Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.077415

ARTICLE

CALoRA: Content-Aware Low-Rank Adaptation for UAV Transfer Learning

Kiseok Kim[#], **Taehoon Yoo**[#], **Sangmin Lee** and **Hwangnam Kim**[*]

School of Electrical Engineering, Korea University, Seoul, Republic of Korea

*Corresponding Author: Hwangnam Kim. Email: hnkim@korea.ac.kr

[#]These authors contributed equally to this work

Received: 09 December 2025; Accepted: 30 January 2026; Published: 09 April 2026

ABSTRACT: Conventional Low-Rank Adaptation (LoRA) constrains weight updates to a static linear low-rank manifold, which is inherently limited when applied to Reinforcement Learning (RL) tasks for Unmanned Aerial Vehicle (UAV) applications. UAVs operate in highly dynamic and nonstationary environments where rapid variations in sensing and state transitions lead to complex, nonlinear input–output relationships. Such environmental complexity cannot be adequately modeled by a static Low-rank approximation, making conventional LoRA approaches insufficient for the high-dimensional dynamics required in UAV applications. To overcome these limitations, we propose an attention-enhanced LoRA that constructs an input-dependent and intrinsically nonlinear adaptation manifold. By integrating a nonstandard attention mechanism into the vanilla LoRA, our method enables the model to dynamically reshape its weight subspace in response to changing environmental conditions. This allows the policy and value networks to capture diverse local patterns as well as global contextual structure during adaptation, ultimately improving robustness under domain shift and nonstationary data distributions. We evaluate the proposed method in UAV adaptation scenario based on the AirSim simulator, where a multi-agent training is conducted with internally collected datasets, including multi-sensor observations and UAV physical state information, and policies are transferred from obstacle-free to cluttered environments. Compared to vanilla LoRA, the proposed method reduces initial reward variance by over 70%, leading to earlier adaptation and more stable generalization, and exhibits richer nonlinear expressive power, allowing the model to accommodate the complex, high-dimensional characteristic of UAV tasks.

KEYWORDS: Low-rank adaptation; transfer learning; reinforcement learning; unmanned aerial vehicle; content-aware

1 Introduction

Unmanned Aerial Vehicles (UAVs) are typically deployed in real-world environments where they must process sensor data in real time, and as a result, they are highly sensitive to distributional shifts induced by environmental conditions [1,2]. As a result, even for the same task, environmental variations such as illumination, terrain, obstacle placement, and signal interference can cause large performance fluctuations [3,4]. These characteristics make it difficult for models to achieve robust generalization. To address this issue, recent studies have attempted to incorporate Transformers or large language models to capture complex environmental context [5], but considering the substantial computational cost and data requirements associated with training such models from scratch, their efficient deployment across diverse UAV applications remains fundamentally limited. In particular, in Reinforcement Learning (RL), unlike supervised learning, the capacity of experience data that an agent can obtain is extremely limited [6,7]. Because the available experiences are highly constrained by the environmental conditions, it takes a long time to accumulate enough data for adequate policy exploration. Ultimately, these constraints lead to a structural

limitation in which policy convergence requires substantially more interactions and longer training time. As a consequence, RL models for UAV applications must recollect sufficient experiential data whenever they encounter a new environment or task, leading to a substantial increase in training time and overall cost [8]. In addition, training models from scratch for each new environment incurs prohibitively high computational costs and fails to provide the real-time adaptability required for autonomous flight. In contrast, reusing shared knowledge enables the agent to avoid redundant exploration and focus on adapting to the distinctive characteristics of new environments, making it the most practical approach for resource-constrained UAV systems. Therefore, transfer learning is essential for effectively transferring the knowledge acquired in previous environments to new ones, thereby reducing the required exploration and improving initial performance [9]. This partially mitigates the distribution mismatch caused by environmental variations and significantly reduces the amount of data and training time required for the policy to converge.

Recently, Parameter-Efficient Fine-Tuning (PEFT) techniques are rapidly spreading along with the rapid development of Large-scale Language Models (LLMs). Whereas traditional transfer learning requires retraining all parameters, PEFTs freeze the base model and update only a small set of lightweight adaptation modules. This significantly reduces training costs while achieving strong adaptability even with limited data. As a leading approach in PEFT, Low-Rank Adaptation (LoRA) has become one of the most widely adopted approaches. By applying low-rank decompositions to correct the linear transformations of large models, LoRA minimizes the number of additional parameters while still providing sufficient expressive power for adapting to new tasks or domains. Due to these advantages, LoRA has rapidly expanded beyond LLMs into various domains including computer vision, speech, and robotics and has emerged as a standard method. As a result, although LoRA was originally designed to reduce the fine-tuning cost of LLMs, there has recently been growing interest in extending it to more general neural network architectures. For example, Khanna et al. [10] shows that when transferring a Vision Transformer (ViT) to a new domain, LoRA-based PEFT enables the model to achieve performance comparable to full fine-tuning while updating less than 10% of the total parameters. Wang et al. [11] showed that applying the LoRA technique to CNN-based edge-AI environments enables even small models to achieve strong adaptation performance with only a small number of additional parameters. Truong et al. [12] showed that LoRA can serve as a general-purpose adaptation module even for non-LLM architectures, such as neural fields.

In this context, this study introduces LoRA to improve adaptation performance in UAV-RL environments, while redesigning its structure to be better tailored to the distinctive dynamical and perceptual characteristics of UAV tasks. Most UAV-based control tasks exhibit highly nonlinear and high-dimensional input–output relationships due to factors such as sensor noise, nonlinear dynamics, and heterogeneous obstacle interactions. As a result, the globally static low-rank linear approximation used in conventional LoRA struggles to capture the diverse patterns induced by environmental variations, ultimately imposing limitations on policy-learning performance. To address these limitations, this study proposes an advanced LoRA architecture that leverages global contextual information to dynamically generate correction matrices conditioned on the input state, thereby enabling the model to learn high-dimensional relationships across diverse patterns. This architecture, unlike conventional approaches that merely add a static low-rank matrix, enables input-dependent weight mixing and nonlinear pattern separation, thereby allowing the model to more effectively approximate the complex functions required in UAV environments. Based on these considerations, the contributions of the proposed LoRA are summarized below:

1. Modular Adaptation Frameworks
   From a functional perspective, the adaptive formulation of the proposed LoRA can be interpreted as an input-dependent mixture of local linear models. The coefficient $\alpha_i(X)$ acts as a gating mechanism that selectively emphasizes the corresponding linear component $A_i B_i^T$ conditioned on the input.

Rather than representing the entire input space with a single linear model, this formulation partitions the space into multiple locally linear regions, allowing the agent to dynamically select an appropriate adaptation pathway.

