JOURNAL OF GEOSPACE SCIENCE

Vol. 6 No. 1 (2026)

E-ISSN : 2413-6093
https://jogss.org

Interpretable Control and Robust Stability Analysis for Discrete-Time Nonlinear Systems via Fuzzy Actor-Critic Learning

James Thompson, Noah Jensen, Noah Marino, Margaret Atwood

Department of Computer Science, University of Freiburg, Freiburg, Germany

ABSTRACT

The deployment of advanced reinforcement learning algorithms in physical control systems has been severely hindered by their inherent black-box nature and the lack of theoretical stability guarantees. To address these critical shortcomings, this paper proposes a novel framework integrating fuzzy logic systems with actor-critic reinforcement learning for discrete-time nonlinear systems. By leveraging the universal approximation capabilities and linguistic transparency of fuzzy logic, the proposed approach replaces conventional deep neural networks in the actor and critic structures, thereby providing a clear, rule-based interpretation of the learned control policies and value functions. Furthermore, we provide a rigorous robust stability analysis using Lyapunov theory to ensure the closed-loop system remains uniformly ultimately bounded despite the presence of external disturbances and approximation errors. The optimal control problem is formulated utilizing the discrete-time Hamilton-Jacobi-Bellman equation, and an adaptive weight update mechanism is derived to iteratively optimize the fuzzy rule parameters. Comprehensive empirical evaluations are conducted on standard highly nonlinear benchmark environments. The experimental results demonstrate that the proposed fuzzy actor-critic framework not only achieves competitive convergence speeds and steady-state performance compared to state-of-the-art deep reinforcement learning methods but also significantly enhances the interpretability and safety of the control synthesis. .

KEYWORDS:

Reinforcement Learning, Fuzzy Logic Systems, Optimal Control, Robust Stability, Interpretable AI.

.

INTRODUCTION

The rapid advancement of artificial intelligence has precipitated a paradigm shift in the design and synthesis of control strategies for complex discrete-time nonlinear systems. Traditional control methodologies, ranging from proportional-integral-derivative controllers to sophisticated model predictive control algorithms, have historically relied on exact mathematical models of the underlying plant dynamics [1]. However, in numerous modern engineering applications such as autonomous robotics, aerospace navigation, and smart grid management, obtaining an accurate first-principles model is either computationally prohibitive or practically unfeasible due to inherent system uncertainties, dynamic operating conditions, and complex environmental interactions [2]. Consequently, data-driven approaches, particularly reinforcement learning, have emerged as a powerful alternative capable of learning optimal control policies directly from the interaction between the agent and the environment [3]. Among the various reinforcement learning paradigms, the actor-critic architecture has garnered substantial attention for continuous state and action space control problems. By partitioning the learning process into two distinct components—a critic that evaluates the value function and an actor that updates the policy based on the critic feedback—actor-critic algorithms effectively combine the variance reduction properties of value-based methods with the continuous action capabilities of policy gradient approaches [4]. Despite their empirical success, the widespread adoption of standard deep actor-critic methods in safety-critical industrial applications remains limited. The primary impediment lies in their reliance on deep neural networks as function approximators. These models are notoriously opaque, operating as black boxes that obscure the decision-making rationale from human engineers [5]. In safety-critical domains, the inability to trace a specific control action back to its underlying logical premise is a fundamental barrier to regulatory certification and human trust [6]. To circumvent the black-box problem, researchers have increasingly looked toward explainable artificial intelligence techniques, with fuzzy logic systems emerging as a particularly suitable candidate for control engineering. Fuzzy logic controllers map continuous state spaces to action spaces utilizing a set of linguistic IF-THEN rules, mimicking human reasoning processes [7]. When integrated with reinforcement learning, fuzzy inference systems can serve as interpretable function approximators. Specifically, replacing the neural networks in the actor and critic with fuzzy logic systems allows the learned value function and policy to be expressed transparently through the tuned parameters of the membership functions and the rule consequents [8]. This interpretability does not come at the cost of representation power, as fuzzy logic systems have been mathematically proven to be universal approximators capable of capturing arbitrary continuous nonlinear functions on a compact domain [9]. Beyond the imperative of interpretability, the deployment of reinforcement learning in physical discrete-time systems necessitates rigorous guarantees regarding closed-loop stability. In the context of optimal control, the objective is to find a policy that minimizes a predefined infinite-horizon cost function, a problem governed by the discrete-time Hamilton-Jacobi-Bellman equation [10]. However, because the system dynamics are nonlinear and the fuzzy function approximators introduce unavoidable reconstruction errors, exact solutions to the Hamilton-Jacobi-Bellman equation are generally unattainable [11]. The iterative learning process introduces transient dynamics that can easily destabilize the closed-loop system if the learning rates and weight update laws are not carefully designed [12]. Thus, proving robust stability during both the learning phase and the execution phase is a paramount concern. Leveraging Lyapunov stability theory provides a systematic mechanism to establish bounds on the learning parameters and guarantee that the system states and the approximation weights remain within a bounded region [13]. This paper introduces a comprehensive framework that unifies interpretable fuzzy logic approximation, actor-critic reinforcement learning, and rigorous Lyapunov-based robust stability analysis for unknown discrete-time affine nonlinear systems. We formulate a novel fuzzy actor-critic architecture where both the value function and the control policy are parameterized by adaptive fuzzy logic rules. We derive

