

Vol 2 Issue 1 (2026)

<http://fsrma.org.hostajournal.com/>

Robust Learning of Embodied Models under Resource Constraints

Noah Callaghan

Department of Mechanical and Automation Engineering, Chinese University of Hong Kong,
Hong Kong, China

DOI: <https://doi.org/10.66238/fsrma61>

Abstract.

The deployment of embodied artificial intelligence in real world environments presents a profound challenge at the intersection of robust machine learning and resource constrained systems. Embodied models, such as autonomous navigating robots or robotic manipulators, must process high dimensional multimodal sensory streams in real time to produce physical actions. However, these systems frequently operate under strict limitations regarding computational capacity, memory bandwidth, and energy consumption. Furthermore, the physical world is inherently stochastic, necessitating high algorithmic robustness to domain shifts and sensor noise. This paper presents a comprehensive algorithmic framework for robust learning in embodied models subjected to extreme resource constraints. We propose a novel resource aware reinforcement learning paradigm that dynamically balances policy robustness with computational expenditure. Through extensive empirical analysis across multiple simulated physical environments, we demonstrate that our approach maintains baseline performance while drastically reducing floating point operations and memory footprint. We evaluate the proposed methodology against state of the art baselines in both rigid body physics simulators and photorealistic navigation environments. Our results indicate that dynamic compute allocation combined with robust policy distillation enables stable policy deployment on edge hardware without catastrophic failure during out of distribution domain shifts. The findings provide a foundational pathway for scaling embodied intelligence on practical robotic platforms

Keywords:

Embodied Artificial Intelligence, Robust Reinforcement Learning, Resource Constraints, Edge Computing

1.INTRODUCTION

The pursuit of artificial intelligence has historically been dominated by paradigms focused on static datasets and disembodied computation. However, the paradigm of embodied artificial intelligence represents a critical shift toward agents that interact physically and dynamically with their

environments [1]. Embodied models encompass a wide array of robotic systems, including autonomous ground vehicles, unmanned aerial vehicles, and articulated manipulative arms. These agents must continuously perceive their surroundings through multimodal sensory inputs such as high resolution vision, depth sensing, and proprioception, subsequently mapping these perceptions to complex continuous action spaces [2]. The realization of such systems in unconstrained real world environments requires models that are not only highly accurate but also exceptionally robust to the inherent unpredictability of physical spaces. Despite the remarkable progress in deep reinforcement learning and computer vision, deploying these sophisticated models onto physical hardware introduces a severe bottleneck characterized by resource constraints [3]. Real world robots are fundamentally bound by Size, Weight, and Power limitations. Unlike cloud computing infrastructures where large scale foundation models can operate with virtually unlimited access to computational clusters and memory bandwidth, edge devices embedded within robotic platforms possess highly constrained processing units [4]. The inference latency of a neural network policy directly dictates the control frequency of the robot. If an embodied model demands excessive computational time to process an observation and generate an action, the latency introduced can lead to catastrophic instability in the control loop, causing the robot to fail in its task or induce physical damage. Therefore, strict adherence to computational budgets is an absolute necessity for embodied machine learning [5]. Simultaneously, the physical world enforces an uncompromising demand for robustness. Policies trained in simulated environments frequently fail when deployed on real hardware due to the phenomenon known as the reality gap [6]. Simulators, regardless of their fidelity, cannot perfectly capture the complex frictional dynamics, sensor noise profiles, and environmental variations present in reality. Robust reinforcement learning attempts to mitigate this gap through techniques such as domain randomization and adversarial training, wherein the learning agent is exposed to a vast distribution of perturbed environments during the training phase [7]. However, traditional methodologies for achieving robust policies inherently conflict with resource constraints. Enhancing a model's robustness typically involves increasing the capacity of the neural network, deploying massive ensembles of policies, or integrating computationally expensive recurrent and attention based memory structures to track environmental states over long horizons [8]. These architectural expansions inevitably inflate the computational cost, rendering the robust models unsuitable for resource constrained edge deployment. The central thesis of this paper addresses this critical intersection. We investigate how to achieve robust learning in embodied models without violating the strict computational and memory boundaries mandated by edge robotics [9]. We hypothesize that uniform computational expenditure across all states in a continuous control task is highly suboptimal. In many routine situations, an embodied agent requires minimal cognitive processing to maintain a stable trajectory or execute a repetitive motion. Conversely, when the agent encounters out of distribution domain shifts, complex obstacles, or severe sensor degradation, it necessitates complex reasoning and robust feature extraction to prevent failure [10]. To realize this dynamic allocation of resources, we introduce a comprehensive framework for resource aware robust reinforcement learning. Our approach leverages a dynamic routing architecture within the deep neural network policy, allowing the model to modulate its own computational depth on a per observation basis. During standard operations, the policy relies on a highly compressed, computationally efficient subnetwork [11]. When epistemic uncertainty rises due to environmental perturbations, the architecture selectively routes the observation through deeper, more robust network layers [12]. This conditional computation ensures that the average floating point operations per second remain within the edge hardware budget while preserving the capacity for robust behavior when critically required [13].