2. Explainability-Oriented Adaptation Methods
We introduce adaptation mechanisms designed not only to improve performance, but also to enable interpretable analysis of the causal and structural relationships between input variations and model updates. The proposed LoRA's input-dependent gating, mixture of linear experts, and structured update explicitly identify the factors underlying weight updates induced by environmental, domain, and input variations, thereby ensuring semantically aligned adaptation.
3. Multimodal Reasoning under Distribution Shift
The proposed LoRA determines its low-rank weight configuration in an input-dependent manner, thereby enabling multimodal inference that induces distinct adaptation responses depending on the dominant modality or contextual patterns. This design provides a robust adaptation mechanism capable of flexibly handling modality-specific variations and cross-modal interactions under nonstationary environmental conditions.

2 Related Work

2.1 Parameter-Efficient Fine-Tuning

With the rapid advancement of large-scale pre-trained models and the growing adoption of transfer-learning paradigms that reuse such models across various tasks and domains, efficiently adapting large models to new environments has emerged as an important research challenge. As model sizes continue to increase, the traditional full fine-tuning approach, which updates all model parameters, has become prohibitively expensive in terms of computational cost, GPU memory consumption, and per-task model storage requirements. These constraints have further exacerbated the limitations inherent in conventional transfer-learning techniques, intensifying the demand for methods that selectively update only the necessary parameters. Within this context, PEFT has become a central technique in modern transfer learning [13–15]. Early PEFT research proposed various designs aimed at reducing the number of trainable parameters while minimizing performance degradation. Adapter-based methods inserted small bottleneck modules into each transformer layer, enabling adaptation with a limited number of additional parameters [16], while prompt and prefix based methods introduced learnable tokens into the input or hidden representations to adapt to specific tasks while keeping the backbone frozen [17]. Later studies reported that the parameter updates resulting from full fine-tuning often exhibit an inherent low-rank structure, leading to the emergence of LoRA techniques that explicitly model such low-dimensional parameter changes [18]. Although these approaches demonstrated that strong transfer performance can be achieved by training only a small subset of parameters, they still suffer from several limitations, including module dependency, variability in transfer performance across tasks, and structural constraints that restrict the effective utilization of pre-trained representations.

2.2 Generalization of LoRA

LoRA is a representative PEFT method designed to efficiently adapt large pre-trained models to new downstream tasks. By decomposing weight updates into low-rank matrices and training only a small subset of parameters while keeping the backbone frozen, LoRA significantly reduces computational cost while maintaining performance comparable to full fine-tuning. Originally introduced to reduce the fine-tuning cost of LLMs, LoRA has since expanded across various neural architectures and has become a core technique for mitigating constraints in transfer learning, remaining one of the most widely deployed

PEFT approaches [18]. Because of its simplicity and efficiency, LoRA rapidly expanded beyond Natural Language Processing (NLP) to applications in computer vision, multimodal learning, speech processing, and time-series modeling [19,20]. LoRA extensions—such as AdaLoRA [21] with its dynamic rank allocation, La-LoRA [22] with layer-wise adaptive scaling, HydraLoRA [23] with its asymmetric multi-branch design, and Trans-LoRA [24] with data-free transferable fine-tuning—collectively enhance flexibility, domain adaptability, and generalization across diverse tasks. Although these variants strengthen LoRA's adaptability, stability, and domain generalization, they remain constrained by the fundamental assumption of low-rank linear approximation. Most extensions focus on adjusting rank, employing adaptive scaling, or introducing architectural modifications, yet they still struggle to provide sufficient expressivity or environment-dependent adaptability for tasks requiring rich high-dimensional feature interaction or dynamic context modeling.

2.3 Enhancing the Representational Power of LoRA

LoRA has been widely adopted across various domains as an efficient fine-tuning method that decomposes weight updates into low-rank components. However, because LoRA constrains the update space to a static linear low-rank subspace, it inherently limits expressive flexibility and fails to capture nonlinear, input-dependent variations in the model's behavior. To address these limitations, several studies have proposed extensions aimed at expanding LoRA's expressive power [25,26]. For example, SRLoRA [27] dynamically reconstructs low-importance rank directions to enlarge the adaptable subspace, while RaSA [28] introduces a shared rank pool across layers to increase the effective rank available during adaptation. Most of the proposed research approaches rely on additional modules or architectural modifications, which inevitably increase the computational overhead compared to standard LoRA. More importantly, because LoRA maintains the same update mechanism regardless of the input, it is difficult for the model to incorporate information such as interactions between adjacent data or the global context of the data. Research attempting to address this issue is very limited and fails to effectively reflect the contextual meaning and high-dimensional correlations of the input data. These limitations in environmental adaptability are particularly pronounced in UAV control environments, where models must continuously respond to rapid changes in sensory input and dynamic physical conditions. LoRA's restricted flexibility and lack of real-time responsiveness can lead to degradation in control, navigation, and obstacle-avoidance performance when deployed in dynamic real-world environments. These observations underscore the need for adaptive mechanisms that move beyond LoRA's static linear subspace and instead support dynamic, input-sensitive, and nonlinear representational adjustments suitable for UAV and other environment-driven applications.

Additionally, recent studies have explored transfer learning and generalization in UAV applications. Prior work has investigated sim-to-real transfer and domain randomization to improve robustness in UAV navigation and control tasks, often focusing on data augmentation or policy-level adaptation strategies [29]. Other studies have examined transferring RL policies across environments with varying obstacle layouts, sensor noise, or dynamics, highlighting the importance of rapid adaptation under limited interaction data [30]. While these approaches demonstrate the effectiveness of transfer learning in UAV systems, they primarily operate at the level of data, policy initialization, or representation learning. In contrast, relatively little attention has been paid to how parameter-efficient adaptation mechanisms themselves can be redesigned to better reflect the nonlinear and context-dependent nature of UAV environments. Our work addresses this gap by introducing an input-dependent low-rank adaptation mechanism tailored to UAV-RL environment.

3 Design of Content-Aware LoRA

3.1 Problem Statements

According to the conventional LoRA, it is applied by adding a pair of low-intrinsic-rank weight matrices (i.e., a low-rank decomposition) in parallel to each dense layer of the pretrained model. Given a base weight $W \in \mathbb{R}^{d_{\text{in}} \times d_{\text{out}}}$, the learnable residual is defined as

$$\Delta W = AB^{\top}, \quad A \in \mathbb{R}^{d_{\text{in}} \times r}, \quad B \in \mathbb{R}^{d_{\text{out}} \times r},$$

where $r \ll min(d_{\text{in}}, d_{\text{out}})$ is the intrinsic rank of the adaptation. This design drastically reduces the number of trainable parameters, allowing the model to save computation and resource usage during adaptation while still achieving performance close to full fine-tuning. However, such a static linear approximation is fundamentally limited in learning the high-dimensional and nonlinear weight updates required by complex tasks as follows:

$$\Delta W_{\text{true}} \not\approx A\,B^{\top}.$$

Especially, RL downstream tasks for UAV applications characterized by multi-modal inputs, high-dimensional state spaces, and long-range temporal dependencies require input-dependent nonlinear mixing and inductive biases compatible with sequence structure. However, the conventional LoRA is constrained by a highly limited gradient subspace, restricting its ability to represent such complex weight updates. Given a loss function $L(W)$, gradient descent updates the weights in the direction of $\nabla_W L$ and the conventional LoRA updates only the gradients of the two low-rank matrices A and B, as follows:

$$\nabla_A L = \nabla_W L \cdot B^{\top}, \quad \nabla_B L = A^{\top} \cdot \nabla_W L.$$

