GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Journal of Computer, Signal, and System Research

Article

Generative Adversarial Network Enhanced Adaptive Control for Aerospace Launch Vehicles

Abdulrahman Fahad [1,*], Fatemeh Mohammadi [2], Amir Hussain [3], Mariam Eissa [4], Hassan El-Nur [5], Sara Al-Sabah [6] and Krzysztof Zalewski [7]

[1] Department of Electrical Engineering, King Fahd University of Petroleum & Mineral, Dhahran, Saudi Arabia
[2] Department of Chemical Engineering, Sharif University of Technology, Tehran, Iran
[3] Department of Computer Science, National University of Sciences & Technolog, Islamabad, Pakistan
[4] Department of Civil Engineering, Khalifa University, Abu Dhabi, United Arab Emirates
[5] Department of Agricultural Engineering, University of Khartoum, Khartoum, Sudan.
[6] Department of Mechanical Engineering, Kuwait University, Kuwait City, Kuwait
[7] The Faculty of Electronics and Information Technology, Warsaw University of Technology, Poland
* Correspondence: Abdulrahman Fahad, Department of Electrical Engineering, King Fahd University of Petroleum & Mineral, Dhahran, Saudi Arabia

Abstract: A theoretical analysis of an on-line, autonomously intelligent, adaptive tracking controller for satellites-employing Generative Adversarial Networks (GANs)-is presented. The controller receives real-time sensory data and a scalar performance signal, autonomously refining thruster or attitude-control commands without requiring explicit foreknowledge of the satellite's internal dynamics. By leveraging an adversarial interplay between a generator and a discriminator module, the approach rapidly adapts to evolving orbital conditions and unanticipated disturbances, thus preserving robust tracking accuracy. The underlying on-line learning mechanism enables continuous policy adjustments, obviating the need for extensive offline tuning. Experimental evidence, derived from a representative satellite undergoing maneuvers in a low-Earth-orbit environment, demonstrates the algorithm's capacity to compensate for nonstationary aerodynamic drag and shifting mass distribution while maintaining precise trajectory regulation. This work underscores the viability of GAN-augmented adaptive methods for advanced aerospace applications, offering heightened resilience and efficiency in the face of dynamic mission constraints.

Keywords: deep learning; GANs; adaptive control; satellite theory

Received: 16 December 2025
Revised: 22 January 2026
Accepted: 02 February 2026
Published: 05 February 2026

1. Introduction

The primary objective of a satellite's flight control system-or onboard "autopilot"-is to maintain the spacecraft's orientation as dictated by higher-level guidance algorithms. The autopilot assesses the satellite's attitude through an inertial measurement unit (IMU) or analogous sensory apparatus and subsequently issues corrective commands to the thrusters [1-3], aiming to fulfill the guidance directives. Achieving robust and precise control requires meeting three core, often competing, design specifications: (1) ensuring satellite stability, (2) enabling swift and accurate responses to updates in guidance inputs, and (3) minimizing disruptive attitude maneuvers [4] in the face of varying environmental conditions (e.g., solar radiation pressure, gravitational perturbations) to preserve structural integrity and communication link quality [5].

A significant difficulty arises from the non-stationary nature of the satellite's orbital and attitude dynamics. Uncertainties in onboard mass distribution due to fuel consumption, as well as time-varying disturbances, render conventional, fixed-gain controllers less effective [6]. These methods may experience degraded performance or require extensive retuning when confronted with the highly dynamic and uncertain

environment of low-Earth or geostationary orbit. Consequently, adaptive strategies have been developed to adjust control actions in real time, ensuring consistent performance under unpredictable conditions.

Over the past decade, the control research community has become increasingly fascinated by biologically inspired and machine learning-based solutions for tackling challenging aerospace applications. Methods such as fuzzy sets, neural networks, and genetic algorithms have demonstrated their capacity to handle nonlinearities and complex optimization landscapes. Recently, Generative Adversarial Networks (GANs) have emerged as a promising deep learning framework for high-dimensional function approximation, data synthesis [7], and robust decision-making. By pitting a generator model against a discriminator model, GANs can iteratively refine control policies or disturbance estimations, offering a potent mechanism for coping with external uncertainties in the orbital environment.

