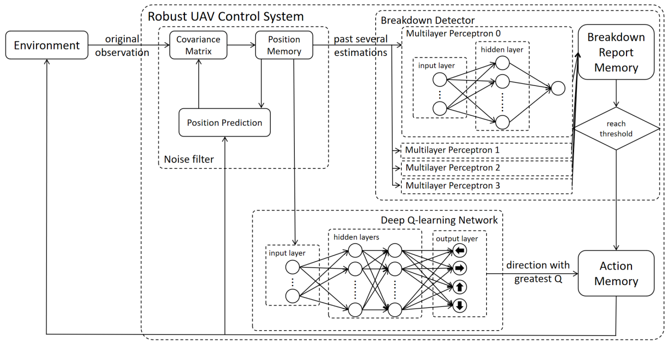

Fig. 2. Inner Structure of the Robust UAV Control System

A potential breakdown that can occur at any time t0 and persist thereafter, manifesting in any of the following three scenarios:

$$x_{observation} = x_{real} + \varepsilon + b, |b| \gg |\varepsilon| \quad (1)$$

$$x_{observation} = random \quad (2)$$

$$x_{observation(t=T)} \equiv x_{observation(t=t_0)} \quad (3)$$

B. Robust Control Problem and its Goal

To be robust, the agent should first ensure its safety.

Assuming that any collision with the walls by the agent results in immediate destruction, the robustness can be evaluated by the probability of colliding with the wall.

On the other hand, once a breakdown happens, the observation significantly deviates from the actual position and thus fails to contain enough information. In such cases, we regard halting as the most robust choice.

In general, to achieve robustness, the agent should find the right way towards the exit, avoid crashing into any wall and stop when a breakdown happens.

III. Robust UAV Control System

To better handle the high interference environment, we proposed a new robust UAV control system consisting of two sections to gradually process the original data and make robust decisions. The structure of our system is shown in Fig. 2.

A. Inner Structure Description

1) Noise Filter: The original observation data received from the environment will be transmitted to the Noise Filter section, which is supposed to reduce the random noise in the observation. This section is constructed roughly based on the Kalman filter. Every time, it make a prediction of its current position based on its memory of the last position and action. Then it uses a covariance matrix to combine its prediction with its current observation to get a much more reliable position and put it into its position memory for further use.

2) Breakdown Detector: When the process in the Noise Filter section completes, the latest several data in the memory will be transmitted to the Breakdown Detector section, which is supposed to detect whether there exists a breakdown in the observation. This section mainly consists of four Multilayer Perceptron. Each of them contains one hidden layer and has 20 nodes in the layer. Every time, they evaluate the data separately and report their results. The reports are saved in the memory and the section makes the final judgment by comparing the reports in the memory with a certain threshold. When a breakdown is believed to exist, it will directly determine the final action.

3) Deep Q Network: If the processed data is believed to be reliable, it will be passed to a the Deep Q Network section, which is supposed to make the final decision. It has two hidden layers with 10 nodes in each layer. Every time, it outputs four values based on the current position, representing the quality of each direction. Then it finds out the direction with the highest value which is supposed to be the optimal action.

B. Training Process

The neural networks mentioned above are all trained separately before they are combined together as the whole system.

1) MLPs: The four MLPs are trained in the same virtual environment with different initial parameters. It have the same interference pattern as the final environment, but the walls are removed and the movement is random. Every MLP receives the input as mentioned in 1.1, and generate a output ω range from 0 to 1 (due to the sigmoid function placed in the end) to represent the possibility to report. Then the label τ of each set of input is 1 if a breakdown exists and otherwise 0.

The loss function is defined as follow:

$$L_{(\omega,\tau)} = (\tau - \omega)^2$$

The parameters are then updated according to the gradient descent algorithm. The learning rate gradually decreases from 0.01 to 0.001. The update momentum is set to 0.2. In the experiment, the models are fully trained after 20000 episodes, when their separate accuracy stay unchanged.

2) DQN: In this paper, we adopt double deep Q-learning method to train the DQN section. The virtual environment is the non-interference version the final environment with intentionally set rewards.