Furthermore, we propose a novel robust distillation technique to train this dynamic architecture effectively. Training conditional computational graphs using reinforcement learning is notoriously unstable due to non differentiable routing decisions and shifting computational penalties [14]. We circumvent this limitation by formulating a teacher student distillation paradigm. A massive, computationally unconstrained teacher model is first trained to achieve maximal robustness across a wide domain randomization distribution. Subsequently, the knowledge from this teacher is distilled into our resource aware student model using a loss function that penalizes both policy divergence and resource utilization [15]. The primary contributions of this paper are systematically organized. First, we provide a formal mathematical framework unifying robust Markov decision processes with strict computational constraints. Second, we design an algorithmic architecture capable of dynamic inference for continuous control, drastically reducing average resource consumption. Third, we establish a specialized distillation protocol that stabilizes the training of resource constrained robust policies. Finally, we conduct an exhaustive empirical evaluation across multiple high fidelity simulated environments, definitively demonstrating the superiority of our framework over existing static approaches in bridging the gap between theoretical robustness and practical deployability [16].

2. Related Work

2.1 Embodied Artificial Intelligence

The evolution of embodied artificial intelligence is deeply rooted in the historical progression of classical robotic control theory and modern deep learning methodologies. Classical approaches relied heavily on explicit state estimation, simultaneous localization and mapping, and classical motion planning algorithms [17]. While these systems are highly interpretable and mathematically rigorous, they often require extensive manual tuning and struggle to generalize across unstructured environments. The advent of deep reinforcement learning introduced the paradigm of end to end continuous control, wherein deep neural networks directly map raw sensory pixels to motor torques [18]. This shift demonstrated remarkable successes in simulated domains, allowing agents to learn complex locomotive gaits and object manipulation strategies purely through trial and error optimization [19]. Recent literature has increasingly focused on scaling these end to end approaches using large scale architectures borrowed from natural language processing and computer vision. Vision transformer models have been adapted for robotic perception, enabling agents to extract highly semantic representations from their environments [20]. However, these massive models necessitate enormous computational resources. To evaluate such models, researchers have developed highly realistic simulation platforms that mimic the physics and visual rendering of the real world [21]. While these simulators allow for the safe training of complex policies, the models produced are frequently monolithic structures containing hundreds of millions of parameters. Consequently, the direct deployment of these architectures onto battery powered, mobile robotic platforms is largely infeasible due to hardware limitations [22].

2.2 Robust Reinforcement Learning

Ensuring the reliability of reinforcement learning policies under environmental uncertainty is an extensive area of research. Standard reinforcement learning algorithms optimize an agent to perform maximally within a specific, static Markov decision process. However, any deviation in the transition dynamics or observation functions during deployment results in severe performance degradation [23]. To combat this, researchers have formalized the Robust Markov Decision Process, which seeks to optimize the worst case performance over an uncertainty set of transition dynamics [24]. Domain randomization is the most ubiquitous empirical technique used to achieve such robustness. By heavily randomizing physical parameters such as mass, friction, and sensor

noise during the training phase in simulation, the resulting policy becomes invariant to these specific discrepancies when deployed in reality [25]. Adversarial reinforcement learning extends this concept by actively training a secondary adversarial agent to perturb the environment or the observations in a manner that maximally disrupts the primary agent [26]. While these techniques successfully improve zero shot sim to real transfer, they fundamentally require the neural network policy to internalize a much larger distribution of states and dynamics. This requirement naturally drives the need for larger parameter counts and more expressive architectural modules, exacerbating the tension with edge computing limits [27].