It means $\nabla_W L$ is constrained to move only in directions that can be expressed as combinations of A and B. In other words, the ΔW is confined to a low-dimensional linear subspace, forcing gradient descent to move only within this narrow plane. These characteristics make low-rank approximation difficult when adapting to nonlinear and high-dimensional tasks. If the task's optimal solution W^* lies outside that subspace, the optimization cannot move toward it, regardless of how long training continues. Even if the gradient attempts to move outside the subspace, the conventional LoRA removes all components in those directions. In another way, increasing r only expands the capacity of the globally static linear subspace $\mathcal{L}_{\text{static}}$, without fundamentally altering the types of mappings it can express, as follows:

$$\mathcal{L}_{\text{static}} = \{\, W_0 + AB^{\top} \mid A, B \in \mathbb{R}^{d \times r} \,\}.$$

3.2 From Static to Input-Dependent Low-Rank Mixtures

In RL for UAV applications, the state inputs typically include LiDAR scans, depth images, and the UAV's velocity and attitude vectors. These modalities collectively encode high-dimensional correlations indicative of complex environmental patterns. Therefore, when a UAV determines an appropriate action in a given situation, it inherently undergoes a nonlinear decision-making process. For instance, 1D LiDAR represents distance measurements across angles as a sequence, where adjacent sequence elements correspond to neighboring directions in physical space. To detect obstacles by analyzing variations in these adjacent distance values, the model must be capable of effectively capturing local spatial patterns. However, static linear approximation methods such as LoRA make it difficult to integrate structural assumptions, such as neighborhood interactions or global context, into the model.

To overcome the expressivity bottleneck of static low-rank corrections, we introduce an adaptive low-rank decomposition. The term *adaptive* denotes that the model employs distinct low-rank projections conditioned on the input pattern, thereby making $\Delta W(x)$ a nonlinear, input-dependent transformation. Reflecting the suggestion, our proposed low-rank decomposition is illustrated in Fig. 1.

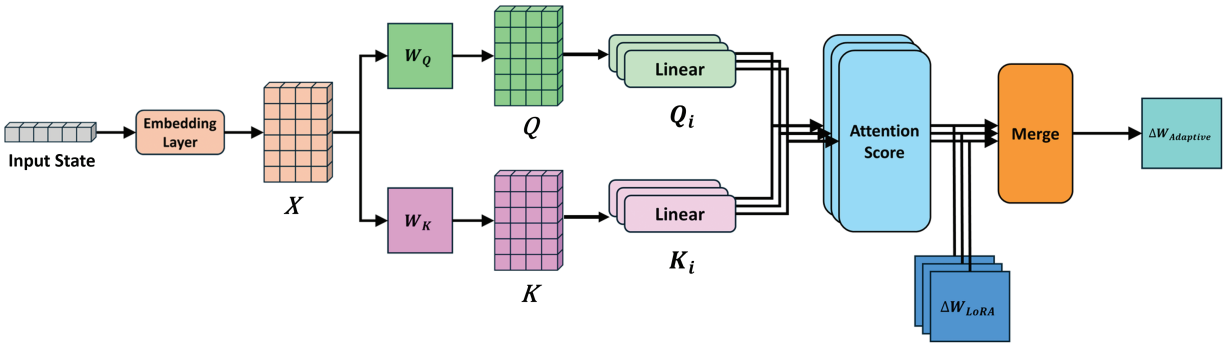


Figure 1: An overview of CALoRA.

Given the embedding sequence $X = [x_1, \ldots, x_N] \in \mathbb{R}^{d_{\text{in}} \times d_{\text{model}}}$, we replace the scaling factor α with a multi-head attention mechanism:

$$Q = XW_Q, \quad K = XW_K,$$
$$\alpha(X) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right).$$

This method allows the model to dynamically recompose ΔW according to the contextual correlation encoded in Q and K. The resulting correction term can be expressed as:

$$\Delta W(X) \approx \sum_{i=1}^{h} \alpha_i(X) AB^\top,$$

where $\alpha_i(X) \in \mathbb{R}^{d_{\text{in}} \times d_{\text{in}}}$ is the input-dependent mixture weight derived from the attention map.

3.3 Adaptive Low-Rank Approximation

The proposed LoRA can be interpreted as an adaptive low-rank decomposition. Each mixture $\alpha_i(X)AB^\top$ forms a rank-r basis while $\alpha_i(X)$ acts as an input-conditioned coefficient that varies per sample. Hence, unlike the static form $\Delta W = AB^\top$, the effective weight update $\Delta W(X)$ spans an input-dependent manifold:

$$\mathcal{M}_{\text{adaptive}} = \left\{ \sum_i \alpha_i(X) AB^\top \;\middle|\; X \in \mathcal{X} \right\},$$

which is significantly richer than the static linear subspace $\mathcal{L}_{\text{static}}$.

Intuitively, while increasing r in conventional LoRA merely enlarge dimensions of $\mathcal{L}_{\text{static}}$, the proposed adaptive variant expands the functional family by introducing a nonlinear dependency of $\alpha_i(X)$ on the input distribution. This nonlinear coupling allows the model to approximate a much broader class of functions without significantly increasing the number of parameters.

3.4 Expressive and Nonlinear Approximation Advantages

From a functional perspective, the proposed LoRA can be described as a input-dependent mixture of local linear approximations:

$$W(X) = (W_0 + \Delta W(X)) = W_0 + \sum_i \alpha_i(X) AB^\top,$$

which introduces two properties.

1. Nonlinear Expressivity
   The softmax attention in $\alpha(X)$ induces explicit nonlinearity, enabling the model to represent piecewise linear or highly nonstationary relationships that static LoRA cannot. In essence, $\alpha(X)$ is not a mere scalar. Rather, it operates as an input-dependent gating mechanism that selectively amplifies or attenuates the contributions of multiple linear submodels. Even with the same number of parameters, $\Delta W(X)$ behaves as a conditional low-rank field, adapting to each input context.
2. Contextual Inductive Bias
   The adaptive mechanism embeds locality and global context simultaneously. Specifically, the multi-head structure generates different weights in parallel and applies distinct weighted combinations to the low-rank bases. As a result, each head can specialize in different sub-patterns (e.g., nearby obstacles, sparse long-range reflections, or noise suppression), facilitating clearer feature separation. This structure-aware inductive bias cannot be achieved simply by enlarging r, which remains purely algebraic and context-agnostic.