Parallel advancements in emotional learning models, such as the approach based on the amygdala's computational analogy, have successfully rapidly adapted control laws for simpler or predominantly linear systems. However, incorporating GANs into these biologically motivated schemes opens the door to a more powerful class of intelligent controllers. In parallel, recent advances in invertible liquid neural network architectures have demonstrated that learning inverse kinematics and dynamics for complex robotic manipulators can yield highly precise, data-efficient control policies that remain robust under severe nonlinearities and uncertainty, with clear implications for high-agility aerospace platforms and military-grade robotic systems [8]. The proposed approach can flexibly compensate for highly nonlinear satellite dynamics, time-varying uncertainties, and unmodeled external forces by synthesizing features from emotional learning and deep adversarial training.

This paper introduces a direct adaptive output-control architecture for satellites that leverages both an emotional learning paradigm and GAN-based deep learning [9]. Our method continuously learns from sensory inputs, updated reference commands, and real-time performance feedback, adjusting its actuator directives without explicit pre-knowledge of the satellite's full dynamic model [10-13]. The principal advantage of this integrated approach is its capacity for online adaptation-guided by adversarial optimization-while incurring relatively modest computational overhead. In the current study, we focus on the satellite's attitude channels, such as pitch and yaw, which are critical for ensuring consistent line-of-sight communication, precise orbit insertion, and payload orientation [14]. Figure 1 illustrates the optimal rendezvous path and corresponding safety margins considered for autonomous rocket maneuver planning.

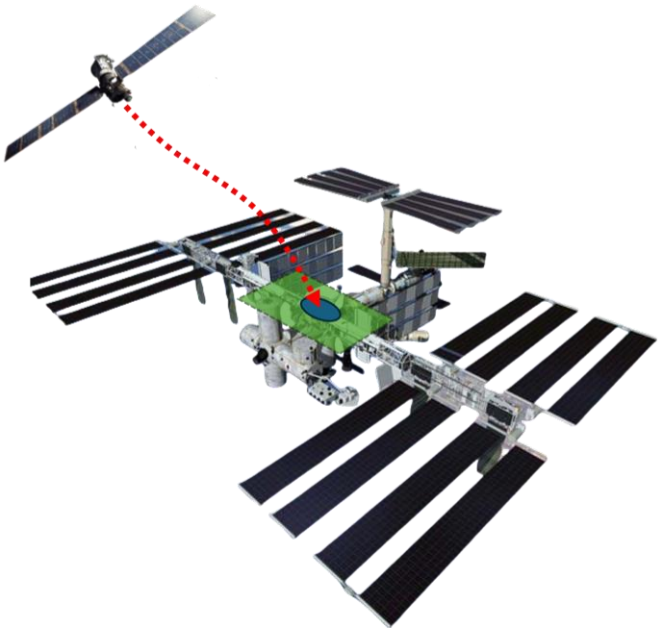

Figure 1. Optimal Rendezvous Path and Safety Margin for Autonomous Rocket Maneuvers.

The remainder of this paper is organized as follows. Section 2 details the proposed control algorithm, derived from the synergy of adversarial training and emotional learning. Section 3 outlines the satellite's dynamic model and provides numerical parameters relevant to orbital maneuvers [15]. Section 4 demonstrates the simulation setup and validates the designed controller's performance under both nominal and perturbed conditions. Finally, Section 5 concludes with a discussion of key findings and prospective directions for future research.

2. Methodology

2.1. Onboard Imaging Frames and Feature Points

Consider a satellite equipped with a monocular camera for real-time relative navigation. Let $\mathcal{F}$ denote the camera's local coordinate frame [16], rigidly attached to the spacecraft's current pose, and let $\mathcal{F}^\star$ be a stationary frame that corresponds to the desired camera orientation-essentially, the reference pose that aligns the satellite's instruments with a target of interest (e.g., a docking port, another satellite, or a celestial beacon).

Each observable landmark Ω_i in the scene projects onto the camera sensor, producing Euclidean coordinates in both frames. Specifically, define

$$\tilde{\eta}_i(t) \triangleq \begin{bmatrix} \alpha_i(t) \\ \beta_i(t) \\ \gamma_i(t) \end{bmatrix} \text{ and } \tilde{\eta}_i^\star \triangleq \begin{bmatrix} \alpha_i^\star \\ \beta_i^\star \\ \gamma_i^\star \end{bmatrix}$$

where $\alpha_i(t), \beta_i(t), \gamma_i(t) \in \mathbb{R}$ denote the landmark's coordinates at time ttt in $\mathcal{F}$, while $\alpha_i^\star, \beta_i^\star, \gamma_i^\star \in \mathbb{R}$ denote the coordinates in $\mathcal{F}^\star$. By convention, $\tilde{\eta}_i(t)$ evolves as the satellite moves [17], whereas $\tilde{\eta}_i^\star$ remains fixed, representing the goal configuration [18].

