Misalignment Tolerance Enhancement and Intelligent Minimal-Activation Pathway Decision for Unmanned Aerial Vehicles WPT Applications

Shuai Wu, *Member, IEEE*, Jingjie Yang, Chunwei Cai, *Member, IEEE*, Wenping Chai, *Member, IEEE*, and Jinpeng Yu, *Senior Member, IEEE*

***Abstract*—This article introduces a reinforcement learning (RL)-enhanced wireless charging system for unmanned aerial vehicles (UAVs), addressing critical challenges of misalignment tolerance and energy-efficient coil matching in autonomous landing scenarios. A reconfigurable magnetic array generates a two-dimensional rotating field to accommodate positional/angular deviations while a Q-learning framework maps stochastic UAV positions into grid states via pick-up voltage feedback, optimizing energy transfer unit activation through multiobjective rewards (efficiency, switching cost, boundaries). Experimental validation shows 354.4 W output at 87.74% dc–dc efficiency, with 92.2% fewer activation steps than exhaustive methods. The RL strategy achieves average 5-step activations and <4.2 cm coarse positioning errors, enabling three-dimensional spatial awareness without prior models. By synergizing adaptive magnetic structures with model-free RL, this work establishes a scalable UAV charging solution, balancing high misalignment tolerance, and efficient power delivery. This article is also accompanied by a video file demonstrating training experiment in PyCharm.**

***Index Terms*—Coil matching, Q-learning algorithm, reinforcement learning (RL), unmanned aerial vehicle (UAV), wireless power transfer (WPT).**

I. Introduction

UNMANNED aerial vehicle (UAV) as an automatic equipment, extensively undertakes a variety of tasks including power inspection, power transmission, and plant protection operations[1], [2]. However, its endurance is significantly constrained by the limitations of onboard battery capacity [3]. Traditional wired charging methods or battery replacement operations require manual intervention. In light of this, Wireless power transfer (WPT) technology, with its unique noncontact advantage, is gradually penetrating and being applied in the field of UAVs, providing a new approach to address endurance issues [4].

Given the peculiar structure and inevitable landing position deviation of UAVs, achieving efficient and stable power transmission necessitates precise alignment between the transmitter and receiver [5], [6]. To overcome these obstacles, some research efforts have been conducted. In [7], an innovative concave energy transmission platform which integrates a specially designed landing pad with a modified landing gear configuration has been proposed, thereby facilitating precise alignment through mechanical aided design. However, it necessitates modifications to the UAV's structural configuration, and potential water accumulation may lead to system failure. Besides, focusing on the WPT system itself, some scholars have proposed an anti-misalignment method utilizing a large-to-small asymmetric magnetic design [8], [9]. However, a significant portion of the emitted magnetic flux does not participate in effective coupling, which result in weak system coupling capability. Furthermore, adopting the transmitting array coil also provide an effective approach to address the misalignment issue and enhance the adjustability and flexibility of WPT systems [10]. In addition, the corresponding efficient power supply topology for exciting the array coils were proposed in papers [11] and [12]. By employing power direction measures within the array transmitter, the degradation of transmission performance due to misalignment under a single transmitting coil can be mitigated and effectively expands the charging area according to the power supply requirements [13], [14]. The authors propose a design for an array transmission platform featuring peak-valley superimposed output, which is suitable for UAVs [15]. However, the UAV-WPT systems with an array configuration necessitates the selective excitation and activation of array coils based on the position of the receiver, or it will result in significant cross-coupling and needless system losses. So, precise receiver positioning and coil matching are prerequisite for achieving power directional transmission, thereby some researches have been conducted in this area.

Currently, the positioning technology primarily employs GPS, infrared and visual identity [16], [17]. UAVs achieve precise landing by recognizing specific markers, such as the

Received 2 July 2025; accepted 14 August 2025. Date of publication 18 August 2025; date of current version 23 December 2025. This work was supported in part by the Natural Science Foundation of Shandong Province under Grant ZR2024QE121, in part by the Major Scientific and Technological Innovation Project of Shandong Province of China under Grant 2022ZLGX04, and in part by the National Natural Science Foundation of China under Grant 52177002. Recommended for publication by Associate Editor A. Kuperman. *(Corresponding author: Chunwei Cai.)*

Shuai Wu, Jingjie Yang, Chunwei Cai, and Wenping Chai are with the School of New Energy, Harbin Institute of Technology-Weihai, Weihai 264209, China, and also with the Key Laboratory of Cross-Domain Synergy and Comprehensive Support for Unmanned Marine Systems, Ministry of Industry and Information Technology, Weihai 264209, China (e-mail: wushuai@hit.edu.cn; 23s030160@stu.hit.edu.cn; caichunwei@hit.edu.cn; wpchai@hit.edu.cn).

Jinpeng Yu is with the School of Automation, Qingdao University, Qingdao 266071, China (e-mail: yjp1109@qdu.edu.cn).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TPEL.2025.3599897.

Digital Object Identifier 10.1109/TPEL.2025.3599897

"H"-shaped landing indicator, through their imaging systems. However, this landing process relies heavily on high-precision image recognition systems, leading to an escalation in costs. Furthermore, the adoption of dynamic landing methodologies necessitates the implementation of robust flight control algorithms to ensure accurate actions. Consequently, the system complexity and control challenges increase significantly. Additionally, inductive information within wireless charging system can be acquired through auxiliary coils or magnetoresistance sensor to achieve position detection of the receiver, and then further realize coil matching [18]. This method based on magnetic induction is widely researched due to its simple structure and low cost. In [19], a position detection method based on the negative coupling characteristic of the extended coils is proposed. However, the derivation of state detection quantity models based on mutual inductance is complex.

Furthermore, researchers have introduced machine learning to perform position detection by analyzing the nonlinear relationship between state variables and positional offsets. In [20], Srivastava et al. applied it to collect voltage samples and predict position of the receiver, so as to effectively supply power to the biomedical implant devices. Zhang et al. [21] further utilized machine learning for supervised learning of characteristic data such as the longitudinal offset of array coils and system transmission efficiency to predict the position of the receiver. However, when the system is located in the boundary area, its position estimation capability is relatively weak due to the difficulty of the BP algorithm in processing abrupt sample information.

The aforementioned machine learning methodologies predominantly employ supervised learning frameworks, which establish input–output mapping relationships through labeled training data to achieve model fitting and subsequent position detection. However, in array-configured WPT systems, mere UAV position inference proves insufficient. The system necessitates advanced coil matching mechanisms—through dynamic array coil switching, the agent can intelligently activate the optimal energy transfer units based on real-time analysis of UAV positioning and current coupling states. Specifically, reinforcement learning (RL) represents a goal-oriented paradigm of machine learning, which learns optimal behavioral strategies through interaction with the environment. It has been extensively applied to topology design and parameter optimization, particularly suited to addressing the aforementioned coil activation sequence decision-making and optimization problems due to its characteristics [22], [23].

Current research limitations in magnetic field positioning methodologies reveal two critical gaps: First, while machine learning techniques have successfully identified implicit correlations within datasets, the transition from data-driven pattern recognition to intelligent decision-making remains underdeveloped, with limited research conducted on intelligent decision-making applications in UAV-WPT systems. Second, existing methodologies in the literatures inadequately address the critical factor of angular misalignment in receiving coils. The prevalent two-dimensional (2-D) positioning framework fundamentally restricts precise coil alignment capabilities, as it disregards the 3-D spatial orientation of UAVs. This spatial dependency directly governs mutual inductance variations between power transmission coils, thereby compromising both electromagnetic field-based positioning accuracy and wireless power transmission efficiency.

To address these issues, this article introduces a reconfigurable and modular WPT system for UAVs arranged in an array configuration, which is addressed in Section II. The system establishes a 2-D dynamic rotating magnetic field based on a reconfigurable array coil structure and three-phase excitation currents, thereby enhancing the capability of misalignment tolerance, as shown in Fig. 1(a). Meanwhile, two reconfigurable three-phase energy transfer units supply power to the receiving coils located on both sides of the UAV's landing gear, as shown in Fig. 1(c). In addition, given the randomness of the UAV's landing position and angle on the array platform, to enable the array platform to selectively activate the optimal energy transfer unit for power transfer based on the position of the receiving coils, RL is introduced to learn the electromagnetic field distribution characteristics under the current environment. Ultimately, this approach achieves the search for and identification of the optimal energy transmission array unit with the shortest coil excitation sequence, thereby enabling efficient power transfer, which is shown in Fig. 1(b). This approach enables the dynamic matching of coils based on the UAV's landing position and determines the shortest activation pathway, ultimately facilitating efficient power transfer, which is incorporated in Section III. The experimental prototype system is set up in Section IV, the performance of coil matching and energy transfer is verified. Finally, Section V concludes this article.

II. Overall Structure of the UAV WPT System Featuring Omnidirectional Energy Transmission

A. Magnetic Structure

The proposed transmitter of charging platform consists of array coils, underneath which ferrite magnetic cores are laid to minimize magnetic leakage. Each array module consists of a total of 19 transmitter coils. This configuration is adopted to maintain a hexagonal-like shape for individual modules, facilitating seamless splicing between modules and cross-module grouping. Equal-amplitude excitation currents with a phase difference of 120° are supplied to three adjacent coils, respectively, forming a reconfigurable three-phase fundamental energy transfer unit (ETU) that generates a 2-D omnidirectional rotating magnetic field. The receiving coils embedded in an orthogonal configuration within the UAV's landing gear, enhancing the coupling of parallel magnetic flux, thereby improving misalignment tolerance.

B. Array Coil Switching Circuit Topology Driven By Intelligent Algorithm

The circuit topology of the transmission module comprises a dc voltage source input, a three-phase half-bridge inverter, and CLC compensation networks, as illustrated in Fig. 2.