Theoretically, the expressive gain can be understood by considering that $\Delta W(X)$ now belongs to a nonlinear manifold parameterized by $\alpha(X)$, whose dimensionality scales with both the rank r and the diversity of input patterns. Thus, the adaptive low-rank decomposition improves functional coverage without linear growth in parameter count, providing a more powerful yet efficient way to model input-dependent variations.

3.5 Design Analysis

The proposed method follows the standard LoRA paradigm by freezing the parameters of the base model and learning only the residual updates ΔW, as shown in Fig. 2. In contrast to conventional static LoRA, however, the proposed approach modulates these low-rank updates in an input-dependent manner via an attention-based weighting mechanism, enabling the effective adaptation subspace to be dynamically reconfigured according to the input context. Algorithm 1 summarizes the complete computational flow. The procedure describes the transition from a static low-rank correction to an input-dependent mixture. It begins by converting the raw input into a sequence representation and applying an embedding projection. Next, the model computes the Query and Key representations that drive the multi-head attention, enabling each head to extract a distinct contextual pattern. Each head subsequently produces an input-dependent coefficient matrix $\alpha_i(X)$, which is then used to modulate a shared low-rank basis $AB^\top$. Finally, the head-wise low-rank corrections are aggregated to form the final adaptation term ΔW, which is injected into the pretrained layer. This process demonstrates that the proposed adaptive structure is not merely an expanded version of conventional LoRA, but a structurally distinct mechanism that integrates attention-based context extraction with low-rank parameterization. Furthermore, by explicitly conditioning ΔW on the input X, it establishes the foundation for the nonlinear and context-aware correction behavior that characterizes the proposed LoRA.

Algorithm 1: CALoRA Inference

Input: The input tensor $\mathbf{x} \in \mathbb{R}^{B \times d_{in}}$
Initialize: Set embedding weight matrix $\mathbf{W}_{\text{embed}} \in \mathbb{R}^{1 \times d_{model}}$, Query weight matrix $\mathbf{W}_Q \in \mathbb{R}^{d_{model} \times d_{model}}$, Key weight matrix $\mathbf{W}_K \in \mathbb{R}^{d_{model} \times d_{model}}$, low-rank weight matrices $\mathbf{A}$, $\mathbf{B}$
Output: The correction matrix $\Delta\mathbf{W}$

Procedure:
Convert a length-N input $\mathbf{x}$ into a $[N, 1]$ sequence $\mathbf{X} \in \mathbb{R}^{B \times d_{in} \times 1}$
Perform embedding on $\mathbf{X}$: $\mathbf{X}_{\text{embed}} \leftarrow \mathbf{X}\mathbf{W}_{\text{embed}}$
Generate $\mathbf{Q}$ and $\mathbf{K}$ by referencing $\mathbf{X}$: $\mathbf{Q} \leftarrow \mathbf{X}_{\text{embed}}\mathbf{W}_Q$, $\mathbf{K} \leftarrow \mathbf{X}_{\text{embed}}\mathbf{W}_K$
Head-wise linear transformation: $\mathbf{Q}_i, \mathbf{K}_i, \quad i = 1, \ldots, h$
for $i = 1$ **to** h **do**
    Compute head-wise attention weight:
    $$\alpha_i(\mathbf{X}) \leftarrow \text{softmax}\left(\frac{\mathbf{Q}_i\mathbf{K}_i^\top}{\sqrt{\mathbf{d}_k}}\right)$$
    Compute head-wise weight distribution:
    $$\Delta\mathbf{W}_i \leftarrow \alpha_i(\mathbf{X})\,\mathbf{A}\mathbf{B}^\top$$
end for
Compute head-wise summation for the correction matrix:
$$\Delta\mathbf{W} \leftarrow \sum_{i=1}^{h} \Delta\mathbf{W}_i$$

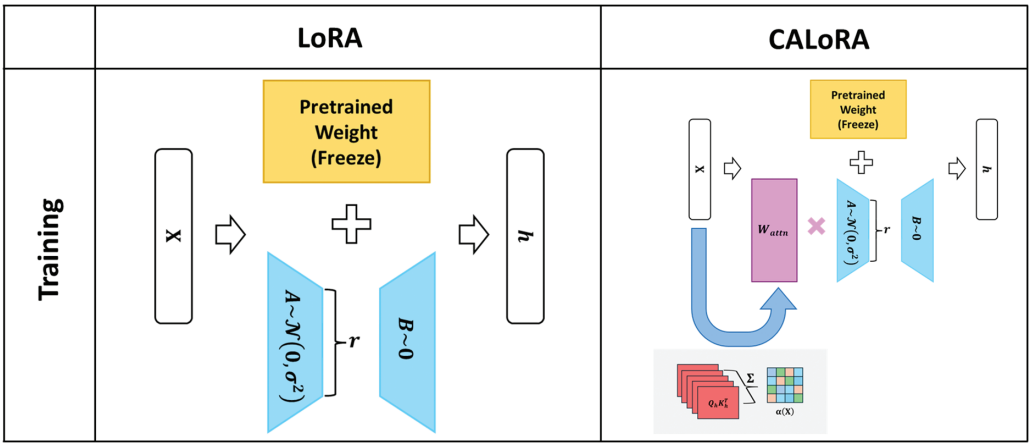


Figure 2: Comparison of residual learning schemes: static LoRA vs. proposed method.

Although the proposed LoRA introduces additional components beyond the basic LoRA, the overall parameter increase remains modest and well-bounded. In the conventional LoRA, the number of trainable parameters added to each layer is given by:

$$|\theta_{\text{LoRA}}| = r(d_{\text{in}} + d_{\text{out}}).$$

In contrast, the proposed LoRA augments this structure with lightweight attention modules that generate input-dependent mixture weights.

Specifically, the query and key projections introduce two additional matrices,

$W_Q, W_K \in \mathbb{R}^{d_{\text{model}} \times d_{\text{model}}}$,

and each head extracts a contextual coefficient $\alpha_i(X)$ via multi-head attention.

The resulting parameter count becomes:

$|\theta_{\text{Adaptive}}| = r(d_{\text{in}} + d_{\text{out}}) + 2d_{\text{model}}^2$.

Although the proposed LoRA introduces a small increase in parameters for each layer compared to basic LoRA, the cost is minimal relative to the functional flexibility and contextual expressiveness it provides.

4 Evaluation

4.1 Experimental Setup

All experiments were conducted in the Microsoft AirSim simulator, which utilizes the latest Unreal Engine–based physics environment. We configured a multi-agent quadrotor setting in which each UAV was equipped with a LiDAR sensor and a depth camera. AirSim's built-in physics engine was used to model UAV dynamics, sensor behavior, and collision events. The simulation environment is an urban-style 3D scene with static obstacles, where obstacle density and layout are randomly varied across episodes. Each episode terminates upon goal arrival, collision, or reaching a maximum step limit, and the reward function combines goal-reaching rewards with collision penalties and step-wise costs. Furthermore, all evaluations were conducted in a multi-UAV setting in which multiple drones were trained concurrently in environments with distinct obstacle distributions, ensuring simultaneous learning under heterogeneous environmental conditions.