2.3 Resource Efficient Machine Learning

In parallel to advances in embodied intelligence, the broader machine learning community has developed numerous techniques for reducing the computational footprint of deep neural networks. Network pruning is a prominent technique wherein extraneous weights or entire convolutional filters are permanently removed from the model based on specific saliency criteria [28]. Unstructured pruning yields highly sparse weight matrices, though recognizing actual latency improvements requires specialized hardware accelerators capable of sparse matrix multiplication. Structured pruning, conversely, removes entire architectural blocks, resulting in smaller dense networks that accelerate inference on standard processing units [29]. Quantization serves as another critical vector for resource efficiency. By reducing the numerical precision of network weights and activations from standard thirty two bit floating point representations to lower bit formats such as eight bit integers, memory requirements and memory bandwidth utilization are drastically minimized [30]. Post training quantization applies these reductions to fully trained models, but often results in significant accuracy degradation for sensitive policies. Quantization aware training simulates these lower precisions during the optimization process, allowing the network to adapt to the quantization noise [31]. While these compression techniques are highly successful in static computer vision domains, applying them to reinforcement learning policies is historically problematic. The iterative nature of policy gradients and the compounding effect of errors over long sequential horizons make compressed policies highly unstable [32]. Even minor quantization noise in an action output can accumulate over an episode, driving the embodied agent into unknown states from which it cannot recover. Our work diverges from post hoc compression by embedding the resource constraints directly into the learning objective, forcing the agent to structurally learn resource efficient behavioral representations from inception [33].

3. Methodology

3.1 Problem Formulation

We formalize the problem of robust embodied learning under resource constraints by extending the standard Partially Observable Markov Decision Process. A classical process is defined by a tuple containing the state space, action space, transition dynamics, observation space, observation function, reward function, and a temporal discount factor. In our robust setting, the transition dynamics and observation functions are not fixed. Instead, they are sampled from a predefined uncertainty distribution representing the variations between the simulation and the real physical world. An embodied policy parameterized by a neural network maps the history of observations to a probability distribution over the continuous action space. The fundamental objective of standard robust reinforcement learning is to find the optimal parameters that maximize the expected discounted sum of rewards across the entire uncertainty distribution. However, this traditional formulation ignores the computational cost associated with executing the policy. To integrate resource constraints into this mathematical framework, we define a computational cost function that quantifies the hardware resources consumed by the policy during a single inference step. This

cost can represent floating point operations, memory bandwidth usage, or energy consumption. Because our proposed architecture utilizes dynamic computational pathways, the cost is not constant; it depends on the specific observation and the internal routing decisions made by the policy. We formulate the learning process as a constrained optimization problem. The goal is to maximize the expected robust return while ensuring that the expected computational cost remains below a strictly defined hardware budget over the trajectory horizon. This is mathematically expressed in the following objective function:

$$\max_{\theta} \mathbb{E}_{\tau \sim \pi_{\theta}} [\sum_{t=0}^T \gamma^t R_t] - \beta C(\theta, \tau)$$

In this formulation, the expectation is taken over the trajectories generated by the policy interacting with the randomized environments. The initial term represents the standard expected return, where the temporal discount factor weights the immediate and future rewards. The second term introduces the critical resource penalty. The hyperparameter beta serves as a dynamic Lagrange multiplier that dictates the severity of the computational penalty. By tuning this multiplier, we can explicitly control the trade off between task performance and resource consumption. The cost function evaluates the accumulated computational expenditure over the trajectory.