The model is expected to estimate the quality of each action at any given state, which can be defined by Bellman Equation:

$$Q^*(s_t, a_i) = r_{(s_t,a_i)} + \gamma * max\left\{Q^*(s_{t+1}, a_i) | a_i \in A_{t+1}\right\}$$

where Q is the quality, s is the state, a is the action, r is the immediate reward and A is the action space.

When training, the model can get the exact reward of its latest move and updates its parameters θ by gradient descent algorithm, minimizing the loss function:

$$L_\theta = (Q_\theta(s_t, a_i) - y_{(t,i)})^2$$

where

$$y_{(t,i)} = r_{(s_t,a_i)} + \gamma * max\left\{Q_{\theta'}(s_{t+1}, a_i) | a_i \in A_{t+1}\right\}$$

where θ' represents the parameters of the target network introduced by DDQN to optimize the training process.

The learning rate gradually decreases from 0.01 to 0.001. The update momentum is set to 0.2. The exploration rate is set to 0.2. The target DQN is updated every 4 episodes.

The rewards in the environment are set to encourage the agent to gradually shift its zone and get closer to the exit.

In our experiment, it is fully trained after about 100000 episodes when it can always succeed in getting to the exit under precise observation.

IV. Performance Comparison Between Robust UAV Control System and Other Methods

In this section, we apply the robust UAV control system to the virtual environment to evaluate its ability to handle extreme conditions. Other methods are also applied to be compared.

The performance of the robust UAV control system and other methods under significant interference and potential breakdowns is shown in Table I.

TABLE I
Performance of the Robust System and Other Methods under Significant Interference and Potential Breakdowns

Method	Collision	Escape	Stop	Successful Detection
DQN	75.9%	24.1%	0.0%	0.0%
Hidden Markov	31.8.0%	68.2%	0.0%	0.0%
Robust System	18.0%	64.8%	17.2%	94.0%

Obviously, our system has a significantly lower collision rate, a significantly higher escape rate than the DQN alone. Although its escape rate is slightly lower than the hidden markov model, its collision rate is far lower. In real life, the collision rate can greatly affect the loss. Therefore, it's actually more important for the control system to be robust. From this prospective, our system is much better than the established hidden markov model.

In order to more intuitively demonstrate how our system and other methods react to the environment, we visualize one of the episodes for each of them.

Fig. 3 and Fig. 4 respectively shows the results of the application of the DQN and our system, where the red line represents the route the agent actually takes and the blue dots are the record of observations.

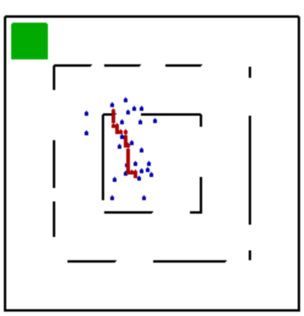

Fig. 3. Application of the DQN

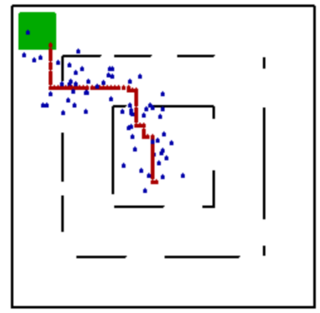

Fig. 4. Application of the robust UAV control system

TABLE II
Performance of the Robust System and Other Methods under Extreme Interference and Potential Breakdowns

Method	Collision	Escape	Stop	Successful Detection
DQN	97.6%	2.4%	0.0%	0.0%
Hidden Markov	45.7%	54.3%	0.0%	0.0%
Robust System	32.9%	52.7%	15.4%	93.5%

To further explore the anti-interference capability of this system, we apply them to an environment with even greater interference. The results are shown in Table II.

Fig. 5 and Fig. 6 respectively shows the results of the application of the DQN and our system under extreme interference.