For policy learning, we employed a RL approach based on the PPO algorithm. At the beginning of every episode, a new target location was randomly sampled, ensuring that the policy generalizes across diverse flight trajectories. The input state to the PPO algorithm included the UAV's kinematic information, control signals, and sensor observations gathered from both the LiDAR and depth camera. The training procedure consisted of two phases. (1) Base Model/Pre-Training: In the first phase, the UAV learned to perform stable point-to-point navigation in an obstacle-free 3D environment. This pre-training step enabled the model to acquire fundamental flight capabilities such as direction control, altitude maintenance, and goal-directed movement. (2) Transfer Learning: The pre-trained policy was then fine-tuned in a more complex environment containing buildings, walls, and other static structures. This phase focused on improving collision avoidance and enhancing goal-arrival reliability under cluttered conditions.

All experiments were executed on a Windows 10 system configured with AirSim 1.8.1 and Unreal Engine 4.27. The software stack consisted of Python 3.10.18, Stable-Baselines3 2.7.0, PyTorch 2.8.0, Numpy 2.2.6, Cloudpickle 3.1.1, and Gymnasium 1.2.1. Training and simulation were accelerated using an NVIDIA RTX 3090 GPU.

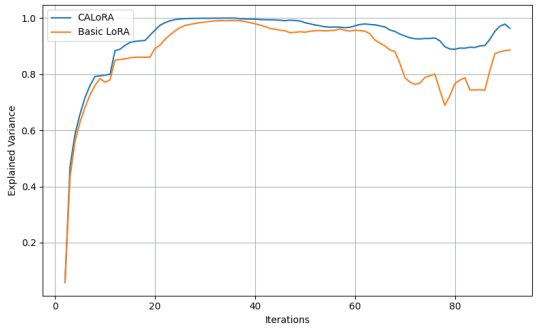
4.2 Adaptation Performance

In this section, we comprehensively evaluate the adaptation performance of the proposed method by comparing three training strategies: Full Fine-Tuning (FT), basic LoRA, and the proposed LoRA. Accordingly, we initialized all models with the same PPO-based pre-trained policy and applied the fine-tuning procedure defined in the previous section uniformly across the three methods, allowing us to quantitatively analyze differences in transfer efficiency caused by their architectural variations.

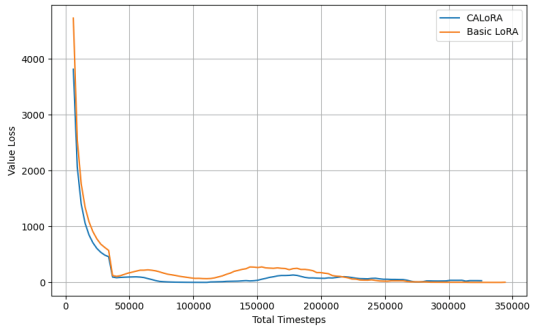
Table 1 summarizes the key metrics measured for the three models in $n\_updates \leq 500$ for a common comparison. First, to evaluate the initial convergence speed, we measured the reward growth rate within the first 30% of timesteps. Full FT exhibited the fastest initial convergence (0.001227), while the proposed LoRA recorded 0.001172, demonstrating nearly the same degree of early adaptation as Full FT. In contrast, Basic LoRA showed the slowest initial convergence among the three models, with 0.000866. This suggests that, despite low-rank updates, the proposed LoRA can effectively preserve early policy-optimization performance. Next, the reward variance in the early stage of transfer learning, which is closely related to training stability, was lowest in the proposed LoRA, reaching only 618. Such a result represents a substantial improvement in stability compared to Full FT (6555) and Basic LoRA (2057). This indicates that the proposed LoRA stabilizes the initial reward dynamics, thereby effectively suppressing drastic policy fluctuations in response to environmental changes. The value loss variance was lowest for Full FT, followed by the proposed LoRA and then Basic LoRA. LoRA-based models structurally approximate the critic's high-dimensional function surface with a low-rank representation, which tends to induce larger oscillations in the loss. However, the proposed LoRA shows an approximately 33% reduction in variance compared to Basic LoRA, indicating a meaningful improvement in critic stability. Finally, in terms of value-function generalization, the proposed LoRA achieved the highest mean Explained Variance (EV) and the lowest standard deviation, demonstrating the strongest value generalization performance among the three models. This result indicates that the proposed LoRA can more effectively capture the underlying transition structures of the environment and task by dynamically reconfiguring its weight composition in response to changes in input patterns. In addition, the proposed LoRA exhibits stable value estimation throughout the entire process, as indicated by Fig. 3a, where its EV remains consistently near 1.0, and Fig. 3b, which shows that after an initial drop, it converges to a lower but more stable baseline. These results suggest that the proposed LoRA structurally has rapid and stable critic convergence, ultimately leading to enhanced policy convergence speed and stability.

Table 1: Comparison of early performance.

Model	Initial Learning Speed	Reward Variance	Value Loss Variance	EV Mean	EV Std
Full Fine-Tuning	0.001227	6555.41	3.73×10^4	0.8891	0.1911
Basic LoRA	0.000866	2057.24	4.18×10^5	0.9196	0.1421
CALoRA	0.001172	618.00	2.82×10^5	0.9493	0.1393



(a) Comparison of Value Function Accuracy



(b) Comparison of Value Function Stability

Figure 3: Stability of value estimation and convergence behavior. (**a**) shows the temporal evolution of explained variance for basic LoRA and CALoRA. (**b**) compares value loss variance across training.

In summary, the proposed LoRA exhibits fast adaptation in the early phase of transfer while maintaining stable generalization throughout training. These observations suggest that the method can operate effectively even in practical transfer-learning scenarios where environmental conditions or data distributions shift frequently.

4.3 Sequence Inductive Bias

The more effectively the model captures diverse sub-patterns and global context from data distributions that vary across situations, the more robustly it can generalize under domain-shift conditions. Furthermore, this indicates that the model can more precisely approximate the complex and nonlinear relationships between inputs and outputs, suggesting the need for an architecture capable of flexibly handling patterns that shift with varying environmental conditions. In this section, we evaluate how the proposed method dynamically reconfigures its weight distribution under varying input conditions. Ultimately, the results visually demonstrate how effectively the model learns sequential structures such as interactions between nearby and distant tokens and the relative and absolute positional relationships. For measurement, we extracted sensor data and UAV state information every 20 steps during the mission defined in Section 4.1. In addition, to more clearly observe the precise feature separation from the input tensor, we applied the proposed method only to the feature extraction layer rather than the entire model.