The adoption of SS compensation on the secondary side is primarily due to the UAV's sensitivity to onboard weight

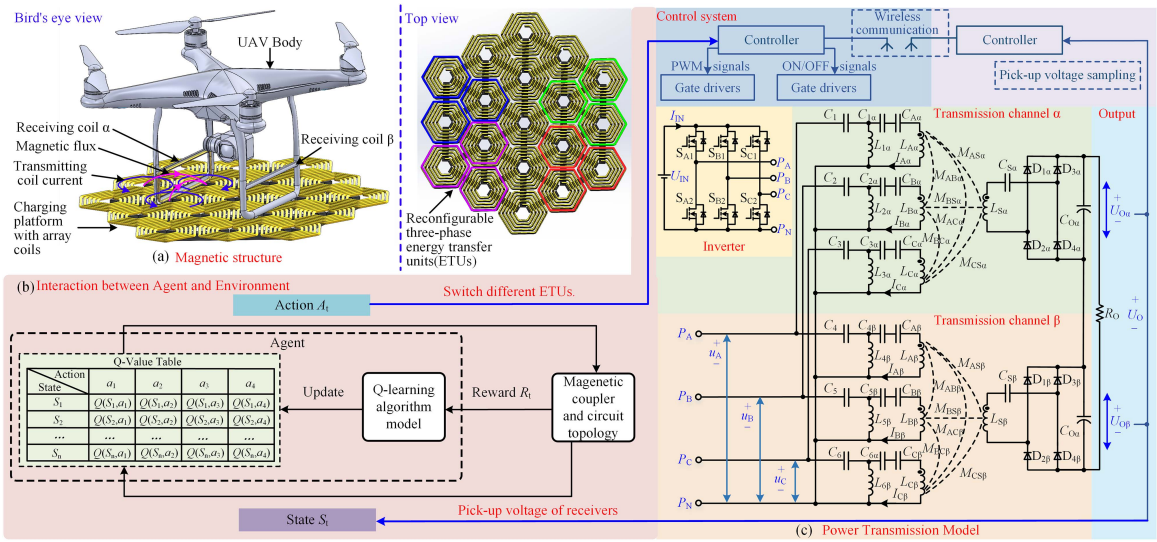


Fig. 1. Overview of this paper. (a) Proposed magnetic structure installed in UAV body. (b) Proposed interactive framework of the algorithm model. (c) Circuit diagram of the receiving system with coupling channels and schematic diagram of control system.

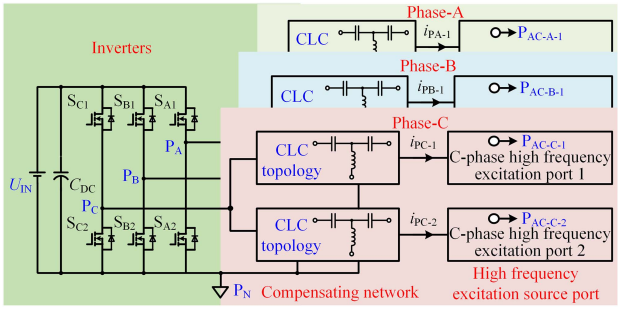


Fig. 2. Three phase magnetizing sources.

requirements. Moreover, compared to the *LCL* compensation topology, which also achieves constant current excitation, the *CLC* compensation topology features capacitors that are smaller in volume relative to inductors, which effectively reduces system complexity and enhances the integration level of the array system. This configuration enables the module to provide high-frequency constant-current excitation for the transmitting coils. The array transmission module integrates 19 transmitting coils organized into six groups (A-1, A-2, B-1, B-2, C-1, and C-2). Each phase of the inverter is connected in parallel with two *CLC* compensation networks for grouping control, effectively preventing short-circuit risks caused by simultaneous activation of two coils with identical phase sequences during coil matching. This design establishes six independent constant-current excitation sources, which are individually connected to the six coil groups through selection switches.

To achieve high-frequency excitation and selective activation of array coils, a flexible switching array with multicoils control capability has been developed as depicted in Fig. 3. Each excitation channel incorporates a dedicated conducting switch to minimize the ON-state resistance of the switching network, which ensures precise current regulation while maintaining operational safety and energy efficiency. In addition, each transmitter coil is serially connected with a gating switch composed of two reverse-series-connected MOSFETs. During charging and matching operations, only one coil per group is activated.

The HRTIM module of the STM32 microcontroller generates pulsewidth modulation signals that meet the requirements for exciting an omnidirectional coupling magnetic field, enabling multiphase inverter output control and driving the array switches. The control action logic is illustrated using Phase A output as an example, as shown in Fig. 4.

The *entire* CLC network exhibits constant-current excitation characteristics. When the switches are turned OFF, the voltage stress rises sharply due to the inductive nature of the array coils, which could burn out the MOSFETS. To address this issue, the specific design incorporates programmed timing priority control, where the inverter output signal actions are given higher priority than the switch ON/OFF timing. Specifically, during array coil switching, the inverter output is first controlled to turn OFF before the array coil is switched, ensuring that there is no current output during switch operations. This prevents the problem of MOSFET burn out caused by high voltage stress. Furthermore, with this design, the array switches do not need to achieve zero-current switching to reduce system losses, as there is no overall system output during switch operations.

C. Modeling and Parametric Investigation of Power Transmission Framework

Since all transmitting and receiving coils possess identical dimensions, the component values and mutual inductance within the system are all the equivalent as

$$\begin{cases} M_{\mathrm{AB}\alpha}=M_{\mathrm{AC}\alpha}=M_{\mathrm{BC}\alpha}=M_P, L_{\mathrm{A}\alpha}=L_{\mathrm{B}\alpha}=L_{\mathrm{C}\alpha}=L_P, \\ C_{\mathrm{A}\alpha}=C_{\mathrm{B}\alpha}=C_{\mathrm{C}\alpha} \\ L_{1\alpha}=L_{2\alpha}=L_{3\alpha}=L_1, C_1=C_2=C_3=C_{1\alpha}=C_{2\alpha}=C_{3\alpha} \end{cases}. \tag{1}$$

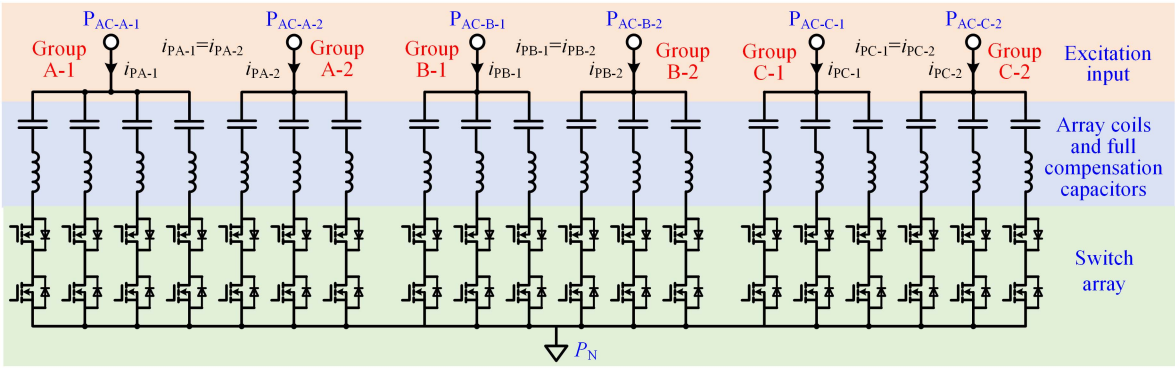


Fig. 3. Array coils and switch structure.

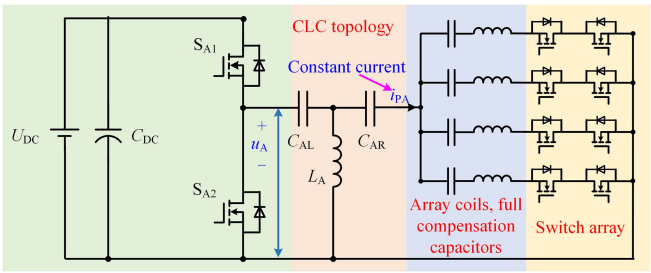


Fig. 4. Phase output topology.

Given the absence of coupling between the two transmission channels, which operate as analogous, parallel power transmission processes, the modeling and analysis presented herein focuses exclusively on a detailed examination of transmission channel α.

As the three-phase inverter output with equal amplitude and 120° difference is adopted, the output voltages of A, B, and C three-phase inverters meet the following requirements:

$$\begin{cases} \dot{U}_A = \frac{\sqrt{2}U_{\mathrm{DC}}}{\pi} = U_P \\ \dot{U}_B = \frac{\sqrt{2}U_{\mathrm{DC}}}{\pi} e^{-j120^\circ} = U_P e^{-j120^\circ} \\ \dot{U}_C = \frac{\sqrt{2}U_{\mathrm{DC}}}{\pi} e^{j120^\circ} = U_P e^{j120^\circ} \end{cases}. \tag{2}$$

The receiving side adopts S-S compensation to meet the resonance state, so the change of circuit parameters at the receiving side will not affect the resonance state at the transmitting side. In addition, considering the mutual coupling of three adjacent coils in the same energy transmission unit, taking the A-phase coil as an example, the reflection impedance of the B and C two-phase transmitting coils to the A-phase transmitting coil is satisfied as follows:

$$\begin{aligned} \dot{Z}_{rA} &= \frac{j\omega M_{AB\alpha}\dot{I}_{B\alpha} + j\omega M_{AC\alpha}\dot{I}_{C\alpha}}{\dot{I}_{A\alpha}} \\ &= j\omega M_{AB\alpha} e^{-j120^\circ} + j\omega M_{AC\alpha} e^{j120^\circ} = -j\omega M_P. \end{aligned} \tag{3}$$

Similarly, it can be concluded that the reflection impedances of the other two phases reflected to the B and C two-phase coils are all jwM_{P}, and further, all compensation capacitors are connected in series with each phase coil to offset the self-inductance of the transmitting coil and the equivalent reflection impedance between the same energy transmission unit, so that the array coils are purely resistive, then the following conditions are met as

$$j\omega L_{A\alpha} - j\omega M_P = j\frac{1}{\omega C_{A\alpha}}, C_{A\alpha} = \frac{1}{\omega^2 (L_{A\alpha} - M_P)}. \tag{4}$$