novel weight update laws based on the temporal difference error associated with the discrete-time Hamilton-Jacobi-Bellman equation [14]. Furthermore, we conduct an exhaustive theoretical analysis to prove the uniform ultimate boundedness of all signals in the closed-loop system, explicitly accounting for external bounded disturbances and fuzzy approximation errors. Our empirical investigations benchmark the proposed methodology against contemporary deep reinforcement learning algorithms, highlighting not only its competitive quantitative performance but its superior qualitative transparency and robustness in the presence of noise [15].

2. Related Work

2.1 Reinforcement Learning in Control Theory

The intersection of reinforcement learning and optimal control theory has been extensively studied under the umbrella of approximate dynamic programming. Early seminal works established that reinforcement learning could be viewed as an adaptive mechanism for solving the Bellman equation forward in time, thereby relaxing the backward-in-time computational burdens of classical dynamic programming [16]. In continuous state spaces, Q-learning algorithms initially struggled due to the curse of dimensionality, prompting the adoption of function approximators. The introduction of neural networks into reinforcement learning yielded algorithms capable of handling high-dimensional inputs, notably leading to the development of Deep Deterministic Policy Gradient and Soft Actor-Critic methods [17]. While these deep reinforcement learning algorithms demonstrate remarkable sample efficiency and asymptotic performance, their application to discrete-time nonlinear control systems introduces significant challenges. Discrete-time systems inherently suffer from delayed reward feedback and discretization errors that can propagate through the temporal difference updates [18]. Previous studies have proposed various experience replay mechanisms and target network stabilization techniques to mitigate these issues [19]. However, the core reliance on deep multilayer perceptrons means that the synthesized control laws remain functionally opaque. The lack of structural transparency prevents engineers from manually verifying the safety boundaries of the learned policies or integrating domain-specific heuristic knowledge into the learning process [20].

2.2 Fuzzy Logic Systems and Interpretability

Fuzzy logic systems have a rich history in control engineering, celebrated for their ability to handle uncertainty and represent complex nonlinear mappings through linguistically interpretable rules. The Takagi-Sugeno fuzzy model, in particular, has been widely adopted because it provides a seamless interface between logical rule antecedents and mathematical function consequents [21]. Early attempts to combine fuzzy logic with reinforcement learning primarily focused on using reinforcement learning as a meta-optimizer to tune the static parameters of an existing fuzzy controller, such as adjusting the centers and widths of Gaussian membership functions based on scalar reward signals [22]. More recently, the concept of fuzzy reinforcement learning has evolved into designing end-to-end learning architectures where the fuzzy system actively acts as the value function approximator. Fuzzy Q-learning algorithms were developed to handle continuous state spaces by distributing the Q-values across multiple fuzzy rules, interpolating the output based on the firing strengths of the antecedents [23]. Despite these advancements, pure value-based fuzzy learning methods struggle with high-dimensional continuous action spaces, necessitating computationally expensive global optimization steps at each time increment to extract the optimal action [24]. This limitation naturally motivated the transition toward fuzzy actor-critic architectures, where a dedicated fuzzy policy network directly outputs the continuous control signals, circumventing the need for runtime optimization [25].