The results are presented in Fig. 4. The graphs in the second column illustrate 1D LiDAR sequences containing diverse sub-patterns such as near-range spikes, multiple peaks, and gradual rising segments. The corresponding heatmaps on the right visualize the self-attention maps, which depict the reference strength between tokens. From the results, we observe that when abrupt changes occur in certain regions, strong local attention tends to concentrate around the neighboring tokens. In addition, for signals containing multiple features, we observed that attention activation regions emerged simultaneously at several locations. This suggests that the model leverages both local and global structures within the input to recompose its weight mixture. Notably, we observe that even with the same model, entirely different attention structures emerge depending on the input pattern. In some samples, attention concentrates on specific key index regions, whereas in other samples, entirely different positions are selected as the primary reference points. These results support that the proposed method does not remain confined to a single static low-rank approximation, but instead flexibly shifts its attention manifold in response to input variations, enabling it to effectively learn diverse forms of sequential information. In other words, the proposed method reconstructs its weight mixture by selecting different references depending on the input, thereby demonstrating flexible representational capabilities that can accommodate diverse forms of input distributions. This dynamic capability to reconfigure weight updates is expected to provide a foundation for achieving more stable generalization performance, even under environmental changes or domain-shift conditions.

4.4 Analysis of Expressive Power

To examine how the proposed LoRA reshapes the model's representational behavior compared to the basic LoRA, we analyze how sensitively the model responds to different inputs through its parameter gradients, and the structural patterns provide direct insight into the model's expressive subspace, inductive bias, and degree of nonlinearity. By jointly visualizing the Neural Tangent Kernel (NTK) matrix and its spectrum, we can intuitively and quantitatively assess whether the proposed LoRA forms richer, higher-dimensional, and more adaptive representations than basic LoRA.

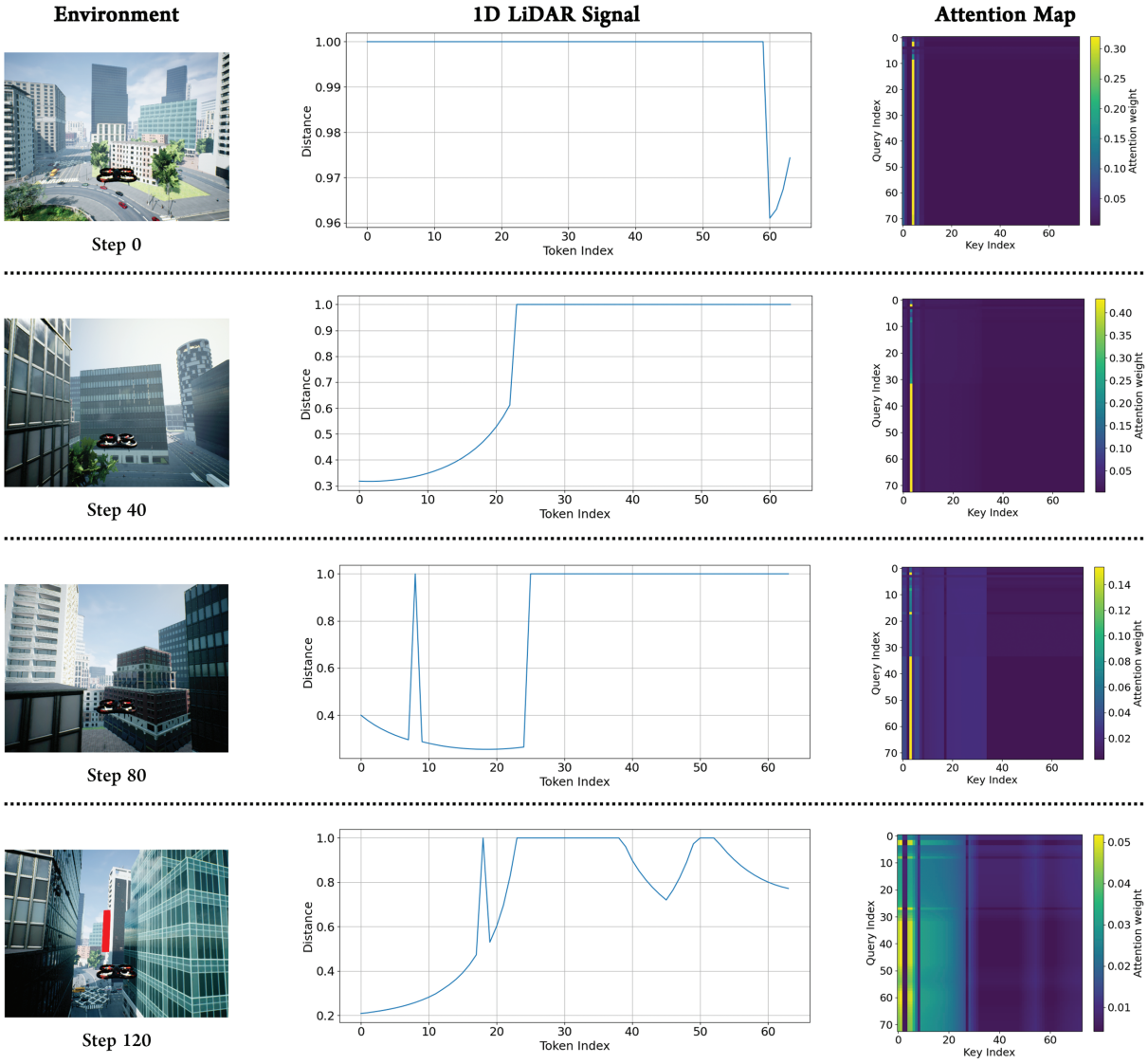
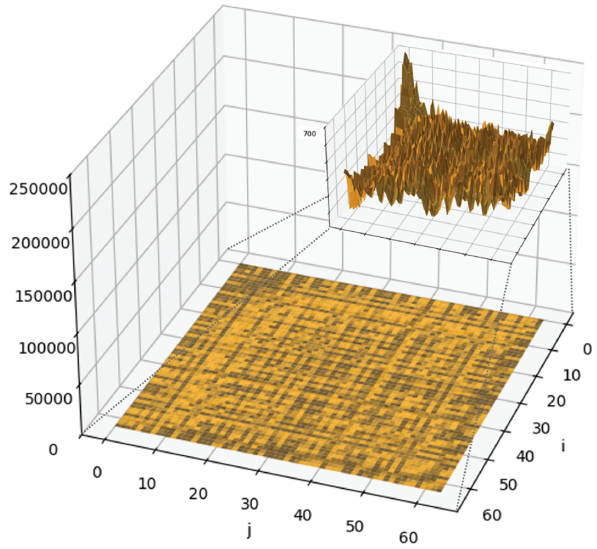


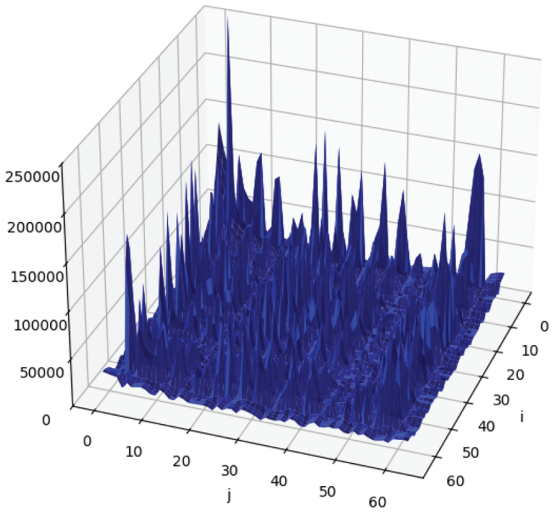
Figure 4: Visualization of AirSim simulation environment (left), 1D LiDAR signal (middle), and Attention Map (right) for several steps.