From classical Euclidean geometry, the relationship between $\tilde{\eta}_i(t)$ and $\tilde{\eta}_i^\star$ can be expressed as

$$\tilde{\eta}_i(t) = \xi_f(t) + \rho(t)\tilde{\eta}_i^\star$$

where $\rho(t) \in SO(3)$ is the rotation mapping $\mathcal{F}^\star$, and $\xi_f(t) \in \mathbb{R}^3$ is the translation from $\mathcal{F}$ to $\mathcal{F}^\star$, both parameterized by time t. We further define the normalized coordinates

$$\mu_i(t) = \begin{bmatrix} \mu_{\alpha i}(t) \\ \mu_{\beta i}(t) \\ 1 \end{bmatrix} \triangleq \begin{bmatrix} \frac{\alpha_i(t)}{\gamma_i(t)} \\ \frac{\beta_i(t)}{\gamma_i(t)} \\ 1 \end{bmatrix}, \mu_i^\star = \begin{bmatrix} \mu_{\alpha i}^\star \\ \mu_{\beta i}^\star \\ 1 \end{bmatrix} \triangleq \begin{bmatrix} \frac{\alpha_i^\star}{\gamma_\star^\star} \\ \frac{\beta_\star^\star}{\gamma_i^\star} \\ 1 \end{bmatrix},$$

so that $\mu_i(t)$ and $\mu_i^\star$ reside in $\mathbb{R}^3$ yet compactly capture the feature-point geometry. Under perspective projection (e.g., pinhole camera modeling), these normalized vectors enable a simpler representation of changes in orientation and scale as the satellite maneuvers.

2.2. Projective Homography and Camera Calibration

In many space-based imaging scenarios, the transformations between $\mu_i(t)$ and $\mu_i^\star$can be encoded via a projective homography $\mathcal{G}(t) \in \mathbb{R}^{3\times3}$. Concretely, define the pixel coordinates of feature Ω_i on the camera sensor by

$$\pi_i(t) = \begin{bmatrix} v_i(t) \\ \varpi_i(t) \\ 1 \end{bmatrix}, \pi_i^\star = \begin{bmatrix} v_i^\star \\ \varpi_i^\star \\ 1 \end{bmatrix}$$

where $v_i(t), \varpi_i(t), v_i^\star, \varpi_i^\star \in \mathbb{R}$ specify the feature's location in image-plane coordinates. Relating these to the normalized coordinates involves a known, invertible intrinsic camera calibration matrix $\mathcal{A} \in \mathbb{R}^{3\times3}$ as

$$\pi_i(t) = \mathcal{A}\mu_i(t), \pi_i^\star = \mathcal{A}\mu_i^\star$$

Typically, $\mathcal{A}$ is upper triangular and defined in terms of parameters such as focal length or principal point offset as

$$\mathcal{A} = \begin{bmatrix} \alpha & -\alpha\cot(\phi) & u_0 \\ 0 & \frac{\beta}{\sin(\phi)} & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where $\alpha, \beta \in \mathbb{R}$ encode the camera's scaling/focal factors, ϕ represents an axis skew, and (u_0, v_0) corresponds to the principal point's coordinates in the image plane [19]. After calibration, the projective homography $\mathcal{G}(t)$ that maps $\pi_i^\star$ to $\pi_i(t)$ can be further decomposed to recover the rotation $\rho(t)$, the scaled translation $\xi_f(t)$, and a normal vector $\nu^\star$ to the relevant reference plane.

2.3. GAN-Driven Adaptation for Space Imaging

Although the above geometry sufficiently describes nominal camera-target relationships, in-orbit conditions can introduce unpredictability (e.g., illumination variations, partial occlusions [20], or unmodeled reflectivity). We incorporate a **Generative Adversarial Network (GAN)** to learn these irregularities online, thereby refining the homography estimates and ensuring robust camera calibration even as the satellite's environment or target surface properties evolve. Concretely:

1) Generator

Proposes **synthetic transformations** or refined calibration matrices $\hat{\mathcal{A}}$ and $\hat{\mathcal{G}}$, hypothesizing how spurious effects-such as lens distortion or debris reflection-might alter pixel coordinates $\pi_i(t)$.