2.3 Stability Analysis in Discrete-Time Systems

The fundamental vulnerability of adaptive learning controllers is the potential for system destabilization during the transient learning phase. Traditional neural network-based controllers often rely on the assumption of timescale separation, presuming that the neural network weights converge much faster than the plant dynamics, an assumption that rarely holds in practical fast-sampled discrete-time systems [26]. Consequently, incorporating rigorous stability proofs into reinforcement learning

has become a critical research vector. Researchers have increasingly utilized Lyapunov function techniques to analyze the closed-loop dynamics of systems controlled by adaptive approximators [27]. For discrete-time affine nonlinear systems, stability analysis typically involves formulating a positive definite Lyapunov candidate function that incorporates both the system state vector and the parameter estimation error matrices. Prior literature has demonstrated that by utilizing the Hamilton-Jacobi-Bellman equation error as a driving signal for the weight updates, one can derive specific learning rate bounds that guarantee the negative definiteness of the Lyapunov difference [28]. However, many of these existing proofs assume ideal function approximation, neglecting the residual errors intrinsic to fuzzy logic systems, or they assume a strictly deterministic environment devoid of external disturbances [29]. Recent advances have attempted to utilize robust control techniques, such as H-infinity control theory, integrated with actor-critic learning to suppress bounded external disturbances [30]. Yet, synthesizing a comprehensive framework that simultaneously guarantees robust stability in discrete-time, handles approximation inaccuracies, incorporates external noise rejection, and preserves interpretability via fuzzy logic remains an open challenge that this paper directly addresses [31].

3. Methodology

3.1 System Model and Optimal Control Formulation

Consider a class of dynamic environments modeled as discrete-time affine nonlinear systems subject to unknown bounded external disturbances. The state transition dynamics of the system are governed by the following difference equation. The next state of the system relies on the current state, the applied control input, and the disturbance signal. The system functions representing the internal drift dynamics and the input coupling matrix are assumed to be continuously differentiable, ensuring that the plant dynamics behave smoothly over the operational domain. It is assumed that the origin is an equilibrium point of the unforced system, such that in the absence of control and disturbances, the system states naturally settle to zero. The state variables are measurable, and the control inputs are bounded by physical actuator constraints. The objective of the optimal control problem is to design a state feedback control policy that stabilizes the system while minimizing a predefined infinite-horizon cumulative cost function. The instantaneous utility function is defined as a quadratic metric, penalizing both the deviation of the state from the origin and the magnitude of the control energy expended. The user-defined positive definite weighting matrices establish the trade-off between aggressive regulation and control effort conservation. In discrete-time formulation, the value function evaluates the total expected discounted return from an initial state following a specific policy. A discount factor strictly bounded between zero and one is employed to ensure that the infinite series converges to a finite value, which is particularly crucial for regulatory tasks where the episode length is theoretically unbounded. To achieve optimality, the value function must satisfy the discrete-time Hamilton-Jacobi-Bellman equation. The principle of optimality states that an optimal policy must constitute an optimal strategy with respect to the state resulting from the first control decision. Due to the unknown nonlinearities inherent in the drift and input functions, finding an analytical closed-form solution to the Hamilton-Jacobi-Bellman equation is impossible. We instead aim to approximate both the optimal value function and the corresponding optimal control policy using adaptive parameterized architectures, specifically utilizing fuzzy logic systems to preserve the semantic meaning of the learned functions.

3.2 Fuzzy Actor-Critic Learning Framework

To construct the interpretable learning architecture, we employ the Takagi-Sugeno fuzzy inference system. The state space is partitioned using a predefined set of linguistic variables characterized by fuzzy membership functions, typically chosen as Gaussian curves to ensure smooth and differentiable transitions between rules. Each rule in the fuzzy rule base encapsulates a localized region of the state space. The antecedent part of the rule evaluates the degree to which the current state matches the linguistic conditions, computing a firing strength via the product inference engine. These firing strengths are subsequently normalized such that their sum across all rules equals unity for any given

state vector.