The transmitting-side circuit and the receiving-side circuit remain in a resonant state at the same frequency, and at this state, the resonance conditions are further satisfied as follows:

$$\omega = \frac{1}{\sqrt{L_{1\alpha}C_1}} = \frac{1}{\sqrt{L_{1\alpha}C_{1\alpha}}} = \frac{1}{\sqrt{L_{\mathrm{S}\alpha}C_{\mathrm{S}\alpha}}}. \tag{5}$$

Therefore, under this resonance condition, the input impedance of the A-phase coil can be further obtained. Similarly, the input impedance of the B-phase coil and the C-phase coil can be obtained as follows:

$$\dot{Z}_A = \frac{L_{1\alpha}}{Z_{sA}C_1}, \dot{Z}_B = \frac{L_{2\alpha}}{Z_{sB}C_2}, \dot{Z}_C = \frac{L_{3\alpha}}{Z_{sC}C_3}. \tag{6}$$

The input impedance of the system is purely resistive and has nothing to do with the change of mutual inductance and load. At this time, according to Ohm's law, it can be further obtained that the excitation current of the input A-phase coil meets the following requirements:

$$\dot{I}_{A\alpha} = \frac{\dot{U}_A}{\dot{Z}_A} \times \frac{j\omega L_{1\alpha}}{j\omega L_{1\alpha} + \left(\frac{1}{j\omega C_{1\alpha}} + Z_{sA}\right)} = -\frac{\dot{U}_A}{j\omega L_{1\alpha}} = -\frac{U_P}{j\omega L_1}. \tag{7}$$

Similarly, the input currents of B and C two-phase coils can be obtained as follows:

$$\begin{cases} \dot{I}_{B\alpha} = -\frac{\dot{U}_B}{j\omega L_{2\alpha}} = -\frac{U_P e^{-j120^\circ}}{j\omega L_1} \\ \dot{I}_{C\alpha} = -\frac{\dot{U}_C}{j\omega L_{3\alpha}} = -\frac{U_P e^{j120^\circ}}{j\omega L_1} \end{cases}. \tag{8}$$

In addition, the harmonic in the output voltage of each half-bridge inverter is filtered by *CLC* compensation network, and the high-frequency excitation source acting on the array coils is constant current excitation. When the system frequency is determined, the excitation current is only related to the input voltage of the compensation network and the compensation inductance L_1. Because L_1 is not affected by dislocation and is a constant value, the excitation current can be constant as

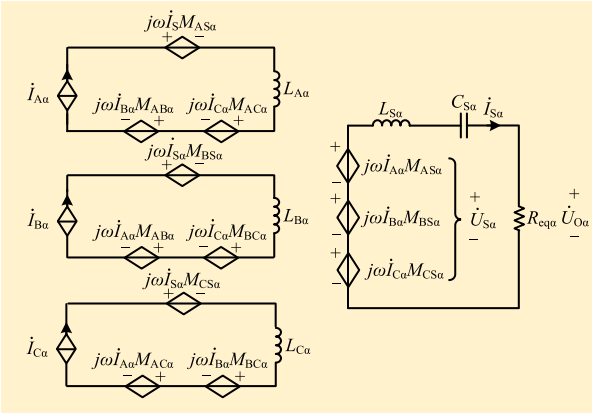


Fig. 5. Equivalent circuit of the transmission channel α.

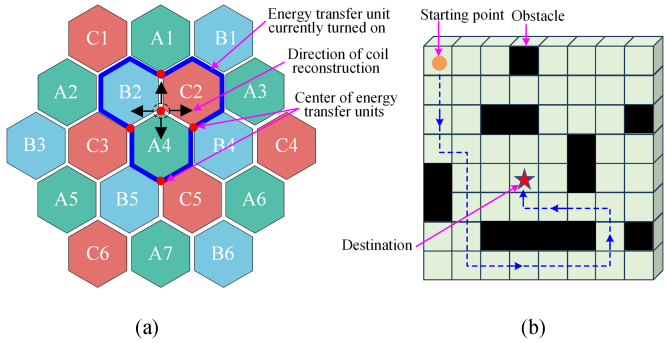


Fig. 6. Algorithm model reconstruction. (a) Optimization of coil matching. (b) Optimization of path planning.

long as the input voltage is fixed, and even if the receiving- side is in no-load operation, the excitation current is still a constant value.

Based on this excitation condition, the induced electromotive force generated at the receiving-side by the transmitting unit composed of three-phase transmitting coils in an energy transmission channel is shown in Fig. 5, which is further satisfied as follows:

$$\dot{U}_{S\alpha} = j\omega M_{\mathrm{AS}\alpha}\dot{I}_{A\alpha} + j\omega M_{\mathrm{BS}\alpha}\dot{I}_{B\alpha} + j\omega M_{\mathrm{CS}\alpha}\dot{I}_{C\alpha}. \quad (9)$$

Substituting (2), (7), and (8) into (9) can further obtain the following:

$$\dot{U}_{S\alpha} = \frac{U_P}{2L_1}\left[(2M_{\mathrm{AS}\alpha} - M_{\mathrm{BS}\alpha} - M_{\mathrm{CS}\alpha}) + j\sqrt{3}\,(-M_{\mathrm{BS}\alpha} + M_{\mathrm{CS}\alpha})\right]. \quad (10)$$

Since the receiving-side adopts S-S compensation to form series resonance between the self-inductance of the receiving coil and the compensation capacitor, it can be further obtained that the rectified dc output voltage $U_{o\alpha}$ is satisfied as follows:

$$U_{\mathrm{O}\alpha} = \frac{\sqrt{2}\pi}{4}|\dot{U}_{S\alpha}| = \frac{U_{\mathrm{IN}}}{2L_1}M_{\mathrm{eq}\alpha} \quad (11)$$

where $M_{\mathrm{eq}\alpha}$ is the equivalent mutual inductance composite value of the transmitting and receiving sides in the energy transmission channel α, where $M_{\mathrm{eq}\alpha}$ is further derived as follows:

$$M_{\mathrm{eq}\alpha} = \sqrt{M_{\mathrm{AS}\alpha}^2 + M_{\mathrm{BS}\alpha}^2 + M_{\mathrm{CS}\alpha}^2 - M_{\mathrm{AS}\alpha}M_{\mathrm{BS}\alpha} - M_{\mathrm{BS}\alpha}M_{\mathrm{CS}\alpha} - M_{\mathrm{AS}\alpha}M_{\mathrm{CS}\alpha}}. \quad (12)$$

In this article, the pick-up voltage delivered to the load is sampled by ADC and then these data are fed back to the primary-side so as to switch different coil combinations and regulate power flow by the controller.

III. Proposed Coil Localization and Intelligent Matching Method Based on Q-Learning Algorithm

The reconfigurable array magnetic structure can effectively mitigate the issue of landing misalignment. It is important to select the transmitting coils to optimize efficiency at different positions. So, this paper aims to integrate an algorithmic model for enabling random landing for UAV on the array charging platform and facilitating precise matching of the receiving coils with a three-phase energy transfer unit in the optimal coupling state. Ultimately, the model achieves the identification of the optimal energy transmission array unit with the shortest coil excitation sequence for efficient power transfer.

A. Matching Principle for UAV Wireless Charging Applications and Algorithmic Reconstruction

RL fundamentally optimizes an agent's behavioral strategy through iterative interactions with its environment. Specifically, an agent executes a sequence of actions within a predefined environment, while exhibiting its instantaneous state representation, as shown in Fig. 1(b). It learns to adjust its strategy based on the feedback triggered by these actions, namely reward or penalty signals, aiming to maximize the cumulative reward. Within this framework, the core mechanism of the Q-learning algorithm involves constructing and maintaining a Q-value table, which records the expected cumulative rewards for performing various actions under various system states in the form of state-action pairs. By establishing a specific algorithmic model and data set within the context of wireless power transfer for UAVs and fully exploring the environmental information during the charging process to achieve convergence of the Q-value table, optimal strategies and outputs can be determined by querying the Q-value table, circumventing the inefficient approach of sequentially iterating through and activating coils.

The fundamental rationale for reconstructing the array coil optimization problem as a path planning task stems further from their intrinsic structural isomorphism in decision-making mechanisms [24], [25]. As illustrated in Fig. 6(a), the primary optimization goal requires the intelligent platform to dynamically identify the optimal energy transfer unit based on real-time coupling state parameters when a UAV randomly lands, thereby achieving impedance matching between the array coils

and receivers for enhanced power transfer efficiency. This multiobjective optimization framework demonstrates remarkable parallels with Markov decision processes in RL-based robotic path planning, as shown in Fig. 6(b). Specifically, both of two systems involve the following.

1) State-space mapping where electrical coupling parameters correspond to spatial coordinates.
2) Action selection mechanisms comparing energy unit activation to movement direction choices.
3) Reward functions balancing transmission efficiency against path length optimization.
4) Under specific condition settings, constraints on system boundaries and system losses are analogous to obstacle avoidance requirements in path planning.

The RL agent, starting from an initial state, i.e., path origin, navigates through constrained environments to reach target destinations i.e., optimal ETU while optimizing trajectory parameters. So, based on these reasons, the selection of array coils optimization shares equivalent mathematical formulation with robotic path planning in terms of selection of state transition, action space dimensionality, and dynamic programming requirements. It is noteworthy that both systems necessitate sequential decision-making under uncertainty, requiring real-time adaptation to environmental variations through iterative policy updates. Such structural homology justifies the algorithmic reformulation of coil array optimization as an intelligent path selection problem within a RL framework.

B. Design of the Proposed Algorithmic Coil Matching Model

1) Define State Space: The design of the state space aims to comprehensively represent the coil matching status within the array system, encompassing key variables that reflect current coil coupling effectiveness. Through the agent's exploration of the environment, i.e., by continuously activating different ETUs, the state space captures the dynamic changes in the system's coupling state within the environment.