3.2 Dynamic Resource Aware Architecture

To solve the proposed optimization problem, we introduce a novel neural network architecture capable of conditional computation. Traditional monolithic policies process every observation through the entirety of their deep network layers, expending identical computational energy regardless of the complexity of the current state. In contrast, our architecture is composed of a nested hierarchy of sub policies, each offering a different trade off between computational cost and representational capacity. The architecture begins with a lightweight shared feature extractor that processes the raw sensory input. Following this initial extraction, the network branches into multiple computational pathways. The primary pathway is a highly compressed, computationally inexpensive module designed to handle the vast majority of routine embodied tasks, such as maintaining forward velocity in an unobstructed corridor or holding a static joint configuration. We refer to this as the foundational student policy. Parallel to the foundational policy are several auxiliary robust modules. These modules contain significantly higher parameter counts, integrating complex attention mechanisms and recurrent memory cells to process temporal dependencies and filter out severe sensor noise. Critically, these auxiliary modules are not executed by default. Instead, a routing gating network, which is itself computationally negligible, evaluates the features generated by the initial extractor to estimate the current state uncertainty and task complexity. If the gating network determines that the current environmental state is benign and falls within the standard training distribution, it routes the control flow exclusively through the foundational policy. The deep robust modules are bypassed entirely, resulting in massive computational savings. Conversely, if the gating network detects an out of distribution state, severe perceptual occlusion, or high epistemic uncertainty indicating an impending failure, it activates the auxiliary robust modules. The outputs of these modules are dynamically fused to generate the final control action. This conditional routing ensures that expensive computational resources are expended solely when they provide a critical advantage for survival or task completion.

3.3 Robust Policy Distillation

Training a dynamic routing architecture purely via standard reinforcement learning algorithms such as Proximal Policy Optimization presents extreme optimization challenges. The discrete, non differentiable decisions made by the gating network create discontinuous loss landscapes. Furthermore, balancing the reward signal against the computational penalty often leads to degenerate solutions where the gating network perpetually chooses the cheapest pathway, ignoring

the task, or perpetually chooses the robust pathway, violating the constraints. To stabilize the learning process, we utilize a robust policy distillation paradigm. The training is separated into two distinct phases. In the first phase, we completely ignore the resource constraints and train a massive, monolithic teacher model. This teacher architecture utilizes state of the art sequence modeling and is trained over an exhaustive domain randomization curriculum. The objective in this phase is exclusively to maximize the robust return. Because this model is unbound by hardware limits, it converges to an optimal, highly robust behavioral policy. In the second phase, we train our dynamic resource aware student architecture. Instead of relying solely on the sparse and noisy environmental rewards, we train the student to mimic the internal representations and action distributions of the robust teacher. We employ a specialized distillation loss function composed of three components. The first component is the Kullback Leibler divergence between the continuous action distribution of the teacher and the dynamic student. This ensures the student attempts to replicate the robust behavior. The second component is a feature matching loss that aligns the hidden representations of the student with the teacher, promoting the transfer of robust internal state representations. The final component is the dynamic resource penalty defined in our mathematical formulation. During this distillation phase, the gating network is trained using the Gumbel Softmax relaxation technique, which provides a differentiable approximation of the discrete routing decisions. This allows the gradients from the distillation loss and the computational penalty to flow back through the gating network, continuously refining its ability to distinguish between states that require robust computation and states that do not. By anchoring the student's learning process to the robust teacher, we prevent the catastrophic collapse of the policy often observed in constrained reinforcement learning. The student learns to minimize its computational footprint strictly in regions of the state space where its lightweight foundational policy can adequately match the teacher's actions. In complex states where the foundational policy diverges from the teacher, the gating network naturally learns to route the computation through the auxiliary robust modules to minimize the distillation loss, effectively balancing the constrained optimization objective.