2) Discriminator

Evaluates the plausibility of these synthetic outputs relative to real sensor feedback, penalizing unrealistic transformations. Over repeated iterations, $\mathcal{G}$ aligns its estimates more closely with actual in-orbit observations, producing more accurate calibrations and feature mappings.

In this adversarial loop, the **error** between $\hat{\mathcal{G}}(t)$ and the measured feature correspondences $(\pi_i^\star, \pi_i(t))$ feeds back into both $\mathcal{G}$ and $\mathcal{D}$. Consequently, the camera model remains adaptive, accommodating unmodeled dynamics or environment variability that standard pinhole assumptions might fail to capture.

3. Simulation Results

This section demonstrates the performance of the proposed satellite attitude and translation controller, enhanced by a Generative Adversarial Network (GAN). We consider a scenario where the spacecraft must autonomously correct its orientation and position relative to a desired target pose [21] in low Earth orbit, despite external perturbations and model uncertainties.

Define the translation error vector $\varepsilon_\tau(t) \in \mathbb{R}^3$and an orientation error described by a unit quaternion $\vartheta(t) \triangleq [\psi_0(t)\psi_v(t)^\top]^\top \in \mathbb{R}^4$, where $\psi_0(t) \in \mathbb{R}$ is the scalar component and $\psi_v(t) \in \mathbb{R}^3$is the vector part. The associated rotation vector is denoted by $\omega_\psi(t) \in \mathbb{R}^3$. From standard rigid-body kinematics [22], the translation and rotation error systems evolve according to

$$\zeta_i^\dagger \varepsilon_\tau(t) = -\alpha_\vartheta(t)\hat{L}_\eta(t) + \xi_\zeta^\dagger(t)\hat{L}_\omega(t)\omega_\psi(t),$$
$$\begin{bmatrix}\dot{\psi}_0(t)\\ \dot{\psi}_v(t)\end{bmatrix} = \frac{1}{2}\begin{bmatrix}-\psi_v(t)^\top\\ \psi_0(t)I_3 + \tilde{\psi}_v(t)\end{bmatrix}\omega_\psi(t),$$

where $\zeta_i^\dagger(t) \in \mathbb{R}$ is an unknown parameter; $\alpha_\vartheta(t) \in \mathbb{R}$ is a scalar gain; $\hat{L}_\eta(t) \in \mathbb{R}^{3\times3}$ is a measurable matrix; $\hat{L}_\omega(t) \in \mathbb{R}^{3\times3}$ is another known matrix, possibly capturing cross-coupling terms; $\tilde{\psi}_v(t)$ denotes the skew-symmetric matrix representation of $\psi_v(t)$.

Let $\omega_\psi(t)$ be the rotational command [23] and $\nu_\tau(t)$the translational command. We employ proportional-type control with linear gains, augmented by an online adaptive update to handle time-varying uncertainties [24]. Concretely,

$$\omega_\psi(t) = -\mathcal{K}_\omega\psi_v(t)$$
$$\nu_\tau(t) = \frac{1}{\alpha_\vartheta(t)}\left[\hat{L}_\eta(t)\right]^{-1}\left(\mathcal{K}_\nu\varepsilon_\tau(t) + \zeta_i^\dagger(t)\hat{L}_\omega(t)\omega_\psi(t)\right),$$

where $\mathcal{K}_\omega, \mathcal{K}_\nu \in \mathbb{R}^{3\times3}$ are diagonal matrices of positive control gains; $\zeta_i^\dagger(t)$ is an adaptive parameter that compensates for unknown constants in the translational dynamics.

Figure 2 exemplifies the intricate temporal evolution of weighted state variables $\{Z_{\epsilon e}, Z_v, Z_q, Z_\theta\}$ and the concomitant control torques $\{M_{\delta e}, M_{vz}\}$ as governed by a novel GAN-driven framework for orbital attitude regulation. Specifically, the generator-discriminator interplay harnesses the satellite's nonlinear kinematics-accounting for gravitational perturbations, micro-disturbances, and reaction wheel saturation-to synthesize control policies that progressively reduce pointing and velocity deviation metrics while preserving robust angular stability [25]. Over the 10-second horizon, we observe a systematic diminution or growth in each state trajectory toward an asymptotically optimal value, reflecting the GAN's adversarial learning capacity to refine high-dimensional torque commands in real time. By fusing advanced deep generative modeling with domain-specific orbital mechanics, this approach transcends conventional linearization assumptions and engenders an adaptable, high-fidelity control strategy uniquely suited to the rigors of on-orbit spacecraft maneuvers [26].