Regarding the selection of state variables, directly using mutual inductance as an evaluation indicator to reflect the coupling state of coils lacks intuitiveness and requires precise calculation with its accuracy depending on the accurate construction of the mutual inductance model. Given that mutual inductance is directly proportional to the pick-up voltage of receiving coils based on the *CLC* topology while the pick-up voltage can be directly fed back to the primary side without the need for explicit parameter recognition. Besides, the measurement of dc voltage in secondary sides usually only needs a voltage divider circuit and an analog-to-digital converter, without complicated processing modules. At the same time, the dc voltage is relatively constant, without considering the change of frequency, phase or waveform, and the stability is stronger. Therefore, the state space is constructed by sampling the pick-up voltage of the receiving coils. Moreover, a grid network map is established within the algorithmic framework to digitally twin the exploration process of physical coil array environments. Each grid vertex corresponds to a distinct energy transfer unit, and the state values of these grid vertex are indicated by the picked-up voltages of receivers, which is defined as follows:

$$S_t = \{U_1, U_2, U_3 \cdots U_t\}. \tag{13}$$

2) Selection of Action Space: The action space defines the activation strategy of the array coils, concurrently mapping to the agent's movement direction across grid points within the algorithmic environment. The action design can be categorized into two situations: activating one single energy transfer unit and activating multiple units simultaneously. An analysis is conducted focused on the latter first, exemplified by the simultaneous activation of two energy transfer units.

a) Scenario 1. Simultaneous activation of two random energy transfer units: In this scenario, compared to activating one single ETU at a time, the addition of a second activated unit introduces an extra degree of freedom to the system. This leads to a significant increase in the solution space, making algorithm convergence more challenging. It necessitates more robust algorithmic logic and results in a dimensional explosion of the Q-value table, requiring substantially larger computational storage. Furthermore, theoretically, it increases the demands on the agent's ability to handle abrupt changes and boundary conditions, placing more stringent requirements on the design of the reward function and other algorithmic components.

b) Scenario 2. Simultaneous activation of two adjacent energy transfer units: Based on this scenario, this configuration requires careful consideration of the tradeoff between the number of actions that the agent needs to explore and the switching cost of the array coils. Each action represents the activation of two energy transfer units, leading to increased twice system losses during the search process. Moreover, the larger step size per search iteration can easily cause the algorithm to miss local optima under the specific size limitation within array structure.

Therefore, to address these challenges, the agent's action space is defined as selecting and activating one of the three-phase energy transmission units adjacent to the current ETU in the four directions of horizontal and vertical, which ensures that the agent can explore any possible reconfiguration of the densely packed array coils on the transmitting platform, avoiding both entrapment in local optima and missing the global optimum due to large step sizes. It also simplifies the complexity of the driving topology control and balances the coil activation cost of the array system. Therefore, the action setting is defined as follows:

$$A = \{a_1, a_2, a_3, a_4\} = \{\text{left}, \text{right}, \text{up}, \text{down}\}. \tag{14}$$

Based on this action design, the agent's state position transition equation is defined as follows:

$$St + 1 = \begin{cases} St + (-1, 0), & a_1 = \text{left} \\ St + (1, 0), & a_2 = \text{right} \\ St + (0, 1), & a_3 = \text{up} \\ St + (0, -1), & a_4 = \text{down} \end{cases}. \tag{15}$$

The composition of Q-value table and exploration environment is shown in Fig. 7. The dimensional optimization of the action space critically governs the convergence behavior and computational tractability of Q-learning implementations. Through deliberate cardinality control of permissible actions,

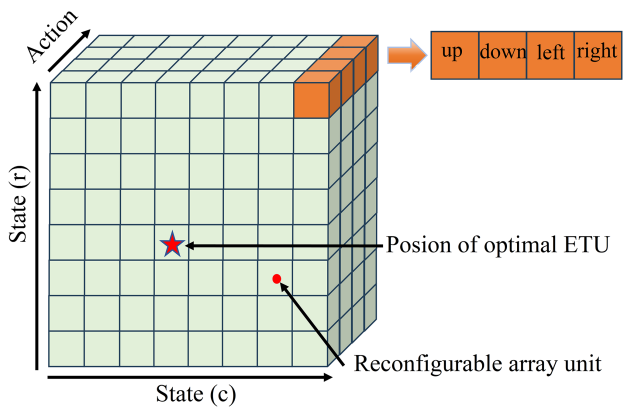


Fig. 7. Schematic diagram of Q-value table composition and exploration environment.

which achieves linear scalability of Q-table dimensions, preventing the curse of dimensionality in value function approximation.

3) Reward Function Setup: The reward function is designed to serve as a mechanism for providing feedback to the agent about the desirability or quality of its actions in a given environment. Aiming at this task model, the reward function framework is mathematically formalized as a multiobjective optimization framework. Specifically, the reward function R_t, is defined as a composite function comprising four components: the immediate reward R_1, the step penalty R_2, the termination reward R_3, and the boundary penalty R_4.

1) *Immediate reward R_1:* This component evaluates the immediate feedback received by the agent after executing an action. It is designed to be proportional to the magnitude of the state variable, specifically the pick-up voltage by the receiving coil. A higher picked-up voltage results in a larger immediate reward, guiding the agent to select ETUs that progressively approach the optimal receiver position. When a power transmission null point occurs, i.e., the pick-voltage is nearly zero at this time, the agent's reward function will significantly decrease, penalizing the agent for executing the corresponding action that leads to the power transmission null point, which prompts the algorithm model to avoid this action and select other optimal paths to reach the designated point of maximum power.
2) *Step penalty R_2:* This component reflects the system cost incurred by activating each energy transfer unit. This penalty encourages the agent to identify the optimal energy transfer unit with the fewest activation steps, thereby minimizing the system's activation cost. In the map of the algorithmic grid, this corresponds to reaching the destination point via the shortest path. Consequently, the penalty increases proportionally with the number of activation steps.
3) *Termination reward R_3:* When the optimal ETU matching state is achieved, i.e., when the agent reaches the location of the optimal energy transfer unit on the algorithmic map, a substantial terminal reward is assigned. This signals to the agent that it has reached the destination.
4) *Boundary penalty R_4:* Theoretically, the array system can expand its array size indefinitely based on the number of drones and charging requirements, leading to a potentially infinite search map for the model. However, due to the limited spatial extent of the magnetic field in near-field coupling, energy transfer is only effectively achieved when energizing energy transfer units in proximity to the receiving coil. Therefore, it is unnecessary and wasteful for the agent to explore and activate energy transfer units outside the region where the picked-up voltage is nearly zero. To address this, the boundary penalty is imposed, restricting the algorithmic model's search map to region where pick-up voltage is nearly zero, preventing excessive exploration beyond this region.

All of these components are designed to ensure rapid model convergence and the generation of a rational algorithmic policy. Therefore, based on the above settings, the calculation formula of the reward function can be expressed as follows:

$$R_t = R_1 - R_2 * \text{step} + R_3 - R_4. \tag{16}$$

After the model training, the specific parameters of the reward function in the algorithmic model are ultimately set as follows:

$$R_t\left(S_t, a_t\right) = \begin{cases} \text{Immediate Reward } R_1 = U_t \\ \text{Step Penalty } R_2 = 0.1 \\ \text{Termination Reward } R_3 = 10 \\ \text{Boundary Penalty } R_4 = 1 \end{cases}. \tag{17}$$

4) Iterative Q-Network Update and Hyperparameter Optimization: The update and iteration of the Q-value table follow the Bellman equation as

$$\begin{aligned} Q_{t+1}\left(s_t, a_t\right) = Q_t\left(s_t, a_t\right) + \alpha[R_{t+1}\left(s_t, s_t, a_t\right) \\ + \gamma \cdot \max_{a \in A} Q_t\left(s_{t+1}, a\right) \\ - Q_t\left(s_t, a_t\right)] \end{aligned} \tag{18}$$

As dictated by the Bellman equation, the Q-table update process inherently selects the maximum Q-value of the subsequent state at each time step for iterative refinement. This implies that Q-learning naturally tends to learn the optimal future path. However, relying solely on a total greedy approach can lead the agent to become trapped in local optima. Therefore, the ε-greedy mechanism in the specific selection of actions is introduced where the agent randomly selects an action with a probability of ε to explore the environment, or selects the action with the highest current Q-value with a probability of 1-ε, thereby achieving a balance between exploration and exploitation.

In the coil matching algorithm model under investigation, there exist three critical hyperparameters: the learning rate (α) and discount factor (γ) within the Q-learning framework, as well as the exploration probability (ε) in the ε-greedy selection strategy that balances exploration and exploitation. To achieve hyperparameter tuning and optimization for these parameters, the design of experiments methodology was employed.

Based on empirically optimized hyperparameter values, this article assigns four distinct levels to each hyperparameter, denoted as: $\alpha \in \{0.05, 0.10, 0.15, 0.20\}$, $\gamma \in \{0.80, 0.85, 0.90, 0.95\}$, and $\epsilon \in \{0.05, 0.10, 0.15, 0.20\}$.This

TABLE I
ORTHOGONAL ARRAYS AND $\bar{\Delta}Q_i$

Experimental Number	Factor Level			$\bar{\Delta}Q_i$
	α	γ	ϵ	
1	1	1	1	0.760962
2	1	2	2	1.405512
3	1	3	3	2.344426
4	1	4	4	5.489999
5	2	1	2	0.683035
6	2	2	1	1.379448
7	2	3	4	1.260179
8	2	4	3	4.707897
9	3	1	3	0.330560
10	3	2	4	1.627957
11	3	3	1	2.427817
12	3	4	2	4.834104
13	4	1	4	0.080474
14	4	2	3	0.367479
15	4	3	2	1.075790
16	4	4	1	4.940370

TABLE II
AVERAGE $\bar{\Delta}Q_i$ AND RANK OF PARAMETERS

Level	α	γ	ϵ
1	2.500	**0.4637**	2.377
2	2.007	1.195	1.999
3	2.305	1.777	**1.937**
4	**1.616**	4.993	2.114
Range	0.884	4.5293	0.44
Rank	**2**	**1**	**3**

configuration yields $4*4*4 = 64$ parameter combinations. To efficiently analyze multifactor and multilevel interactions, an orthogonal array was employed to systematically arrange experiments, ensuring statistically uniform distribution of factor-level combinations while minimizing experimental trials.

Specifically, an $L_{16}(4^3)$ orthogonal array was utilized to generate 16 representative parameter combinations, facilitating evaluation of their differential impacts on the mean variation of Q-values ($\bar{\Delta}Q_i$), the algorithm model is tested on a dataset that includes voltage environments with various UAV landing scenarios. The number of iterations for each experiment is set to 10 000 generations, and each parameter combination is run in parallel three times to minimize experimental errors, with the final results averaged. as shown in Table I.