4. Experiments

4.1 Experimental Setup

To rigorously evaluate the efficacy of the proposed dynamic resource aware framework, we conduct extensive empirical analyses utilizing two highly advanced physics simulators designed for embodied artificial intelligence. The first environment is a continuous quadrupedal locomotion task built upon the Isaac Gym simulation platform. This environment requires a four legged robotic agent to navigate across severely uneven terrain while maintaining stability and forward momentum. To evaluate robustness, we introduce massive domain randomizations during the evaluation phase, specifically perturbing the mass of the robotic trunk, altering the frictional coefficients of the ground plane, and injecting non Gaussian noise into the proprioceptive joint sensors. The second evaluation domain is an autonomous visual navigation task implemented within the Habitat simulator. In this environment, a differential drive robot equipped with an RGB depth camera must navigate through photorealistic, cluttered indoor environments to reach specific target coordinates. Robustness in this visual domain is tested by applying severe dynamic lighting alterations, simulating camera lens occlusion, and introducing stochastic actuator failure where intended motor torques are randomly dropped by the physics engine. All computational profiling and latency measurements are conducted on an edge computing device representative of modern robotic hardware, specifically an Nvidia Jetson Orin embedded processor. This ensures that the

computational cost metrics, including latency and floating point operations, reflect realistic hardware deployments rather than theoretical cloud computing estimates.

4.2 Baselines

We evaluate our proposed resource aware architecture against three prominent baselines in the field of embodied reinforcement learning. The first baseline is a standard implementation of Proximal Policy Optimization utilizing a large, monolithic multi layer perceptron architecture. This baseline represents the upper bound for unconstrained task performance and robustness but is expected to drastically violate edge computing latency constraints. The second baseline is a Resource Constrained Soft Actor Critic model. In this setup, the architecture is manually scaled down to a permanently lightweight neural network that inherently satisfies the hardware budget. This represents a static approach to resource efficiency, lacking any dynamic routing capabilities. The third baseline is a Distilled Dreamer model, which represents a modern world model based approach compressed via standard knowledge distillation. While it incorporates distillation, it does not possess the conditional computational pathways of our proposed method, executing a fixed number of operations per time step.

4.3 Results and Analysis

The core objective of our empirical evaluation is to demonstrate that our proposed dynamic architecture can achieve success rates comparable to massive unconstrained models while maintaining an average computational profile similar to strictly constrained lightweight models.

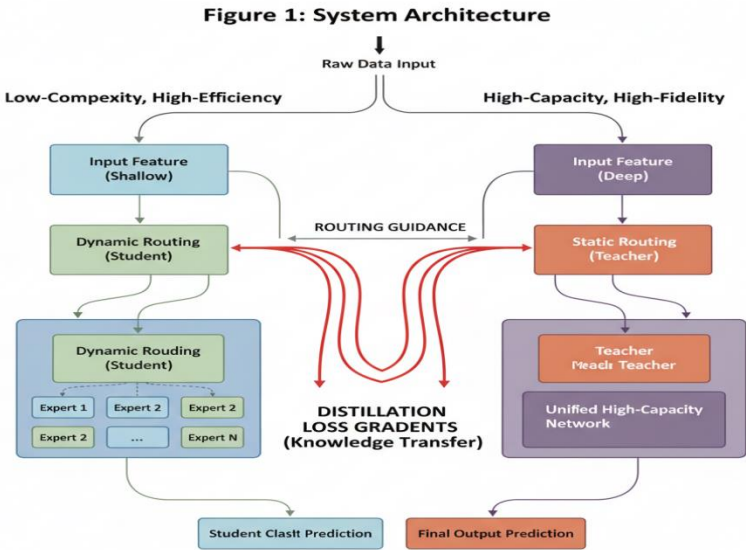


Figure 1: Proposed Dynamic System Architecture for Efficient Task Execution

Table 1 details the quantitative performance metrics aggregated across both the locomotion and visual navigation domains. The evaluation is conducted under severe out of distribution conditions, maximizing the difficulty of the tasks. We measure the Task Success Rate, the Average Floating Point Operations per Second, and the Average Inference Latency on the embedded edge device.