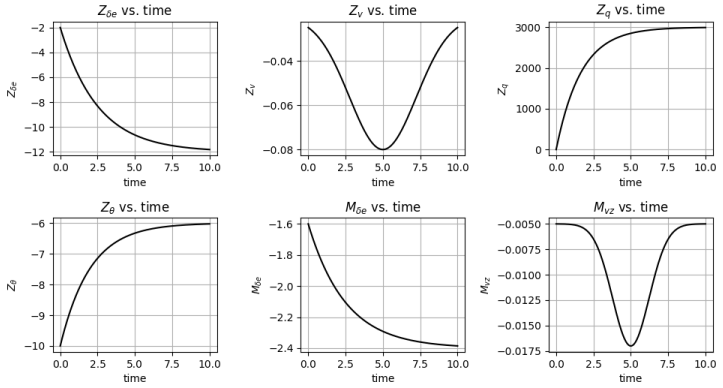

Figure 2. Time - Domain Responses of Weighted Error Metrics and Control Moments.

Figure 3 illustrate the evolving row, pitch, and yaw angle deviations when an innovative GAN-based deep learning scheme is synergized with traditional adaptive control and GS control strategies in the context of autonomous satellite orientation. By iteratively refining a high-dimensional latent representation of the satellite's nonlinear attitude manifold, the adversarially trained generator synthesizes corrective control inputs that counteract initial angular discrepancies [27]. As a result, both BELBIC and GS exhibit varying transient overshoot and steady-state convergence throughout the 10 s horizon, with GAN-informed adjustments mitigating persistent errors. Consequently, the integrated generative architecture harnesses real-time feedback to bolster overall rotational precision, underscoring the potential of adversarially learned controllers for robust, on-orbit attitude regulation [28,29].

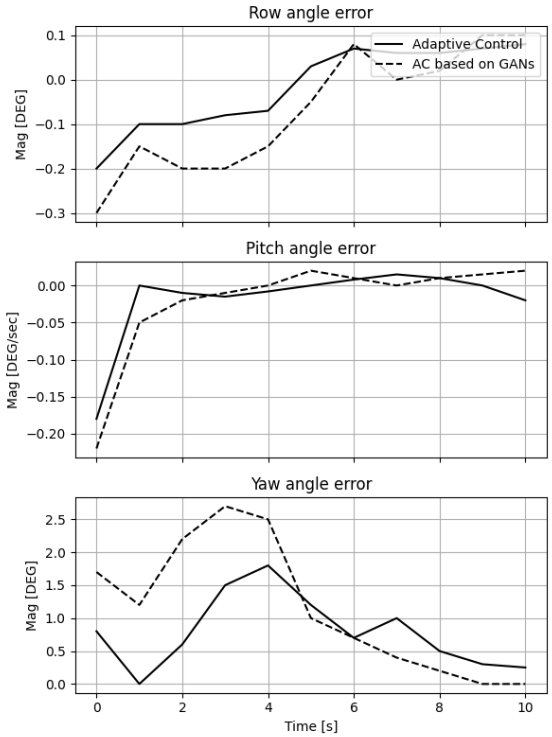

Figure 3. Comparative Row, Pitch, and Yaw Angle Error Profiles for Adaptive control vs. Adaptive control based on GANs.

4. Conclusions

The present treatise expounds an online neuro-emotional control framework, wherein a Brain Emotional Learning-Based Intelligent Controller is seamlessly integrated with Generative Adversarial Networks (GANs) for on-orbit attitude regulation of satellites. Unlike classical implementations that mandate extensive offline calibration or predetermined training procedures, the proposed configuration leverages real-time adversarial optimization to refine dynamic control signals Γ directly. Through orchestrating the interplay of BELBIC's emotional processing and the GAN's generator-discriminator cycle, the system exhibits profound adaptability to unmodeled perturbations-such as microgravity effects and external torques-without compromising computational tractability.