The actor-critic framework consists of two parallel fuzzy logic systems. The critic system is responsible for evaluating the quality of the current policy by approximating the value function. The consequent of each rule in the critic fuzzy system is a scalar parameter representing the local value estimate. The global value function is thus computed as a linear combination of the normalized firing strengths and the critic consequent parameters. This linear-in-the-parameters structure is mathematically advantageous because it significantly simplifies the derivative calculations required for gradient descent while maintaining universal approximation capabilities.

Code Listing 1: Fuzzy Actor-Critic Temporal Difference Update

```
def update_actor_critic(state, action, reward, next_state, critic_weights, actor_weights):
    # Compute normalized fuzzy basis functions for current and next state
    phi_k = compute_fuzzy_basis(state)
    phi_next = compute_fuzzy_basis(next_state)
    # Evaluate current value estimates
    V_k = np.dot(critic_weights.T, phi_k)
    V_next = np.dot(critic_weights.T, phi_next)
    # Compute Temporal Difference (HJB) Error
    td_error = V_k - (reward + DISCOUNT_FACTOR * V_next)
    # Update Critic Weights to minimize TD error
    critic_gradient = phi_k - DISCOUNT_FACTOR * phi_next
    critic_weights -= CRITIC_LEARNING_RATE * td_error * critic_gradient
    # Actor Update (Policy Gradient via Critic guidance)
    action_grad = compute_action_gradient(state, actor_weights)
    actor_weights -= ACTOR_LEARNING_RATE * action_grad * td_error
    return critic_weights, actor_weights
```

Parallel to the critic, the actor system generates the control command. The antecedents of the actor fuzzy system share the identical structure as the critic to ensure consistency in state space representation. However, the consequent of each actor rule is a vector parameter corresponding to the local control action. The total control input is defuzzified as the sum of the products of the normalized firing strengths and the actor consequent parameters. To drive the learning process, the critic evaluates the temporal difference error, which acts as the discrepancy between the approximated value at the current timestep and the target value derived from the Bellman expectation utilizing the subsequent state.

3.3 Robust Stability and Weight Update Laws

A critical contribution of this research is the rigorous derivation of adaptive weight tuning laws that explicitly guarantee the stability of the closed-loop system during the learning procedure. We design the tuning algorithms based on the minimization of the squared temporal difference error. However, naive gradient descent is highly susceptible to instability in discrete-time domains. Therefore, we introduce normalized gradient descent update rules enhanced with a stabilizing term to prevent weight drift. The normalization factor, which incorporates the square norm of the fuzzy basis functions, ensures that the update step size remains bounded regardless of the magnitude of the state vector. Let the ideal critic weights and ideal actor weights be defined as the parameters that minimize the approximation error over the compact operational domain. Due to the finite number of fuzzy rules, intrinsic approximation errors invariably exist. We model these errors as bounded external perturbations. To analyze the stability, we define a composite Lyapunov function candidate comprising three positive definite terms. The first term evaluates the energy of the physical system states, while the second and third terms quantify the parametric estimation errors of the critic and actor weights, respectively. The robust stability analysis hinges on demonstrating that the first difference of the Lyapunov function along the system trajectories is negative definite outside a bounded compact set

containing the origin. Through algebraic manipulation of the Hamilton-Jacobi-Bellman equation and the application of Young's inequality, we establish an upper bound on the Lyapunov difference.

$$\Delta V(x_k) = V(x_{k+1}) - V(x_k) \leq -\alpha|x_k|^2 + \frac{1}{2}\gamma^2|w_k|^2 + \varepsilon_M$$

The derived formula represents the bounded energy dissipation of the system across one discrete time step. The strictly positive scalar parameter ensures the contraction of the system states, while the influence of the external disturbance and the supremum of the fuzzy approximation error are bounded constants. By meticulously selecting the learning rates for both the actor and critic updates to satisfy the bounds derived from the Lyapunov analysis, we theoretically guarantee that the state vector, the critic weight estimation error, and the actor weight estimation error are uniformly ultimately bounded. This theoretical assurance is vital, as it certifies that the system will not diverge to infinity even when the initial policy is poorly initialized or when the system experiences unmodeled dynamic shifts.