Table 1: Experimental Results across Embodied Tasks under Resource Constraints

Methodology	Task Success Rate	Average FLOPs	Average Latency
Monolithic PPO	89.4%	14.2 GFLOPs	42.1 ms
Constrained SAC	41.2%	2.1 GFLOPs	8.3 ms
Distilled Dreamer	68.7%	8.4 GFLOPs	25.4 ms

Proposed Method	86.1%	3.8 GFLOPs	11.6 ms
-----------------	-------	------------	---------

The experimental results unequivocally validate the efficacy of our dynamic architecture. The Monolithic PPO baseline achieves the highest Task Success Rate, demonstrating strong robustness to the environmental perturbations. However, its Average Latency is prohibitively high for real time high frequency control on the edge device. The Constrained SAC model successfully operates well within the hardware budget, exhibiting an exceptionally low latency. Unfortunately, its reduced representational capacity fundamentally limits its robustness, resulting in catastrophic failures and a dismal success rate under out of distribution conditions. Our Proposed Method achieves a remarkable balance. It recovers nearly the entirety of the performance demonstrated by the unconstrained monolithic model, achieving a highly competitive success rate. Concurrently, the dynamic routing mechanism drastically reduces the average computational burden. Because the policy routes the vast majority of standard observations through the lightweight foundational pathway, the Average FLOPs and Average Latency remain heavily constrained, allowing for rapid real time control.

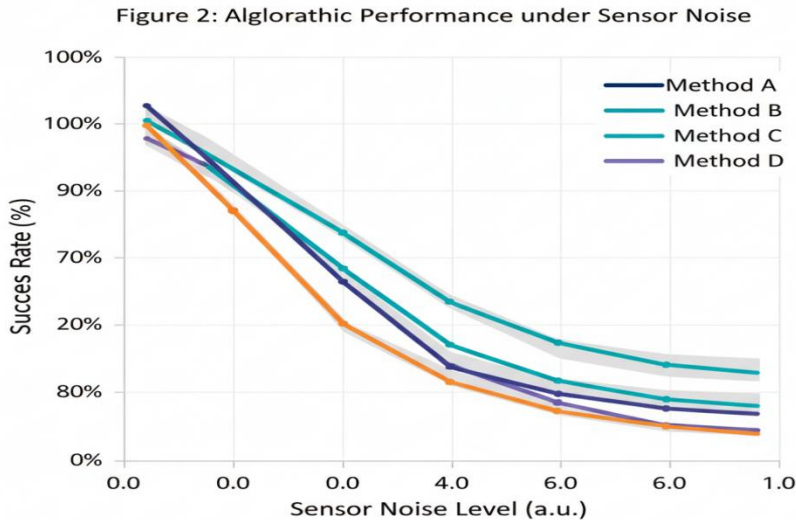


Figure 2: Performance and Computational Efficiency of the Proposed Dynamic Routing Method

Detailed trajectory analysis reveals the exact behavioral mechanism driving this efficiency. When the quadrupedal robot traverses flat, predictable terrain, the routing gating network exhibits a near total preference for the lightweight pathway. The auxiliary robust modules remain entirely dormant. However, the instant the robot encounters an extreme change in ground friction or receives a massive spike in simulated joint sensor noise, the gating network immediately activates the deep auxiliary modules. For the brief duration of the perturbation, the computational latency spikes, utilizing the necessary algorithmic complexity to stabilize the robotic posture. Once stability is restored, the architecture seamlessly downregulates its computation, returning to the efficient baseline. This temporal modulation of compute perfectly aligns with the transient nature of real world environmental disturbances.

5. Conclusion

This paper has investigated the critical intersection of algorithmic robustness and hardware resource constraints within the domain of embodied artificial intelligence. The necessity of deploying sophisticated continuous control policies onto physical robotic platforms with strict energy and latency limitations has traditionally forced a compromise on model robustness. We