We first compute the NTK matrices of both models, as shown in Fig. 5a,b. This figure visually illustrates the sensitivity and interaction structure between input tokens, showing how the model's nonlinear expressive subspaces are spatially distributed. The basic LoRA's surface exhibited a homogeneous, nearly isotropic pattern overall, indicating limited expressive subspace expansion in various directions during training. In contrast, the proposed LoRA's surface exhibited anisotropic structure with high intensity to interactions between tokens. This indicates that the model is more sensitive to specific input patterns or local structures, implying the presence of nonlinear feature dependencies. Similar tendencies are also observed in Fig. 5c. The corresponding eigenvalue spectrum indicates the number and complexity of expression directions utilized by each model, showing quantitatively the size of the expressive subspace and the level of nonlinear expressive power. The results show that the eigenvalues of the proposed LoRA not only maintain significantly larger values in higher-order modes, but also slowly decline in the tail even in the eigenvalue lower-order modes. This indicates that the model forms a wider expressive subspace in higher-order modes, implying its ability

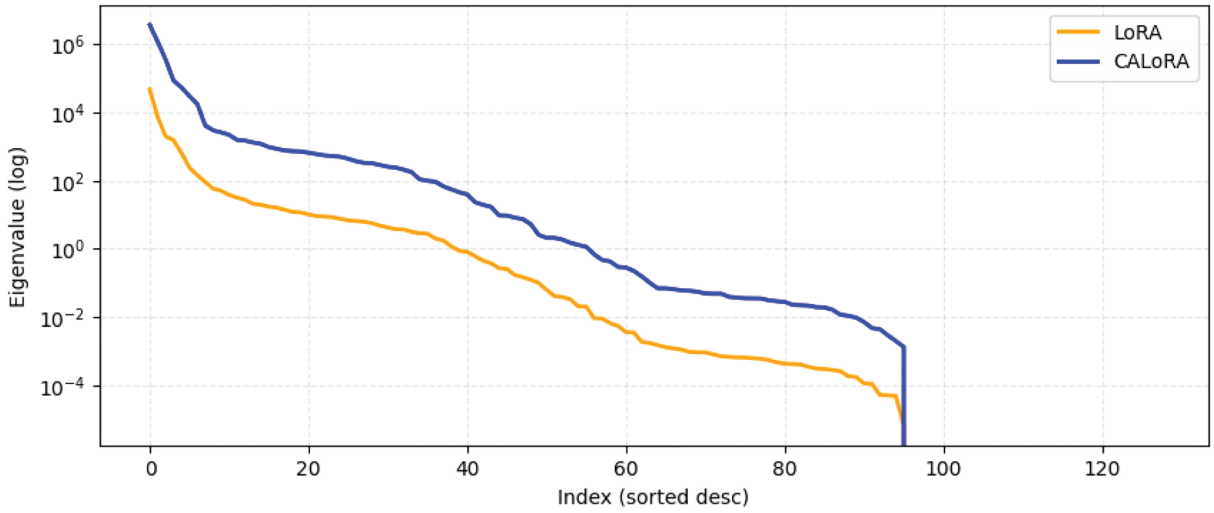
to learn complex functional forms. In summary, the proposed LoRA clearly demonstrates a relatively larger expressive subspace, a nonlinear structure, and a more flexible inductive bias than the basic LoRA, resulting in more effective learning of complex, high-dimensional patterns.



(a) NTK Matrix of LoRA



(b) NTK Matrix of CALoRA



(c) Comparison of NTK Eigenvalue Spectrum

Figure 5: NTK analysis of basic LoRA and CALoRA. (**a**,**b**) show the NTK matrices obtained after adaptation for LoRA and CALoRA, respectively. (**c**) compares the sorted NTK eigenvalue spectra (log scale).

5 Conclusion

We introduce CALoRA, a content-aware low-rank adaptation method designed for UAV applications with nonlinear and high-dimensional representation spaces, which constructs an input-dependent adaptation manifold by integrating a nonstandard attention-based mixing mechanism into conventional LoRA, thereby enabling the model to dynamically reconfigure its weight updates in response to environmental context and overcoming the representational limitations of static LoRA during adaptation.

Through extensive AirSim-based UAV experiments, we demonstrated that CALoRA achieves faster early adaptation, significantly improved training stability, and stronger value function generalization compared to both full fine-tuning and basic LoRA. In particular, CALoRA reduced early-stage reward variance by more than 70% relative to full fine-tuning and improved EV consistency throughout training, indicating more reliable critic convergence under domain shifts. Furthermore, NTK-based analysis revealed that CALoRA forms a wider and more anisotropic expressive subspace, confirming that the proposed method enables richer nonlinear representations without reducing parameter efficiency. These results collectively demonstrate that CALoRA is not merely a structural variant of LoRA, but a practically effective adaptation mechanism tailored to high-dimensional, nonstationary UAV environments. By combining parameter efficiency with adaptive expressivity, CALoRA provides a scalable and robust solution for real-world UAV transfer learning. We expect this method to serve as a PEFT approach for control systems operating under dynamic and nonstationary conditions.

Acknowledgement: The authors would like to express their deepest gratitude to Prof. Hwangnam Kim for his invaluable guidance, insightful feedback, and continuous support throughout the development of this research. The authors also thank the members of the WINE Laboratory at Korea University for their helpful discussions and technical support during the course of this study. In addition, the authors appreciate the administrative assistance provided by the department during the preparation of this manuscript.

Funding Statement: This research was supported by the MSIT (Ministry of Science and ICT), Republic of Korea, under the ITRC (Information Technology Research Center) support program (IITP-2025-RS-2021-II211835) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation). And this work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Climate, Energy & Environment (MCEE) of the Republic of Korea (RS-2022-KP002860).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; methodology, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; software, Kiseok Kim; validation, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; formal analysis, Kiseok Kim and Hwangnam Kim; investigation, Kiseok Kim, Taehoon Yoo and Sangmin Lee; resources, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; data curation, Kiseok Kim, Taehoon Yoo and Sangmin Lee; writing—original draft preparation, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; writing—review and editing, Kiseok Kim, Taehoon Yoo, Sangmin Lee and Hwangnam Kim; visualization, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; supervision, Kiseok Kim Taehoon Yoo and Hwangnam Kim; project administration, Kiseok Kim, Taehoon Yoo and Hwangnam Kim; funding acquisition, Hwangnam Kim. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data used in this study are not owned by the authors and therefore cannot be made publicly available. Access to the data is restricted by the data owner in accordance with project policies, and the authors are not permitted to distribute or disclose the data.