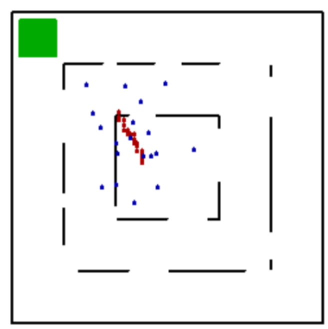

Fig. 5. Application of the DQN under extreme interference

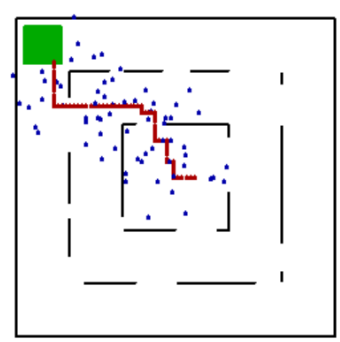

Fig. 6. Application of the robust UAV control system under extreme interference

In such condition, our system has even greater advantage over other methods, maintaining low collision rate and thus being robust and suitable for real life usage.

Therefore, we regard our system to be the most robust one and is superior to other methods under significant interference and potential breakdowns.

V. Conclusion and Expectation

In this paper, we provide an in-depth study of the path planning problem of unmanned aerial vehicles (UAVs) in complex environments and explore various strategies to improve the robustness of the control system. We propose a novel robust UAV control system (robust-UAV-control-system) that combines Kalman filter (KF) and multilayer perceptron (MLP) architectures. Our system performs significantly superior in a predefined experimental environment compared to the traditional pure Deep Q-Network (DQN) approach and the traditional Hidden Markov Model. The experiments demonstrate that enhancing the complexity of the UAV control system and introducing advanced filtering algorithms are key choices for improving the robustness of the system.

Looking ahead, we plan to investigate whether similar or even better control results can be achieved by modifying the architecture of the DQN (e.g., changing it to a robust-DQN, or by introducing an auxiliary neural network (e.g., a long short-term memory network LSTM). Meanwhile, we also hope to explore the scalability and applicability of the system and evaluate its robustness and generalisability in other control problems to further validate its potential and value in practical applications.

References

[1] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, "Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination," *Computer Networks*, vol. 101, pp. 158–168, 2016.

[2] M. Bertolini, D. Mezzogori, M. Neroni, and F. Zammori, "Machine learning for industrial applications: A comprehensive literature review," *Expert Systems with Applications*, vol. 175, 2021.

[3] G. Fragapane, D. Ivanov, M. Peron, F. Sgarbossa, and J. O. Strandhagen, "Increasing flexibility and productivity in industry 4.0 production networks with autonomous mobile robots and smart intralogistics," *Annals of Operations Research*, vol. 308, no. 1-2, pp. 125–143, 2022.

[4] G. Fragapane, R. de Koster, F. Sgarbossa, and J. O. Strandhagen, "Planning and control of autonomous mobile robots for intralogistics: Literature review and research agenda," *European Journal of Operational Research*, vol. 294, no. 2, pp. 405–426, 2021.

[5] C.-S. Karavas, G. Kyriakarakos, K. G. Arvanitis, and G. Papadakis, "A multi-agent decentralized energy management system based. on distributed intelligence for the design and control of autonomous polygeneration microgrids," *Energy Conversion and Management*, vol. 103, pp. 166–179, 2015.

[6] S. A. H. Mohsan, M. A. Khan, F. Noor, I. Ullah, and M. H. Alsharif, "Towards the unmanned aerial vehicles (uavs): A comprehensive review," *Drones*, vol. 6, no. 6, 2022.

[7] P. K. Reddy Maddikunta, S. Hakak, M. Alazab, S. Bhattacharya, T. R. Gadekallu, W. Z. Khan, and Q.-V. Pham, "Unmanned aerial vehicles in smart agriculture: Applications, requirements, and challenges," *Ieee Sensors Journal*, vol. 21, no. 16, pp. 17 608–17 619, 2021.

[8] N. Zhao, W. Lu, M. Sheng, Y. Chen, J. Tang, F. R. Yu, and K.-K. Wong, "Uav-assisted emergency networks in disasters," *Ieee Wireless Communications*, vol. 26, no. 1, pp. 45–51, 2019.