To further quantify the influence of each parameter level on the mean variation of Q-values ($\bar{\Delta}Q_i$), the average variation for each level of every hyperparameter was calculated and further determined the corresponding range of variation across different levels, which is shown in Table II. This approach enables comparative evaluation of parameter sensitivity by isolating the

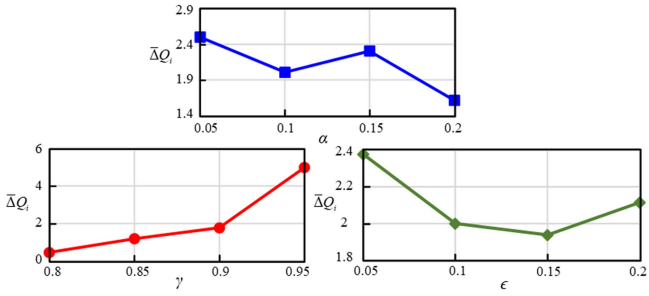


Fig. 8. Trends of factor levels.

marginal effects of individual hyperparameter settings while controlling for multifactor interactions inherent in the experimental design.

Therefore, the parameter level corresponding to the smallest average mean variation of Q-values ($\bar{\Delta}Q_i$) across all parameters was selected as the final optimized hyperparameter value. Simultaneously, the range of variation for each parameter was utilized to quantify its relative influence on algorithm performance, enabling hierarchical ranking of parameter impacts. A larger range value indicates a more pronounced effect of the parameter on the algorithm's behavior. The influence trends of different parameter levels on $\bar{\Delta}Q_i$are further illustrated in Fig. 8, providing visual insights into their respective contributions to algorithmic performance variation.

As demonstrated in Table II and Fig. 8, among the three parameters under investigation, parameter γ exhibits the most pronounced influence on algorithm performance, followed by parameters α and ε in descending order of impact magnitude. Specifically, the discount factor (γ) quantifies the agent's emphasis on future rewards, where smaller γ values indicate a preference for immediate rewards, steering the agent toward strategies that yield instant returns. The learning rate (α) determines the degree to which newly acquired information updates existing Q-values, governing the weighting balance between historical knowledge and new observations during each update iteration. A higher α accelerates Q-value updates, potentially enabling faster convergence to optimal strategies. The exploration factor (ε) regulates the tradeoff between exploration and exploitation by controlling the probability of random action selection. Elevated ε values compel the agent to favor stochastic behavior, thereby facilitating exploration of novel states and actions while mitigating the risk of premature convergence to suboptimal solutions. Based on this comprehensive analysis, the optimal parameter configuration for the multi-UAV coil matching algorithm is determined as $\gamma = 0.8$, $\alpha = 0.2$, and $\varepsilon = 0.15$.

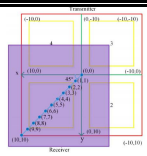
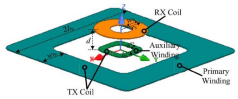
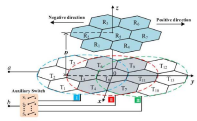
C. Comparison of Different Optimization Algorithms and Coil Matching Methods

Furthermore, to better highlight the superiority of the proposed method, the Sarsa optimization algorithm framework is introduced for comparison, which is also applicable to the task model. Both the Sarsa algorithm and the Q-learning algorithm are value-convergence iterative algorithms, particularly well-suited for the coil matching problem in this task. The two

TABLE III
COMPUTATIONAL COMPLEXITY COMPARISON RESULT

Method	Policy type	convergence Episode	Training time	Average step	Average voltage	Memory usage
Sarsa	On-Policy	14 000	1.7 s	4.91 steps	285.93 V	83.87 MB
Q-learning	Off-Policy	3000	0.87 s	5.03 steps	290.324 V	53.43 MB

TABLE IV
COMPARISON OF POSITIONING AND COIL MATCHING METHODS

Method	 [18]	 [19]	 [21]	 This paper
Coupling mode	Face to face	Face to face	Face to face	Orthogonal
Search scale	2*2	2*2	7*3	19*n
Source for position detection	Mutual inductance	Mutual inductance	Excitation current	Pick-up voltage
System goal realization	Two-dimensional positioning	Two-dimensional matching	Two-dimensional matching	Three-dimensional matching
Whether to support multi-UAVs?	NO	YES	NO	YES
System complexity	⭐⭐	⭐⭐	⭐	⭐⭐⭐
Computational complexity	⭐⭐	⭐	⭐⭐	⭐⭐⭐

The more the number of "⭐", the better the performance.

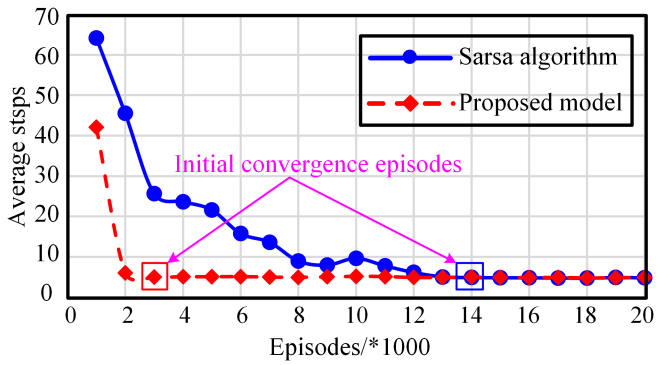


Fig. 9. Convergence curve of the two optimization methods.

optimization methods were compared within the same testing environment to observe their training processes and convergence behaviors. The respective convergence curves are illustrated in Fig. 9. Given that the Sarsa algorithm employs an on-policy strategy, it adopts a relatively conservative learning approach compared to Q-learning, requiring a greater number of iterations and longer training time to reach convergence. However, once convergence is achieved, its strategy becomes more stable. Therefore, the Sarsa algorithm framework takes longer to reach the convergence state and requires correspondingly longer training times compared with Q-learning framework.

Moreover, the computational complexity of the two optimization methods were tested, as shown in Table III. After reaching convergence, the average step size of the Sarsa algorithm is smaller than that of the Q-learning algorithm, indicating that the Sarsa algorithm can relatively reach the optimal energy transmission unit with fewer actions in the coil matching strategy after convergence. Additionally, the average voltage magnitude after each search process executed by the two optimization methods following convergence was recorded to reflect the algorithms' access thresholds to high-voltage state spaces. According to the test results, the Q-learning algorithm exhibits a higher average voltage compared to the Sarsa algorithm, indicating a higher access density to high-voltage states. This reflects that the current agent can approximate the optimal energy transmission unit using high-voltage states. Furthermore, the memory usage of both optimization methods was simultaneously recorded using the Python library Psutil. Therefore, compared to the Sarsa algorithm, the proposed optimization method occupies less memory and has lower computational complexity.

A comparison between the multicoil detection methods currently applied in WPT systems and the proposed method is shown in Table IV. The proposed method demonstrates corresponding advantages in terms of theoretical computational complexity, system cost, and the scale of the array platform. And the proposed system is particularly well-suited for applications involving wireless power transmission for multiple UAVs. As for system goal realization, the model task objective is to ultimately achieve three-dimensional spatial matching and selective activation of coils for efficient energy transfer.

Based on the above analysis, by continuously exploring the voltage distribution under random landing positions and orientations of UAV, after convergence the algorithmic model can quickly and accurately determine the activation sequence of three-phase units in the array based on the current pick-up voltage of the receiver. Taking one situation for an example, the voltage distribution explored by the model is illustrated in Fig. 10.

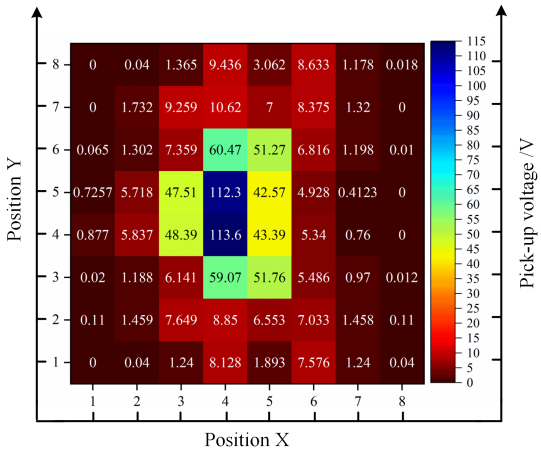


Fig. 10. Pick-up voltage distribution under the exploration environment.

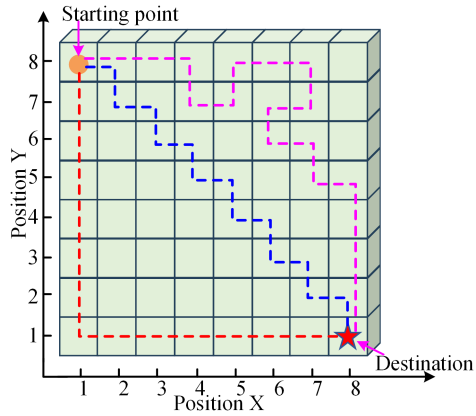


Fig. 11. Schematic diagram of moving distance of agent.

D. Convergence Analysis and Performance Evaluation of the Proposed Algorithm Model

The system's operational framework introduces a novel paradigm distinct from traditional path planning models, as shown in Fig. 11. Due to the random nature of UAV landings on array platforms, the initial activation of energy transmission units is randomized at each exploration onset. This randomization mechanism fundamentally differs from conventional fixed-start point models by simultaneously satisfying task model requirements in the problem context and enhancing algorithmic generalization capabilities through diversified training scenarios. The ultimate optimization objective resides in enabling agents to identify optimal energy transmission units through minimal switching paths that achieve predetermined voltage thresholds.

While Q-table stabilization might indicate potential convergence, such observation lacks both intuitiveness and optimality verification. The inherent randomness of exploration starting points within fixed-voltage-threshold environments precludes direct assessment through simplistic metrics like action counts or traversal distances. Furthermore, evaluation validity faces interference from variable task difficulty levels—proximity between randomly activated units and optimal targets might yield deceptively low action counts without guaranteeing optimal strategies.