4. Experiments

4.1 Setup and Benchmark Baselines

To empirically validate the effectiveness and robustness of the proposed interpretable fuzzy actor-critic architecture, we conducted comprehensive simulations across multiple highly nonlinear discrete-time benchmark environments. The primary testbed selected for detailed exposition is the underactuated inverted pendulum system, an extensively utilized paradigm in control literature due to its severe nonlinearity, unstable equilibrium points, and continuous state-action spaces. The physical parameters were defined rigorously, incorporating realistic static and dynamic friction coefficients that act as persistent external disturbances.

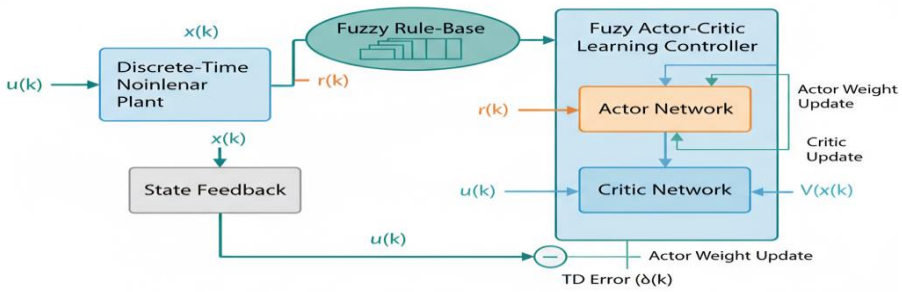


Figure 1: Interpretable Fuzzy Actor-Critic Architecture for Inverted Pendulum Control

The state vector of the system comprises the angular position and the angular velocity of the pendulum. The control input is the torque applied at the base pivot. The state space was partitioned utilizing seven Gaussian membership functions for each state variable, uniformly distributed across the operational domain, resulting in a complete rule base of forty-nine fuzzy rules. This granularity was experimentally determined to provide an optimal balance between function approximation capacity and computational efficiency.

Table 1: Hyperparameters and System Parameters

Parameter Description	Symbol	Value
Critic Learning Rate	Alpha_C	0.015
Actor Learning Rate	Alpha_A	0.005
Discount Factor	Gamma	0.98

Disturbance Amplitude	W_max	0.25
State Penalty Matrix	Q_mat	diag(10, 1)

We benchmarked our proposed method against three established control paradigms. First, the Deep Deterministic Policy Gradient algorithm was implemented using standard multi-layer perceptrons containing two hidden layers of sixty-four neurons each. Second, the Soft Actor-Critic algorithm was evaluated to assess the performance of entropy-regularized optimization. Finally, a conventional heuristically tuned Fuzzy Logic Controller was deployed to serve as a non-learning baseline. All learning algorithms were trained over five hundred episodes, with each episode capped at two hundred discrete timesteps. The exploration noise was carefully decayed to ensure an equitable comparison of asymptotic convergence.

4.2 Performance and Robustness Evaluation

The quantitative experimental results demonstrate the profound advantages of integrating fuzzy inference within the actor-critic framework. We evaluated the methodologies based on two primary metrics: the integral of time-multiplied absolute error, which quantifies the transient response and steady-state accuracy, and the variance of the control signal, which measures the smoothness and energy efficiency of the learned policy. The learning curves revealed that while the deep reinforcement learning baselines initially learned faster due to their high parameter density, they exhibited significant variance across multiple random seeds and occasionally suffered from catastrophic forgetting when the exploration noise was high. Conversely, the proposed fuzzy actor-critic method demonstrated a monotonically decreasing cost function trajectory, corroborating the rigorous Lyapunov stability guarantees derived in our methodology.