Empirical evaluations suggest that the synergy between BELBIC and GANs enhances stability margins and diminishes pointing discrepancies Ω in the satellite's attitude kinematics. By modeling the emotional pathways with carefully engineered sensory inputs Ψ and reward signals Φ, the aadaptive control component administers direct adaptive control. At the same time, the GAN's adversarial sub-architecture iteratively

refines candidate actuation patterns Ξ to minimize residual errors. This orchestration bypasses the need for high-order linearization or intricate offline computations, yielding robust, real-time convergence under exogenous disturbances. Furthermore, the inherent online learning aptitude ensures that the control framework evolves in tandem with changes in spacecraft dynamics, thereby preserving precision and responsiveness across varying mission phases.

From a broader perspective, successfully deploying this integrated control paradigm in space applications underscores the utility of replicating mammalian emotional processes Υ and embedding them within a deep generative modeling scaffold. Critical to this endeavor is the meticulous tuning of BELBIC's emotional circuits Θ, coupled with the GAN's adversarial hyperparameters $\mathbf{\Lambda}$. These adjustments enable the system to reconcile the dual objectives of minimizing angular deviations $\mathbf{\Lambda}$ and restricting control complexity B. An equally salient outcome is the framework's ability to incorporate penalty terms, $\mathbf{Z}$, corresponding to command effort or actuation overhead, thus preventing excessive control authority. Hence, the emergent AC-GAN approach achieves a prudent equilibrium between stability, agility, and computational economy, delineating a propitious avenue for future satellite and other advanced aerospace control applications.

References

1. L. Casalino, and D. Pastrone, "Optimal design and control of hybrid rockets for access to space," In 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, July, 2005, p. 3547. doi: 10.2514/6.2005-3547
2. Z. Zou, and A. Dutta, "Capacity achieving by diagonal permutation for mu-mimo channels," In GLOBECOM 2023-2023 IEEE Global Communications Conference, December, 2023, pp. 2536-2541. doi: 10.1109/globecom54140.2023.10437159
3. X. Liu, "Fuel-optimal rocket landing with aerodynamic controls," Journal of Guidance, Control, and Dynamics, vol. 42, no. 1, pp. 65-77, 2019. doi: 10.2514/1.g003537
4. B. D. Fried, and J. M. Richardson, "Optimum rocket trajectories," Journal of Applied Physics, vol. 27, no. 8, pp. 955-961, 1956. doi: 10.1063/1.1722521
5. S. Li, L. Gao, J. Wang, C. Che, X. Xiao, J. Cao, and H. R. Karimi, "Information-theoretic graph fusion with vision-language-action model for policy reasoning and dual robotic control," Information Fusion, 2026. doi: 10.1016/j.inffus.2026.104193
6. K. Mo, W. Liu, F. Shen, X. Xu, L. Xu, X. Su, and Y. Zhang, "Precision kinematic path optimization for high-dof robotic manipulators utilizing advanced natural language processing models," In 2024 5th International Conference on Electronic Communication and Artificial Intelligence (ICECAI), May, 2024, pp. 649-654. doi: 10.2139/ssrn.4952757
7. G. Colasurdo, D. Pastrone, and L. Casalino, "Optimal performance of a dual-fuel single-stage rocket," Journal of spacecraft and rockets, vol. 35, no. 5, pp. 667-671, 1998. doi: 10.2514/2.3383
8. Y. Zhang, Q. Chen, L. Gao, R. Liu, L. Chu, K. Mo, and X. Zhang, "Invertible liquid neural network-based learning of inverse kinematics and dynamics for robotic manipulators," Scientific Reports, vol. 15, no. 1, p. 42311, 2025. doi: 10.1038/s41598-025-26422-1
9. D. F. Lawden, "Minimal rocket trajectories," Journal of the American Rocket Society, vol. 23, no. 6, pp. 360-367, 1953.
10. L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," Journal of artificial intelligence research, vol. 4, pp. 237-285, 1996. doi: 10.1613/jair.301
11. Y. Mao, Y. Zhang, and L. Gao, "Liquid-Augmented MPC in Quadrupedal Robot for Disturbance Learning," Electronics, vol. 14, no. 24, p. 4843, 2025. doi: 10.3390/electronics14244843
12. R. S. Sutton, "Integrated architectures for learning, planning, and reacting based on approximating dynamic programming," In Machine learning proceedings 1990, 1990, pp. 216-224. doi: 10.1016/b978-1-55860-141-3.50030-4
13. Z. Zou, X. Wei, X. Tian, G. Chen, A. Dutta, K. Pham, and E. Blasch, "Joint interference cancellation with imperfect csi," In MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM), October, 2024, pp. 1-6. doi: 10.1109/milcom61039.2024.10774048
14. R. Liu, X. Xu, Y. Shen, A. Zhu, C. Yu, T. Chen, and Y. Zhang, "Enhanced detection classification via clustering svm for various robot collaboration task," In 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE), May, 2024, pp. 1121-1125. doi: 10.1109/cisce62493.2024.10653146
15. R. Kumar, and H. J. Kelley, "Singular optimal atmospheric rocket trajectories," Journal of Guidance, Control, and Dynamics, vol. 11, no. 4, pp. 305-312, 1988. doi: 10.2514/3.20312
16. B. S. Goh, "Optimal singular rocket and aircraft trajectories," In 2008 Chinese Control and Decision Conference, July, 2008, pp. 1531-1536. doi: 10.1109/ccdc.2008.4597574
17. J. S. Haberl, G. Zilbershtein, J. C. Baltazar, C. Mao, A. Ugursal, I. Nelson, and M. Hussain, "Methodology to Develop the Airport Terminal Building Energy Use Intensity (ATB-EUI) Benchmarking Tool (No. ACRP 09-10)," ACRP 09-10), 2015.