Ethics Approval: Not applicable. This study does not involve human participants or animal subjects.

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

References

1. Shakhatreh H, Sawalmeh AH, Al-Fuqaha A, Dou Z, Almaita E, Khalil I, et al. Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges. IEEE Access. 2019;7:48572–634. doi:10.1109/ACCESS.2019.2909530.
2. Ryan A, Zennaro M, Howell A, Sengupta R, Hedrick JK. An overview of emerging results in cooperative UAV control. In: 2004 43rd IEEE Conference on Decision and Control (CDC). Vol. 1. Piscataway, NJ, USA: IEEE; 2004. p. 602–7.

3. Shakeri R, Al-Garadi MA, Badawy A, Mohamed A, Khattab T, Al-Ali AK, et al. Design challenges of multi-UAV systems in cyber-physical applications: a comprehensive survey and future directions. IEEE Commun Surv Tutor. 2019;21(4):3340–85. doi:10.1109/COMST.2019.2924143.
4. Elmeseiry N, Alshaer N, Ismail T. A detailed survey and future directions of unmanned aerial vehicles (UAVs) with potential applications. Aerospace. 2021;8(12):363. doi:10.3390/aerospace8120363.
5. Kheddar H, Habchi Y, Ghanem MC, Hemis M, Niyato D. Recent advances in transformer and large language models for UAV applications. arXiv:2508.11834. 2025.
6. Wiering MA, Van Otterlo M. Reinforcement learning. Adapt Learn Optim. 2012;12(3):729.
7. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. J Artif Intell Res. 1996;4:237–85.
8. Chen H, Lin Y, Fu M, Yao L, Sheng M. A survey on reinforcement learning methods for UAV systems. ACM Comput Surv. 2025;58(4):1–37. doi:10.1145/3769426.
9. Taylor ME, Stone P. Transfer learning for reinforcement learning domains: a survey. J Mach Learn Res. 2009;10(7):1633–85. doi:10.1145/1273496.1273607.
10. Khanna S, Irgau M, Lobell DB, Ermon S. ExPLoRA: parameter-efficient extended pre-training to adapt vision transformers under domain shifts. arXiv:2406.10973. 2024.
11. Wang Z, Ma H, Zhai J. Low-rank adaptation for edge AI. Sci Rep. 2025;15(1):33109. doi:10.1038/s41598-025-16794-9.
12. Truong A, Mahmoud AH, Konaković Luković M, Solomon J. Low-rank adaptation of neural fields. In: Proceedings of the SIGGRAPH Asia 2025. New York, NY, USA: ACM; 2025. p. 1–12.
13. Houlsby N, Giurgiu A, Jastrzebski S, Morrone B, De Laroussilhe Q, Gesmundo A, et al. Parameter-efficient transfer learning for NLP. In: International Conference on Machine Learning. London, UK: PMLR; 2019. p. 2790–9.
14. Ding N, Qin Y, Yang G, Wei F, Yang Z, Su Y, et al. Parameter-efficient fine-tuning of large-scale pre-trained language models. Nat Mach Intell. 2023;5(3):220–35. doi:10.1038/s42256-023-00626-4.
15. Fu Z, Yang H, So AMC, Lam W, Bing L, Collier N. On the effectiveness of parameter-efficient fine-tuning. In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. Palo Alto, CA, USA: AAAI Press; 2023. p. 12799–807.
16. Pfeiffer J, Kamath A, Rücklé A, Cho K, Gurevych I. Adapterfusion: non-destructive task composition for transfer learning. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Stroudsburg, PA, USA: ACL; 2021. p. 487–503.
17. Liu X, Ji K, Fu Y, Tam W, Du Z, Yang Z, et al. P-tuning: prompt tuning can be comparable to fine-tuning across scales and tasks. In: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA, USA: ACL; 2022. p. 61–8.
18. Hu EJ, Shen Y, Wallis P, Allen-Zhu Z, Li Y, Wang S, et al. LoRA: low-rank adaptation of large language models. arXiv:2106.09685. 2021.
19. Sun Z, Yang H, Liu K, Yin Z, Li Z, Xu W. Recent advances in LoRa: a comprehensive survey. ACM Trans Sens Netw. 2022;18(4):1–44.
20. Sundaram JPS, Du W, Zhao Z. A survey on LoRa networking: research problems, current solutions, and open issues. IEEE Commun Surv Tutor. 2019;22(1):371–88.
21. Zhang Q, Chen M, Bukharin A, Karampatziakis N, He P, Cheng Y, et al. AdaLoRA: adaptive budget allocation for parameter-efficient fine-tuning. arXiv:2303.10512. 2023.
22. Gu J, Yuan J, Cai J, Zhou X, Fan L. La-LoRA: parameter-efficient fine-tuning with layer-wise adaptive low-rank adaptation. Neural Netw. 2025;194(10):108095. doi:10.1016/j.neunet.2025.108095.
23. Tian C, Shi Z, Guo Z, Li L, Xu CZ. HydraLoRA: an asymmetric lora architecture for efficient fine-tuning. Adv Neural Inf Process Syst. 2024;37:9565–84.
24. Wang R, Ghosh S, Cox D, Antognini D, Oliva A, Feris R, et al. Trans-LoRA: towards data-free transferable parameter efficient finetuning. Adv Neural Inf Process Syst. 2024;37:61217–37. doi:10.52202/079017-1957.
25. Li C, Ren Y, Tong S, Siam SI, Zhang M, Wang J, et al. Chirptransformer: versatile lora encoding for low-power wide-area IoT. In: Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services. New York, NY, USA: ACM; 2024. p. 479–91.

26. Meng F, Wang Z, Zhang M. Pissa: principal singular values and singular vectors adaptation of large language models. Adv Neural Inf Process Syst. 2024;37:121038–72.
27. Yang H, Wang L, Hossain MZ. SRLoRA: subspace recomposition in low-rank adaptation via importance-based fusion and reinitialization. arXiv:2505.12433. 2025.
28. He Z, Tu Z, Wang X, Chen X, Wang Z, Xu J, et al. RaSA: rank-sharing low-rank adaptation. arXiv:2503.12576. 2025.
29. Zhao W, Queralta JP, Westerlund T. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI). Piscataway, NJ, USA: IEEE; 2020. p. 737–44.
30. Huda SA, Moh S. Transfer learning algorithms in unmanned aerial vehicle networks: a comprehensive review. In: Proceedings of the 11th International Conference on Smart Media and Applications (SMA 2022). New York, NY, USA: ACM; 2022. p. 9–14.