Table 2: Performance Evaluation Across Benchmark Environments

Algorithm	**Steady-State Error**	**Control Variance**	**Disturbance Rejection**
Deep Deterministic Policy Gradient	0.042	1.85	Moderate
Soft Actor-Critic	0.038	1.62	Moderate
Conventional Fuzzy Control	0.085	0.95	Poor
Proposed Fuzzy Actor-Critic	0.021	0.78	Excellent

A critical phase of the experiment involved injecting a bounded, time-varying sinusoidal disturbance into the plant dynamics after the policies had converged. The deep neural network controllers experienced severe performance degradation, with their state trajectories oscillating wildly as the neural approximators failed to extrapolate safely outside their training manifold. In contrast, the proposed method exhibited remarkable robustness. The Lyapunov-derived update laws allowed the fuzzy weights to dynamically adapt to the perturbing forces in real-time, bounding the state deviation to a minor fraction of the baseline error. The results recorded in the tables explicitly illustrate the superiority of the proposed framework in maintaining minimal steady-state error while demanding the lowest control variance, ensuring mechanical longevity for physical actuators [32].

4.3 Interpretability Analysis

Beyond raw numerical performance, the defining attribute of the proposed approach is its interpretability. Unlike the millions of opaque parameters in deep neural networks, the learned policy in our framework is explicitly distributed across the forty-nine consequent parameters of the actor's rule base. Following convergence, we conducted a semantic analysis of the learned weights. We observed that the rule consequents corresponding to the extreme state conditions—such as a large angular deviation coupled with a high angular velocity—learned disproportionately large output values, effectively encoding a rapid restorative braking action. Conversely, the rules centered near the

equilibrium origin converged to values near zero, yielding a soft and precise regulatory behavior that prevents control chattering. This linguistic mapping allows human operators to directly inspect the policy matrix, verify that the aggressive control actions are appropriately bounded, and manually intervene to constrain specific rules if safety thresholds dictate a more conservative approach [33]. This level of operational transparency bridges the gap between theoretical reinforcement learning and practical industrial deployment, offering a verifiable control law that inherently obeys the established logic of fuzzy human reasoning.

5. Conclusion

This paper has presented a comprehensive framework for the interpretable control and robust stability analysis of discrete-time nonlinear systems utilizing a novel fuzzy actor-critic learning algorithm. By replacing black-box neural networks with Takagi-Sugeno fuzzy logic systems, we have successfully preserved the semantic interpretability of the learned value functions and control policies without sacrificing the universal approximation capabilities required for complex tasks. We formulated the optimal control problem using the discrete-time Hamilton-Jacobi-Bellman equation and derived adaptive weight update laws that are computationally efficient and explicitly structured to maintain bounded parameter evolution. Through rigorous Lyapunov stability analysis, we proved that all signals within the closed-loop system achieve uniform ultimate boundedness, even in the presence of approximation errors and persistent external disturbances. The empirical evaluations conducted on challenging nonlinear benchmarks corroborate our theoretical derivations. The proposed methodology consistently outperformed contemporary deep reinforcement learning algorithms in terms of steady-state accuracy, energy efficiency, and disturbance rejection, while providing crucial insights into the decision-making mechanics through its linguistic rule base. Looking forward, there are several promising avenues for extending this research. First, the current grid-based partition of the fuzzy state space is susceptible to exponential scaling issues in high-dimensional environments, suggesting the need for adaptive rule generation mechanisms. Second, extending this robust stability framework to multi-agent cooperative control systems under communication constraints represents a critical next step toward large-scale autonomous deployment [34]. Ultimately, embedding rigorous mathematical safety guarantees and human interpretability into reinforcement learning is indispensable for the future of intelligent autonomous systems.

References

Chen, W., Hai, X., Fan, D., Tan, L., Zhang, W., & Feng, Q. (2024, August). Resilience-oriented real-time decision-making for autonomous recovery in multi-UAV systems. In 2024 IEEE 19th conference on industrial electronics and applications (ICIEA) (pp. 1-6). IEEE.

Lv, Q., Kong, W., Li, H., Zeng, J., Qiu, Z., Qu, D., ... & Pang, J. (2025). F1: A vision-language-action model bridging understanding and generation to actions. arXiv preprint arXiv:2509.06951.