To address these complexities in balancing exploration difficulty with policy effectiveness, the proposed algorithmic model first defines the Manhattan distance between two points (X_1, Y_1) and (X_2, Y_2) as

$$d_t = |X_1 - X_2| + |Y_1 - Y_2|. \tag{19}$$

The convergence metric m is defined as the ratio of the number of actions executed by the agent during an iteration to the Manhattan distance between its initial position and the optimal energy transmission unit. When the m value asymptotically approaches unity during model training, it indicates that the agent, starting from a random initial point on the current grid map, achieves an action count equal to the Manhattan distance between the two points. This observation directly validates the attainment of the optimization objective, confirming that the agent has completed the search task via an optimal path. Consequently, the converged Q-value table and derived policy can be reliably extracted for deployment, thereby confirming the model's readiness for stable, optimal decision-making under stochastic initial conditions.

The original input for model optimization is derived from the coupling mutual inductance variation matrix obtained through ANSYS-Maxwell simulations. These matrixes involve randomly varying the landing positions of UAVs and activating different array elements. Subsequently, the coupling mutual inductance matrix is fed into a cosimulation of the designed multicoil driving topology to obtain the magnitude of the pick-up voltage. Furthermore, the corresponding data can be collected through experiments, establishing a mapping relationship between the types of array coil combinations and the pick-up voltage, which serves as the dataset source for model training.

Above all, the design flow and training process of the proposed algorithm model are shown in Fig. 12. The designed algorithm model has the following characteristics.

1) Self-learning and adaptive capabilities to dynamically changing environments with random UAV positions and orientations, eliminating need for prior magnetic field models or fixed rules. A digital twin of electromagnetic field distribution is constructed by coding the variation of the receiving terminal voltage as the grid environment state.
2) With minimal system cost compared to traversal, the integrated optimal coil matching approach significantly enhances the system's tolerance to misalignment and energy transfer performance. The setting of multiobjective reward function balances the matching efficiency and cost of the system.
3) High matching accuracy without relying on complex computations, simplifying the detection process, the physical array coil is transformed into a discrete grid environment, thus realizing intuitive decision-making and model-free calculation.

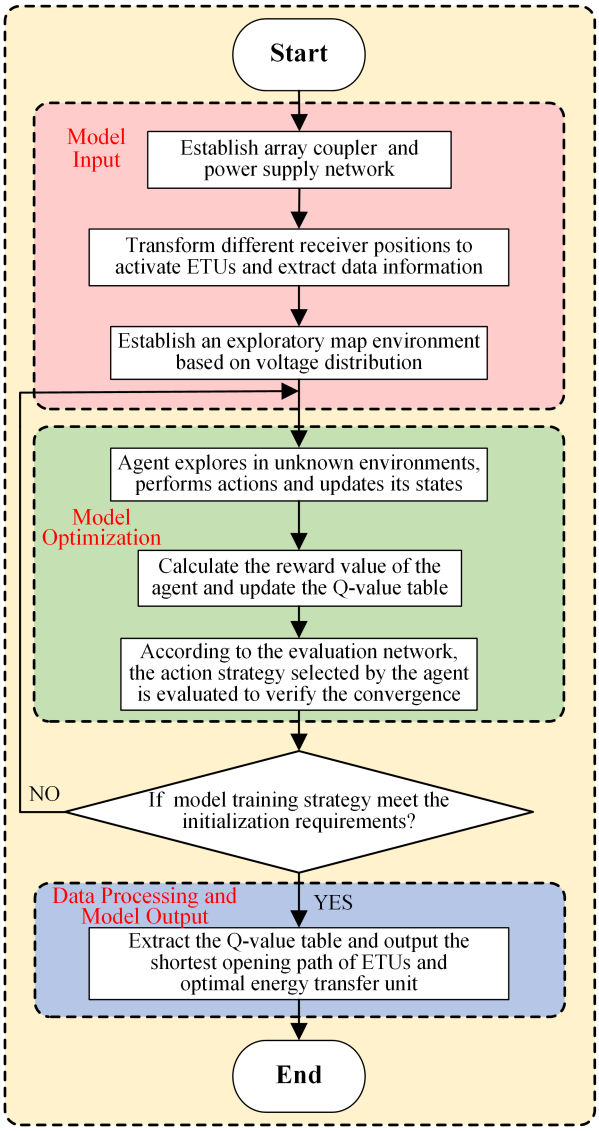


Fig. 12. Flowchart for optimizing the selection of array coils.

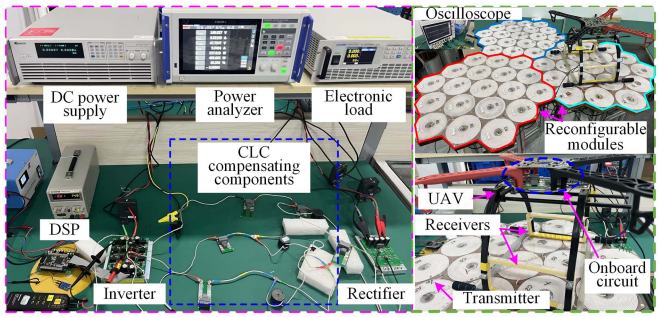


Fig. 13. Experimental prototype.

IV. Experimental Verification

A. Experimental Verification With Power Transfer Ability

1) Experimental Protype Setup: To validate the proposal, an experimental prototype based on the proposed magnetic structure is constructed, as shown in Fig. 13. To mitigate the impact of skin effect and proximity effect during high-frequency alternating current operation, each array transmitting coil was wound using 14 AWG (0.1 × 300) Litz wire. The outer side length of the hexagonal coil was 72 mm, with an outer diameter of 63 mm and an inner diameter of 15 mm. The receiving coil was wound using 15 AWG (0.1 × 200) Litz wire, with each coil having a length and height of 180 and 82 mm, respectively. The transmitting and receiving coils had 19 and 15 turns, respectively. Regarding the selection of the coil quality factor Q, while a higher Q factor can enhance the coupling coefficient and transmission distance, it also narrows the bandwidth of the resonant circuit, making it more susceptible to frequency drift and misalignment. Therefore, a moderate Q factor was chosen to strike a balance between efficiency, transmission distance, and robustness against misalignment. The energy transfer distance was set at 5 mm.

TABLE V
Experimental System Parameters

Parameter	Value	Parameter	Value
U_{DC}	150 V	$L_{1\alpha}$	15 μH
C_1	233.7 nF	$L_{A\alpha}$	40.58 μH
$C_{1\alpha}$	233.7 nF	$C_{A\alpha}$	82 nF
$M_{AS\alpha}$	1.7 μH	$L_{S\alpha}$	60.6 μH
$C_{S\alpha}$	57.85 nF	f	85 kHz
R_O	15 Ω	receiver weight	115 g

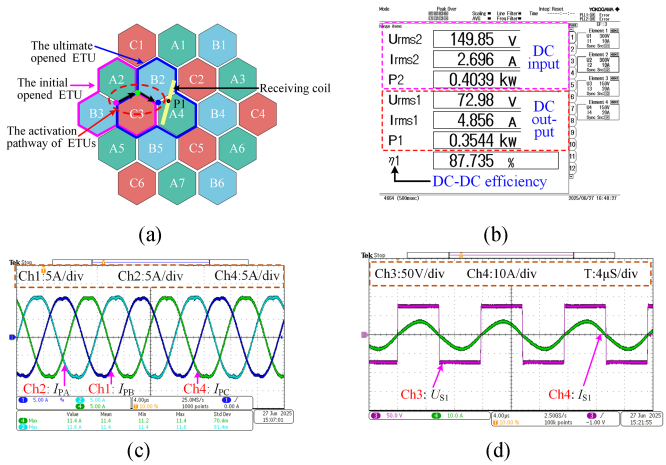


Fig. 14. Power transfer ability test. (a) Coil matching process at P1 position. (b) System output power and efficiency. (c) Waveforms of magnetizing currents of transmission channel α. (d) Waveforms of system output.

The overall circuit parameters are listed in Table V. Experimental tests were conducted based on the array platform and the corresponding circuit design parameters.

The array platform can be further extended freely according to the number of UAVs. The coil matching process when the receiving coil at P1 position as shown in Fig. 14(a), furthermore the transfer power and efficiency of the system are investigated by Power Analyzer, the tested result is shown in Fig. 14(b), which indicates that the output power can delivery 354.4 W with the dc–dc efficiency of 87.74%. This verified that the system enables efficient energy transfer for UAVs. The three-phase

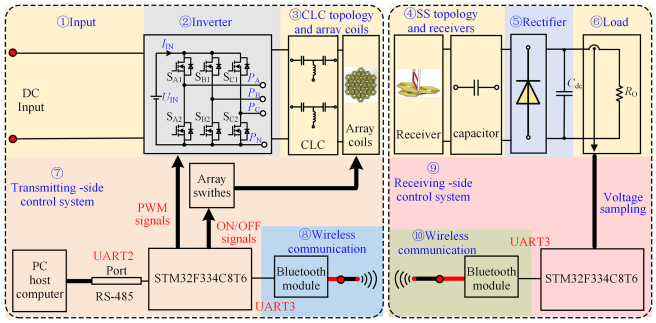


Fig. 15. Schematic diagram of control system.

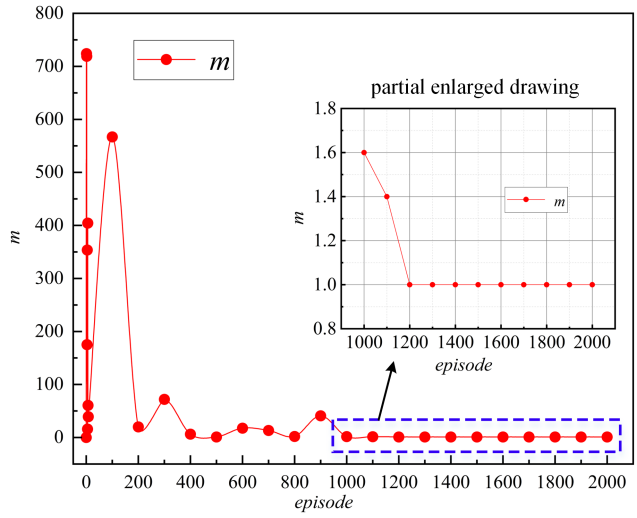


Fig. 16. Experimental convergence test results of the proposed model.

excitation current waveforms in the energy transmission unit were obtained, as depicted in Fig. 14(c). The three-phase excitation currents have equal amplitudes and a phase difference of 120°, meeting the excitation conditions for the coils. Meanwhile, the rectifier input waveform on the secondary side is shown in Fig. 14(d), verifying that the current system is in resonant state.