[9] S. Xu and H. Peng, "Design, analysis, and experiments of preview path tracking control for autonomous vehicles," *Ieee Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 48–58, 2020.

[10] Z. Peng, J. Wang, and Q.-L. Han, "Path-following control of autonomous underwater vehicles subject to velocity and input constraints via neurodynamic optimization," *Ieee Transactions on Industrial Electronics*, vol. 66, no. 11, pp. 8724–8732, 2019.

[11] B. Kiumarsi, K. G. Vamvoudakis, H. Modares, and F. L. Lewis, "Optimal and autonomous control using reinforcement learning: A survey," *Ieee Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2042–2062, 2018. [Online]. Available: ¡Go to ISI¿://WOS:000432398300002

[12] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, "A survey of deep learning applications to autonomous vehicle control," *Ieee Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 712–733, 2021.

[13] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning ¡i¿a brief survey¡/i¿," *Ieee Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[14] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. Al Sallab, S. Yogamani, and P. Perez, "Deep reinforcement learning for autonomous driving: A survey," *Ieee Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2022.

[15] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction second edition introduction," Editorial Material; Book Chapter, 2018, times Cited: 6993. [Online]. Available: ¡Go to ISI¿://WOS:000481873900002

[16] V. Talpaert, I. Sobh, B. R. Kiran, P. Mannion, S. Yogamani, A. El-Sallab, and P. Perez, "Exploring applications of deep reinforcement learning for real-world autonomous driving systems," in *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP)*, 2019, pp. 564–572.

[17] T. Zhang, J. Lei, Y. Liu, C. Feng, and A. Nallanathan, "Trajectory optimization for uav emergency communication with limited user equipment energy: A safe-dqn approach," *Ieee Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1236–1247, 2021.

[18] X. Wu, H. Chen, C. Chen, M. Zhong, S. Xie, Y. Guo, and H. Fujita, "The autonomous navigation and obstacle avoidance for usvs with anoa deep reinforcement learning method," *Knowledge-Based Systems*, vol. 196, 2020.

[19] M. Zhu, Y. Wang, Z. Pu, J. Hu, X. Wang, and R. Ke, "Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving," *Transportation Research Part C-Emerging Technologies*, vol. 117, 2020.

[20] J. Chen, S. E. Li, and M. Tomizuka, "Interpretable end-to-end urban autonomous driving with latent deep reinforcement learning," *Ieee Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5068–5078, 2022.

[21] A. Singla, S. Padakandla, and S. Bhatnagar, "Memory-based deep reinforcement learning for obstacle avoidance in uav with limited environment knowledge," *Ieee Transactions on Intelligent Transportation Systems*, vol. 22, no. 1, pp. 107–118, 2021.

[22] N. H. Amer, H. Zamzuri, K. Hudha, and Z. A. Kadir, "Modelling and control strategies in path tracking control for autonomous ground vehicles: A review of state of the art and challenges," *Journal of Intelligent Robotic Systems*, vol. 86, no. 2, pp. 225–254, 2017.

[23] Q. Zhang, J. Lin, Q. Sha, B. He, and G. Li, "Deep interactive reinforcement learning for path following of autonomous underwater vehicle," *Ieee Access*, vol. 8, pp. 24 258–24 268, 2020.

[24] S. Guo, X. Zhang, Y. Du, Y. Zheng, and Z. Cao, "Path planning of coastal ships based on optimized dqn reward function," *Journal of Marine Science and Engineering*, vol. 9, no. 2, 2021.

[25] W. Wei, J. Wang, Z. Fang, J. Chen, Y. Ren, and Y. Dong, "3u: Joint design of uav-usv-uuv networks for cooperative target hunting," *Ieee Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 4085–4090, 2023.

[26] J. Li, X. Xiang, D. Dong, and S. Yang, "Prescribed time observer based trajectory tracking control of autonomous underwater vehicle with tracking error constraints," *Ocean Engineering*, vol. 274, 2023.