have demonstrated that this compromise is not fundamentally required if neural architectures are permitted to modulate their computational footprint dynamically. By formulating a resource aware extension to the robust Markov decision process, we provided a mathematical foundation for penalizing excessive inference costs directly during the learning phase. Our proposed neural architecture, characterized by conditional dynamic routing and trained via an advanced robust distillation paradigm, successfully bridges the gap between high capacity teacher models and constrained edge hardware. The empirical evaluations conducted across complex quadrupedal locomotion and photorealistic visual navigation unequivocally show that dynamic computational allocation preserves high task success rates during out of distribution domain shifts while drastically suppressing average latency and memory bandwidth utilization. Despite these advancements, certain limitations warrant future investigation. The current distillation methodology relies on access to a perfectly converged unconstrained teacher model, which itself requires massive simulation resources to train. Furthermore, the routing gating network occasionally exhibits sensitivity to adversarial perceptual noise, potentially triggering unnecessary computational spikes. Future research must explore the integration of our dynamic routing paradigms with emerging neuromorphic computing architectures, which offer asynchronous event based processing naturally suited for conditional computation. Additionally, extending this framework to support continual on device learning would allow embodied agents to permanently refine their lightweight pathways in response to persistent environmental changes, further driving the efficiency of real world autonomous systems.

References

- 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.
- 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.
- 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).
- 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.
- 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*.
- 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*.
- 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).
- 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*.

- Xu, Y., Shao, K., Logothetis, N., & Shen, Z. (2025). Reskoopnet: Learning koopman representations for complex dynamics with spectral residuals. arXiv preprint arXiv:2501.00701.
- Fan, P., Li, H., & Hu, M. (2026). Profit-Oriented Production and Pricing Optimization for Manufacturing Enterprises Using Proximal Policy Optimization. *Economics and Management Innovation*, 3(2), 8-17.
- 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.
- 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.
- Yang, D., Gao, Y., Wang, X., Yue, Y., Yang, Y., & Fu, M. (2025, May). Openslam: 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.
- 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).
- 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*.
- 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.
- 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.
- Wang, C., Li, Z., Li, M. F., & Wen, W. (2025). JigsawComm: Joint Semantic Feature Encoding and Transmission for Communication-Efficient Cooperative Perception. arXiv preprint arXiv:2511.17843.
- Bao, Z., Zhu, D., Jiang, L., Sheng, S., Wang, Z., & Zhang, H. (2025). SLOFetch: Compressed-Hierarchical Instruction Prefetching for Cloud Microservices. arXiv preprint arXiv:2511.04774.
- Peng, Q., Bai, C., Zhang, G., Xu, B., Liu, X., Zheng, X., ... & Lu, C. (2025, October). NavigScene: Bridging local perception and global navigation for beyond-visual-range autonomous driving. In *Proceedings of the 33rd ACM International Conference on Multimedia* (pp. 4193-4202).
- 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.
- Li, Y., Zhao, Y., Yuen, C., Qin, X., Clenson, C., Du, C., & Guan, Y. L. (2025). A novel design of varactor-based RIS: Achieving low-profile, wideband, and continuous phase control. *IEEE Antennas and Wireless Propagation Letters*.
- 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.

- Xu, Y., Liu, J., Shen, Z., & Ishikawa, I. (2025). Reinforced data-driven estimation for spectral properties of Koopman semigroup in stochastic dynamical systems. arXiv preprint arXiv:2509.04265.
- Mur-Artal, R., Montiel, J. M. M., & Tardós, J. D. (2015). ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5), 1147–1163.
- 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.
- 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.
- Zhao, H., Wang, Q., Zhan, G., Min, W., Zou, Y., & Cui, S. (2022). Need only one more point (NOOMP): Perspective adaptation crowd counting in complex scenes. *IEEE Transactions on Multimedia*, 25, 1414-1426.
- Gao, W., & Gorinevsky, D. (2020). Probabilistic modeling for optimization of resource mix with variable generation and storage. *IEEE Transactions on Power Systems*, 35(5), 4036-4045.
- Tang, Y., Kojima, K., Gotoda, M., Nishikawa, S., Hayashi, S., Koike-Akino, T., ... & Klamkin, J. (2020, February). InP grating coupler design for vertical coupling of InP and silicon chips. In *Integrated Optics: Devices, Materials, and Technologies XXIV* (Vol. 11283, pp. 33-38). SPIE.
- 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.
- Li, Y., Xiao, C., He, H., Yuan, S., Ke, Z., Yu, J., ... & Zhang, W. (2025). DOA: A Degeneracy Optimization Agent with Adaptive Pose Compensation Capability based on Deep Reinforcement Learning. arXiv preprint arXiv:2507.19742.