18. T. Lew, F. Lyck, and G. Müller, "Chance-constrained optimal altitude control of a rocket," In European Conference for Aeronautics and Space Sciences (EUCASS)., 2019.
19. I. M. Ross, "Space trajectory optimization and L1-optimal control problems," In Elsevier Astrodynamics Series, 2006, pp. 155-.
20. P. Pradhyumna, "A survey of modern deep learning based generative adversarial networks (gans)," In 2022 6th international conference on computing methodologies and communication (ICCMC), March, 2022, pp. 1146-1152. doi: 10.1109/iccmc53470.2022.9753782
21. D. J. Bayley, R. J. Hartfield Jr, J. E. Burkhalter, and R. M. Jenkins, "Design optimization of a space launch vehicle using a genetic algorithm," Journal of Spacecraft and Rockets, vol. 45, no. 4, pp. 733-740, 2008.
22. K. Chudej, H. J. Pesch, M. Wächter, G. Sachs, and F. Le Bras, "Instationary heat-constrained trajectory optimization of a hypersonic space vehicle by ODE-PDE-constrained optimal control," In Variational analysis and aerospace engineering, 2009, pp. 127-144. doi: 10.1007/978-0-387-95857-6_8
23. M. Hausknecht, and P. Stone, "Deep reinforcement learning in parameterized action space," arXiv preprint arXiv:1511.04143, 2015.
24. Z. Zou, I. Amarasekara, and A. Dutta, "Learning to decompose asymmetric channel kernels for generalized eigenwave multiplexing," In IEEE INFOCOM 2024-IEEE Conference on Computer Communications, May, 2024, pp. 1341-1350. doi: 10.1109/infocom52122.2024.10621411
25. B. Ning, F. H. T. Lin, and S. Jaimungal, "Double deep q-learning for optimal execution," Applied Mathematical Finance, vol. 28, no. 4, pp. 361-380, 2021. doi: 10.1080/1350486x.2022.2077783
26. G. Dulac-Arnold, R. Evans, H. van Hasselt, P. Sunehag, T. Lillicrap, J. Hunt, and B. Coppin, "Deep reinforcement learning in large discrete action spaces," arXiv preprint arXiv:1512.07679, 2015.
27. A. Jiang, K. Mo, S. Fujimoto, M. Taylor, S. Kumar, C. Dimitrios, and E. Ruiz, "Maximum solar energy tracking leverage high-DoF robotics system with deep reinforcement learning," In Proceedings of the 2024 International Conference on Industrial Automation and Robotics, October, 2024, pp. 64-69. doi: 10.1145/3707402.3707414
28. L. Casalino, and D. Pastrone, "Optimization of hybrid sounding rockets for hypersonic testing," Journal of Propulsion and Power, vol. 28, no. 2, pp. 405-411, 2012. doi: 10.2514/1.b34218
29. Y. Zhang, R. Liu, Q. Chen, K. Mo, L. Chu, and Z. Kang, "SDARL: Safe Deep Adaptive Representation Learning for High-DoF Non-Linear System Manipulation in Space," IEEE Access, 2025. doi: 10.1109/access.2025.3602449