[27] F. Tian, Z. Li, F.-Y. Wang, and L. Li, "Parallel learning-based steering control for autonomous driving," *Ieee Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 379–389, 2023.

[28] H.-T. Sun, C. Peng, X. Ge, and Z. Chen, "Secure event-triggered sliding control for path following of autonomous vehicles under sensor and actuator attacks," *Ieee Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 981–992, 2024.

[29] C. Hu, H. Jing, R. Wang, F. Yan, and M. Chadli, "Robust ¡i¿h¡/i¿¡sub¿¡/sub¿ output-feedback control for path following of autonomous ground vehicles," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 414–427, 2016.

[30] J. Lu, L. Han, Q. Wei, X. Wang, X. Dai, and F.-Y. Wang, "Event-triggered deep reinforcement learning using parallel control: A case study in autonomous driving," *Ieee Transactions on Intelligent Vehicles*, vol. 8, no. 4, pp. 2821–2831, 2023.

[31] F. Dang, D. Chen, J. Chen, and Z. Li, "Event-triggered model predictive control with deep reinforcement learning for autonomous driving," *Ieee Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 459–468, 2024.

[32] L. Chen, X. Hu, B. Tang, and Y. Cheng, "Conditional dqn-based motion planning with fuzzy logic for autonomous driving," *Ieee Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 2966–2977, 2022.

[33] Y. Yao, J. Zhao, Z. Li, X. Cheng, and L. Wu, "Jamming and eavesdropping defense scheme based on deep reinforcement learning in autonomous vehicle networks," *Ieee Transactions on Information Forensics and Security*, vol. 18, pp. 1211–1224, 2023.

[34] L. Yang, Y. Xu, J. Tian, and D. Wang, "Autonomous path planning and anti-wind disturbance control of tilt-rotor aircraft in complex and dense environment," *Journal of Applied Science and Engineering*, vol. 28, no. 1, pp. 61–73, 2025.

[35] Q. Yang, Y. Zhu, J. Zhang, S. Qiao, J. Liu, and Ieee, "Uav air combat autonomous maneuver decision based on ddpg algorithm," in *IEEE 15th International Conference on Control and Automation (ICCA)*, ser. IEEE International Conference on Control and Automation ICCA, 2019, pp. 37–42.

[36] K. Wan, X. Gao, Z. Hu, and G. Wu, "Robust motion control for uav in dynamic uncertain environments using deep reinforcement learning," *Remote Sensing*, vol. 12, no. 4, 2020.

[37] G. Munoz, C. Barrado, E. Cetin, and E. Salami, "Deep reinforcement learning for drone delivery," *Drones*, vol. 3, no. 3, 2019.

[38] O. Perez-Gil, R. Barea, E. Lopez-Guillen, L. M. Bergasa, C. Gomez-Huelamo, R. Gutierrez, and A. Diaz-Diaz, "Deep reinforcement learning based control for autonomous vehicles in carla," *Multimedia Tools and Applications*, vol. 81, no. 3, pp. 3553–3576, 2022.

[39] X. Ruan, D. Ren, X. Zhu, J. Huang, and Ieee, "Mobile robot navigation based on deep reinforcement learning," in *31st Chinese Control And Decision Conference (CCDC)*, ser. Chinese Control and Decision Conference, 2019, pp. 6174–6178.

[40] L. Li, D. Wu, Y. Huang, and Z.-M. Yuan, "A path planning strategy unified with a colregs collision avoidance function based on deep reinforcement learning and artificial potential field," *Applied Ocean Research*, vol. 113, 2021.

[41] N. Gao, Z. Qin, X. Jing, Q. Ni, and S. Jin, "Anti-intelligent uav jamming strategy via deep q-networks," *Ieee Transactions on Communications*, vol. 68, no. 1, pp. 569–581, 2020.

[42] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. G. van Sloun, and Y. C. Eldar, "Kalmannet: Neural network aided kalman filtering for partially known dynamics," *Ieee Transactions on Signal Processing*, vol. 70, pp. 1532–1547, 2022.