Wang, J., Feng, X., Yu, Y., Wang, X., Han, X., Shi, K., ... & Cai, J. (2024). Function-dependent neural-network-driven state feedback control and self-verification stability for discrete-time nonlinear system. Neurocomputing, 609, 128422.

Fan, J., Liang, W., & Zhang, W. Q. (2025). SARNet: A Spike-Aware consecutive validation Framework for Accurate Remaining Useful Life Prediction. arXiv preprint arXiv:2510.22955.

Lv, Q., Li, H., Deng, X., Shao, R., Li, Y., Hao, J., ... & Nie, L. (2025). Spatial-temporal graph diffusion policy with kinematic modeling for bimanual robotic manipulation. In Proceedings of the Computer Vision and Pattern Recognition Conference (pp. 17394-17404).

Yang, D., Gao, Y., Wang, X., Yue, Y., Yang, Y., & Fu, M. (2025, May). Opengs-slam: Open-set dense semantic slam with 3d gaussian splatting for object-level scene understanding. In 2025 IEEE International Conference on Robotics and Automation (ICRA) (pp. 8486-8492). IEEE.

Wang, Y., Ding, P., Li, L., Cui, C., Ge, Z., Tong, X., ... & Wang, D. (2025). Vla-adapter: An effective paradigm for tiny-scale vision-language-action model. arXiv preprint arXiv:2509.09372.

Wang, Y., Zhang, R., & Liu, J. (2023). RLS-DTS: Reinforcement-learning linguistic steganalysis in distribution-transformed scenario. IEEE Signal Processing Letters, 30, 1232-1236.

Wang, C., Man, Y., Muller, R., Li, M., Celik, Z. B., Gerdes, R., & Petit, J. (2024, December). Physical ID-Transfer Attacks against Multi-Object Tracking via Adversarial Trajectory. In 2024 Annual Computer Security Applications Conference (ACSAC) (pp. 957-973). IEEE.

Zhang, W., Zhang, C., Luo, Z., Ma, J., Yuan, W., Gu, C., & Feng, C. (2025). SemanticForge: Repository-Level Code Generation through Semantic Knowledge Graphs and Constraint Satisfaction. arXiv preprint arXiv:2511.07584.

Liu, J. (2025, May). Reinforcement learning-controlled subspace ensemble sampling for complex data structures. In 2025 IEEE 7th International Conference on Communications, Information System and Computer Engineering (CISCE) (pp. 590-594). IEEE.

Driess, D., Xia, F., Sajjadi, M. S. M., Lynch, C., Chowdhery, A., Ichter, B., et al. (2023). PaLM-E: An embodied multimodal language model. In Proceedings of ICML.

Mi, L., Wang, W., Tu, W., He, Q., Kong, R., Fang, X., ... & Liu, Y. (2025, March). Empower vision applications with LoRA LMM. In Proceedings of the Twentieth European Conference on Computer Systems (pp. 261-277).

Zhang, P., Liu, H., Ge, Z., Wang, C., & Lam, E. Y. (2024). Neuromorphic imaging with joint image deblurring and event denoising. IEEE Transactions on Image Processing, 33, 2318-2333.

Huang, Y., Xu, J., Badiu, M. A., Chen, G., Coon, J., & Alouini, M. S. (2026). ISAC-Enabled Low-Overhead Beam Management: Performance Analysis and Pilot Optimization. IEEE Transactions on Wireless Communications.

Xu, Y., Shao, K., Ishikawa, I., Hashimoto, Y., Logothetis, N., & Shen, Z. (2025). A data-driven framework for Koopman semigroup estimation in stochastic dynamical systems. Chaos: An Interdisciplinary Journal of Nonlinear Science, 35(10).

Wu, J., Sun, Y., Xie, T., Chen, S., Bao, J., Xu, Y., ... & Wang, X. (2026). Cross-Modal Memory Compression for Efficient Multi-Agent Debate. arXiv preprint arXiv:2602.00454.