2) Description of the Feedback Control System: The feedback control system primarily consists of the primary-side and secondary-side control units. The main energy transmission components and the control system are illustrated in Fig. 15.

Specifically, on the receiving side, the feedback signal from pick-up voltage is first sampled through the built-in ADC of the STM32F334C8T6 microcontroller. The sampling frequency is set to 1 Msps, which means that the entire duration from the start of sampling to the completion of a full analog-to-digital conversion is 1 μs. Based on the microcontroller's built-in 12-bit ADC, the digital signal is divided into 4096 levels, resulting in a sampling resolution of 0.8 mV (within $V_{\text{ref}} = 3.3$ V). This resolution provides sufficiently high precision to reflect true voltage variations. Subsequently, the data enters the postprocessing stage. For filtering, a total of ten sampling points are collected, with the two maximum and two minimum values removed. The mean of the remaining six sampling points is then calculated and input to the ADC to complete signal processing. Nonlinear errors of the ADC are eliminated through AD calibration. Finally, communication with the transmitting side is established using a Bluetooth module (BLE5.0) via the UART3 interface, employing a broadcasting method to achieve continuous data transmission.

After the STM32F334C8T6 on the transmitting-side control unit receives Bluetooth data, it establishes communication with the PC host computer via the RS-485 (UART2) communication mode to achieve algorithm integration. The RS-485 communication mode is configured with a baud rate of 115 200 bps, 8 data bits, 1 stop bit, and no parity. The data parsing method specifies that the byte order is high-byte-first, adopting the big-endian mode, where the byte with the larger numerical value is stored at the lower address end. Additionally, the voltage value unit is defined as 10 mV. Furthermore, UART1 is utilized for serial port printing and data debugging. The data frame format adheres to the Modbus-RTU protocol.

As for computational constraints of the STM32F334C8T6 microcontroller, it operates at a clock frequency of 72 MHz, with 12 KB of RAM and 64 KB of Flash memory. Regarding signal processing time delays: on the secondary side, the ADC sampling and conversion time is 1 μs; digital filtering (mean filtering with 6 points) takes 7 μs; and Bluetooth transmission requires 50 ms. On the primary side, Bluetooth reception and parsing takes 1 ms; RS-485 command transmission (16 bytes at 115 200 bps) consumes 1.1 ms. Therefore, the overall system delay is in the tens of milliseconds, primarily attributed to communication overhead in the Bluetooth module. Additionally, it should be noted that the realization of the system design objectives in this article is based on the intelligent matching of UAV receiving coils, considering only the wireless power transfer process. Consequently, the system is configured not to connect to the actual battery load until the coil matching process is completed. The design proposal does not include a post-stage charging controller, which can be implemented based on primary-side or secondary-side regulation to achieve precise constant current/constant voltage (CC/CV) control to meet the charging requirements of actual battery loads. Therefore, for the unmanned array platform, this signal delay is also acceptable.

B. Coil Matching and Position Estimation Verification

To validate the effectiveness of the coil matching algorithm integrated into the platform, it first examined the change in the algorithm's convergence criterion m, during the training process. As illustrated in Fig. 16, m converges and stabilizes at a value of 1 after approximately 1200 iterations.

This indicates that within the current grid map, the agent consistently achieves the shortest path for each exploration, effectively minimizing the number of coil activation sequences and fulfilling the training's optimization objective.

Subsequently, the Q-table was extracted to analyze the algorithmic policy under the current map configuration, as depicted in Fig. 17. Each arrow represents the direction of the next action selected by the agent under the current given voltage state. The results demonstrate that the algorithm correctly directs the agent towards the destination from any grid point, without unnecessary

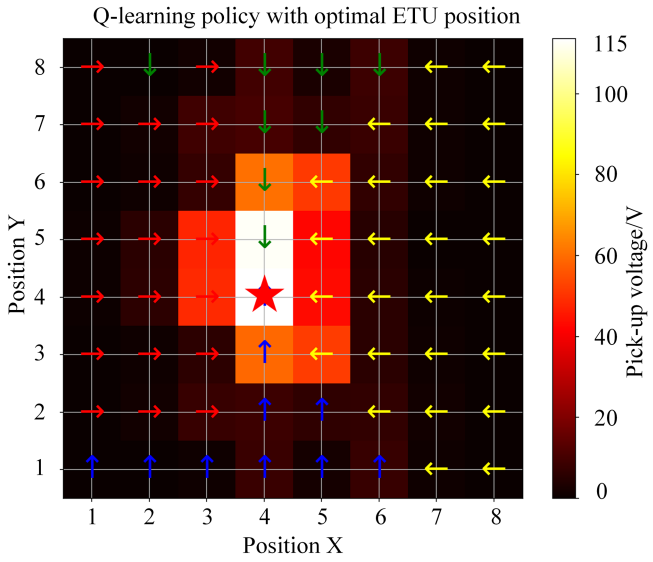


Fig. 17. Output strategy based on the proposed algorithm model.

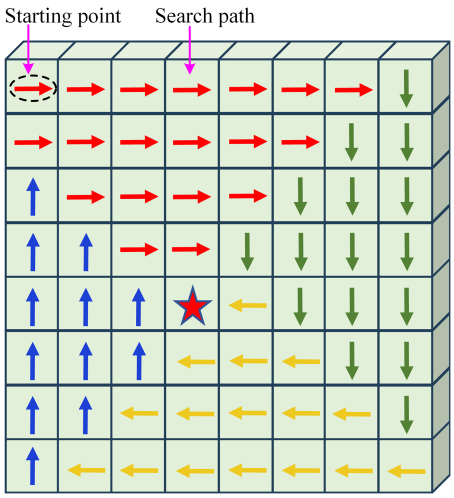


Fig. 18. Traversal search strategy compared to the proposed algorithm model.

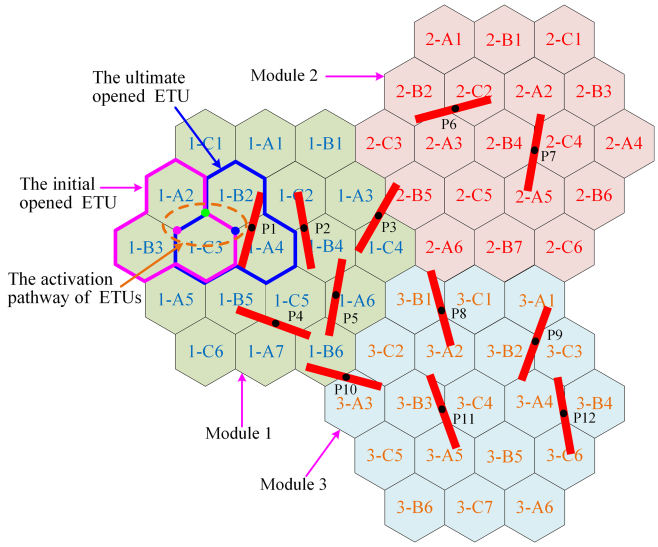


Fig. 19. Schematic diagram of position estimation test.

TABLE VI
COIL MATCHING AND POSITION ESTIMATION

Position	Opening times	Detected ETU	Relative error
P1	3	(1-A4, 1-B2, 1-C3)	1.8 cm
P2	3	(1-A4, 1-B4, 1-C2)	2.4 cm
P3	5	(1-A3, 2-B5, 1-C4)	2.6 cm
P4	2	(1-A7, 1-B5, 1-C5)	2.8 cm
P5	4	(1-A6, 1-B4, 1-C5)	4.2 cm
P6	2	(2-A3, 2-B2, 2-C2)	1.6 cm
P7	3	(2-A5, 2-B4, 2-C4)	1.9 cm
P8	4	(3-A2, 3-B1, 3-C1)	2.1 cm
P9	3	(3-A1, 3-B2, 3-C3)	1.2 cm
P10	2	(3-A3, 1-B6, 3-C2)	3.4 cm
P11	5	(3-A5, 3-B3, 3-C4)	2.8 cm
P12	4	(3-A4, 3-B4, 3-C6)	1.9 cm

or circuitous actions. Furthermore, the algorithm exhibits robust performance at the magnetic field coupling boundaries, with the agent avoiding excessive exploration in these regions. The average number of coil activations for this algorithm is 5. As shown in Fig. 18, compared to an exhaustive array coil scanning algorithm, this represents a 92.2% improvement in searching efficiency with-in the 8∗8 grid map, effectively reducing system losses and enabling intelligent decision-making in the matching process.

Generalization testing was further conducted using 12 randomly distributed UAV landing configurations on the three-module splicing array platform, as shown in Fig. 19. Notably, due to the high symmetry of the array structure between modules, a complete phase sequence cycle can be realized without errors in coil phase sequence excitation, which verifies the modular design and scalability of the array platform. The coil matching test results are presented in Table VI, where the algorithm's output strategy matching results all achieve optimal coil matching, validating the effectiveness of the algorithm model. Additionally, the UAVs still correctly execute action decisions in cross-module areas, further validating the effectiveness of the algorithm model.

Moreover, since the central coordinates of the optimal energy transmission unit after search process can be output, the distance between these coordinates and the center of the UAV's receiving coil can be estimated accordingly. And the algorithm model maintains positional estimation errors below 4.2 cm. It should be emphasized that the primary objective of the proposed model exclusively focuses on optimal transmitting coil matching rather than precise UAV position localization. However, the designed model demonstrates auxiliary positioning capability with maximum errors under 4.2 cm based on the proposed array configuration, thus effectively integrating dual functionalities of coarse localization and energy-efficient coil matching. These results collectively validate the model's robustness in addressing random landing scenarios while maintaining capability of spatial awareness for the rotation angle of the receivers.