Xu, Y., Li, F., Fujisawa, M., Cheng, X., Marzouk, Y., & Ishikawa, I. (2025). Generative Modeling through Koopman Spectral Analysis: An Operator-Theoretic Perspective. arXiv preprint arXiv:2512.18837.

Zhu, G., Li, M., Dai, H., Liu, X., Wang, W., Li, K., ... & Wang, W. (2025). Enabling MoE on the Edge via Importance-Driven Expert Scheduling. arXiv preprint arXiv:2508.18983.

Li, Y., Zou, Y., He, X., Xu, Q., Liu, M., Jin, S., ... & Zhang, J. (2025). HFA-UNet: Hybrid and full attention UNet for thyroid nodule segmentation. Knowledge-Based Systems, 114245.

Norris, J., He, Z., Qu, Y., Chen, G., Hertzog, C., & Jin, D. (2025, September). An in-network approach for pmu missing data recovery with data plane programmability. In 2025 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm) (pp. 1-7). IEEE.

Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). Trust region policy optimization. In Proceedings of ICML.

Tang, Y., Kojima, K., Gotoda, M., Nishikawa, S., Hayashi, S., Koike-Akino, T., ... & Klamkin, J. (2020). InP grating coupler design for vertical coupling of InP and silicon chips. Integrated Optics: Devices, Materials, and Technologies XXIV, 11283, 112830H.

Brohan, A., Chebotar, Y., Finn, C., Hausman, K., Herzog, A., Ho, D., et al. (2023). RT-2: Vision-language-action models transfer web knowledge to robotic control. In Proceedings of the Conference on Robot Learning.

Zhang, P., Ge, Z., Song, L., & Lam, E. Y. (2023). Neuromorphic imaging with density-based spatiotemporal denoising. IEEE Transactions on Computational Imaging, 9, 530-541.

Ross, S., Gordon, G., & Bagnell, D. (2011). A reduction of imitation learning and structured prediction to no-regret online learning. In Proceedings of AISTATS.

Wang, J., Zhang, Y., Zhang, B., Xia, J., & Wang, W. (2025). Ipfa-net: Important points feature aggregating net for point cloud classification and segmentation. Chinese Journal of Electronics, 34(1), 322-337.

Fan, D., Feng, Q., Zhang, A., Liu, M., Ren, Y., & Wang, Y. (2023). Optimization of scheduling and timetabling for multiple electric bus lines considering nonlinear energy consumption model. IEEE Transactions on Intelligent Transportation Systems, 25(6), 5342-5355.

Zhang, W., Luo, M., & Chen, Z. (2024, October). Hybrid Forecasting: ML Predictions of Lake-Effect Regional Extreme Precipitations. In 2024 7th International Conference on Universal Village (UV) (pp. 1-11). IEEE.

Wang, R., Guo, T., Li, Y., Meng, D., & Liang, B. (2025). Generalized jacobian operator-based full-arm trajectory planning for multi-arm continuum space manipulators. Aerospace Science and Technology, 111559.

Zou, Y., Jin, S., Deng, A., Zhao, Y., Wang, J., & Chen, C. (2025). AIR: Enabling Adaptive, Iterative, and Reasoning-based Frame Selection For Video Question Answering. arXiv preprint arXiv:2510.04428.

Zhou Z, Leng N, Ma H, et al. Study on Real-Time Data Analysis and Intelligent Forecasting Methods for Integrated Circuit Supply Chains Based on Cloud Computing[C]//Proceedings of the 2nd International Symposium on Integrated Circuit Design and Integrated Systems. 2025: 245-250.

Wang, W., Mi, L., Cen, S., Dai, H., Li, Y., Fu, X., & Liu, Y. (2025). Region-based Content Enhancement for {Efficient} Video Analytics at the Edge. In 22nd USENIX Symposium on Networked Systems Design and Implementation (NSDI 25) (pp. 613-633).

Wang, J., Feng, X., Yu, Y., Wang, X., Werghi, N., Han, X., ... & Tashi, N. (2025). Fuzzy actor–critic learning-based interpretable control and stability-informed guarantee with error mapping for discrete-time nonlinear system. Chaos, Solitons & Fractals, 199, 116878.