V. Conclusion

This article proposes an innovative scheme for an autonomous landing array charging platform for UAVs. The reconfigurable magnetic structure enhances the system's misalignment tolerance. Additionally, a UAV positioning and coil matching model based on RL is proposed. Experimental results demonstrate that the system can achieve power output of 354.4 W with system efficiency of 87.74%. According to the output of the proposed model, coil matching and positioning can be accurately achieved under random landing positions. The average activation count of ultimate strategy is 5 times and the minimum is 2 times with the maximum not exceeding 7 times, representing a 92.2% improvement in detection and matching efficiency compared to the traversal of activating all array coils. This not only effectively reduces system losses but also realizes intelligent decision-making and simplify the workflow in the matching process, which providing robust support for the development of future intelligent UAV operation and maintenance systems.

References

[1] S. Wu, C. Cai, L. Jiang, J. Li, and S. Yang, "Unmanned aerial vehicle wireless charging system with orthogonal magnetic structure and position correction aid device," *IEEE Trans. Power Electron.*, vol. 36, no. 7, pp. 7564–7575, Jul. 2021.

[2] C. Cai et al., "Multi-State voltage balancing of UAV's cell string: A reconfigurable WPT based multiport hybrid charging approach," *IEEE Trans. Ind. Electron.*, vol. 72, no. 1, pp. 266–277, Jan. 2025.

[3] C. Zhang et al., "A strong misalignment-tolerance wireless power transfer system based on dynamic diffusion magnetic field for unmanned aerial vehicle applications," *IEEE Trans. Power Electron.*, vol. 39, no. 11, pp. 14129–14134, Nov. 2024.

[4] Z. Shi, Q. Sun, J. Wang, and Z. Liu, "Free-position dynamic WPT system supporting In-flight concurrent charging for multiple unmanned aerial vehicles," *IEEE Trans. Power Electron.*, vol. 39, no. 12, pp. 15389–15397, Dec. 2024.

[5] Y. Li et al., "A new magnetic coupler with high rotational misalignment tolerance for unmanned aerial vehicles wireless charging," *IEEE Trans. Power Electron.*, vol. 37, no. 11, pp. 12986–12991, Nov. 2022.

[6] Y. Guan, Y. Qiao, J. Mai, Y. Wang, and D. Xu, "Compensation parameter optimization of inductive wireless power transfer system for low stray magnetic field," *IEEE Trans. Power Electron.*, vol. 40, no. 5, pp. 7537–7548, May 2025, doi: 10.1109/TPEL.2025.3527723.

[7] C. Song et al., "EMI reduction methods in wireless power transfer system for drone electrical charger using tightly coupled three-phase resonant magnetic field," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 6839–6849, Sep. 2018.

[8] H. Zhang, Y. Chen, C.-H. Jo, S.-J. Park, and D.-H. Kim, "DC-link and switched capacitor control for varying coupling conditions in inductive power transfer system for unmanned aerial vehicles," *IEEE Trans. Power Electron.*, vol. 36, no. 5, pp. 5108–5120, May 2021.

[9] S. Wu, C. Cai, X. Liu, W. Chai, and S. Yang, "Compact and free-positioning omnidirectional wireless power transfer system for unmanned aerial vehicle charging applications," *IEEE Trans. Power Electron.*, vol. 37, no. 8, pp. 8790–8794, Aug. 2022.

[10] S. Y. R. Hui and W. W. C. Ho, "A new generation of universal contactless battery charging platform for portable Consumer Electronic equipment," *IEEE Trans. Power Electron.*, vol. 20, no. 3, pp. 620–627, May 2005.

[11] X. Tian, K. T. Chau, W. Liu, H. Pang, and C. H. T. Lee, "Maximum power tracking for magnetic field editing-based omnidirectional wireless power transfer," *IEEE Trans. Power Electron.*, vol. 37, no. 10, pp. 12901–12912, Oct. 2022.

[12] X. Zhou, Y. Wang, and L. Yang, "An LCC-LCC compensated WPT system with inherent CC-CV transition function for battery charging applications," *Wireless Power Transfer*, vol. 11, no. 2, Jul. 2024, Art. no. e002.

[13] J. Li, F. Yin, L. Wang, B. Cui, and D. Yang, "Electromagnetic induction position sensor applied to anti-misalignment wireless charging for UAVs," *IEEE Sens. J.*, vol. 20, no. 1, pp. 515–524, Jan. 2020.

[14] J. Li, N. Kou, S. Yu, Z. Ding, and Z. Zhang, "Reconfigurable PCB coil array loaded with PIN diodes for improving transmission efficiency in MCR-WPT systems," *IEEE Antennas Wireless Propag. Papers*, vol. 22, no. 7, pp. 1667–1671, Jul. 2023.

[15] S. Wu, C. Cai, H. Zhang, X. Liu, and W. Chai, "A free-positioning IPT system via reconfigurable coil array transmitter for unmanned aerial vehicle applications," *IEEE Trans. Transp. Electrific.*, vol. 10, no. 4, pp. 8746–8757, Dec. 2024.

[16] Y. Cui et al., "A novel positioning method for UAV in GNSS-denied environments based on mechanical antenna," *IEEE Trans. Ind. Electron.*, vol. 71, no. 10, pp. 13461–13469, Oct. 2024.

[17] H. Wang, F. Zhou, and Q. Wu, "Accurate vision-enabled UAV location using feature-enhanced transformer-driven image matching," *IEEE Trans. Instrum. Meas.*, vol. 73, Apr. 2024, Art. no. 5502511.

[18] Z. Liu et al., "Receiver position identification method of wireless power transfer system based on magnetic integration inductance," *IEEE Trans. Ind. Appl.*, vol. 58, no. 1, pp. 1136–1145, Jan./Feb. 2022.

[19] C. Cai, J. Wang, H. Nie, P. Zhang, Z. Lin, and Y.-G. Zhou, "Effective-configuration WPT systems for drones charging area extension featuring quasi-uniform magnetic coupling," *IEEE Trans. Transp. Electrific.*, vol. 6, no. 3, pp. 920–934, Sep. 2020.

[20] V. K. Srivastava, A. Ahmad, and A. Sharma, "A machine-learning-assisted localization and magnetic field forming for wireless powering of biomedical implant devices," *IEEE Trans. Antennas Propag.*, vol. 72, no. 11, pp. 8590–8599, Nov. 2024.

[21] H. Shen, P. Tan, B. Song, X. Gao, and B. Zhang, "Receiver position estimation method for multitransmitter WPT system based on machine learning," *IEEE Trans. Ind. Appl.*, vol. 58, no. 1, pp. 1231–1241, Jan./Feb. 2022.

[22] B.-G. Choi, E. S. Lee, and Y.-S. Kim, "Optimal structure design of ferromagnetic cores in wireless power transfer by reinforcement learning," *IEEE Access*, vol. 8, pp. 179295–179306, 2020.

[23] X. Liu, J. Chao, C. Rong, Z. Liao, and C. Xia, "Compatibility and performance improvement of the WPT systems based on Q-learning algorithm," *IEEE Trans. Power Electron.*, vol. 39, no. 8, pp. 10582–10593, Aug. 2024.

[24] Y. Yu, H. Zheng, and W. Xu, "Learning and sampling-based informative path planning for AUVs in ocean current fields," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 55, no. 1, pp. 51–62, Jan. 2025.

[25] W. Hu et al., "Multi-UAV coverage path planning: A distributed online cooperation method," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 11727–11740, Sep. 2023.

Shuai Wu (Member, IEEE) was born in Shanxi, China, in 1995. He received the B.S. degree in electrical engineering and automation from Shanxi Agricultural University, Jinzhong, China, in 2017, and the M.S. and Ph.D. degrees in electrical engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2019 and 2023, respectively.

Since 2023, he has been an Associate Professor with the School of New Energy, HIT, Weihai, China. His current research interests include wireless power transfer for unmanned vehicles and implanted biomedical devices.

Jingjie Yang was born in Sichuan, China, in 2000. He received the B.S. degree in mechanical and electrical engineering from North China Electric Power University, Baoding, China, in 2023. He is currently working toward the M.S. degree in electrical engineering in the Academy of New Energy, Harbin Institute of Technology, Weihai, China.

His current research interests include power electronics and wireless power transfer system for unmanned aerial vehicles.

Chunwei Cai (Member, IEEE) was born in Shandong, China, in 1977. He received the B.S. and M.S. degrees in control theory and control engineering from Shan Dong University, Jinan, China, in 2001 and 2004, respectively, and the Ph.D. degree in electrical engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2013.

He has been a Lecturer with HIT, Weihai, China, since 2006 and as a Professor since 2014. His current research interests include wireless power transfer systems, power converters, and inverters.

Wenping Chai (Member, IEEE) was born in Liaoning, China, in 1991. She received the B.S. degree in electrical engineering and automation and also management from the Harbin Institute of Technology, Harbin, China, in 2014, and the Ph.D. degree in electrical engineering from the Department of Electrical and Electronic Engineering, Hanyang University, Seoul Special City, South Korea, in 2020.

She is currently an Assistant Professor of electrical engineering with the School of New Energy, Harbin Institute of Technology at Weihai, Weihai, China. Her research interests include the design, analysis, and optimization of electric machines, and wireless power transfer technologies.

Jinpeng Yu (Senior Member, IEEE) received the B.Sc. degree in automation from Qingdao University, Qingdao, China, in 2002, the M.Sc. degree in system engineering from Shandong University, Jinan, China, in 2006, and the Ph.D. degree in system theory from the Institute of Complexity Science, Qingdao University, Qingdao, China, in 2011.

He is currently a Distinguished Professor with the School of Automation, Qingdao University. He is a recipient of the Shandong Province Taishan Scholar Special Project Fund and Shandong Province Fund for Outstanding Young Scholars. His research interests include electrical energy conversion and motor control, applied nonlinear control and intelligent systems.

Dr. Yu has served as an Associate Editor of several journals, including IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I: REGULAR PAPERS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: EXPRESS BRIEFS, *Information Sciences*, *Journal of The Franklin Institute*, *ISA Transactions*, *Journal of System Science and Complexity*, etc.