End-to-End AI-based IP-GNC architecture for spacecraft proximity operations

Andrea Brandonisio*[1], Michele Bechini[1], Gaia Letizia Civardi[1], Lorenzo Capra[1], Michèle Lavagna[1]

[1] *Department of Aerospace Science and Technology, Politecnico di Milano, Via La Masa, 34, Milan, Italy*

Autonomy is increasingly crucial in space missions due to several factors driving the exploration and utilization of space. In the meanwhile, Artificial Intelligence (AI) methods begin to play a crucial role in addressing the challenges associated with and enhancing autonomy in space missions. The proposed work develops a closed-loop simulator for proximity operations scenarios, particularly for the inspection of an unknown and uncooperative target object, with a fully AI-based closed-loop image processing (IP) and Guidance Navigation & Control (GNC) chain. This tool is based on four main blocks: image generation, CNN-based image processing, navigation filter, and DRL-based guidance and control blocks. The proposed AI-based architecture is first trained and tuned to investigate the interface problems between each GNC block. Afterwards, the architecture is deployed in an Montecarlo testing campaign to verify and validate the performance of the proposed IP-GNC loop.

1 Introduction

Nowadays enhanced autonomy is the research driver of most of the leading space agencies, as spacecraft independence would allow for reliable, cost-effective, lower-risk services, and much more flexibility in mission planning, specifically concerning the wide spectrum of in-orbit servicing activities. A crucial aspect in these operations is the spacecraft's GNC (Guidance, Navigation and Control) autonomy, which shall ensure precise state reconstruction, together with trajectory generation and tracking. Moreover, with vision-based systems, image processing (IP) becomes extremely relevant for autonomously extrapolating the necessary information to perform state estimation. In recent years, the number of works on machine learning techniques applied to space-related problems is growing, especially in the context of guidance, navigation and control [1–3], affecting almost all the most important aspects of the GNC in a great variety of scenarios ranging from feasibility studies to on-board applications [4, 5]. Among the studied applications, autonomous relative pose estimation with known uncooperative targets is tackled in [6–8], where Convolutional Neural Networks (CNN) [7, 9, 10] The latter is then refined to perform both target detection and keypoint regression in a single inference in [11]. At the same time, Deep Reinforcement Learning (DRL), as a sub-discipline of machine learning, is gaining much interest, especially for autonomous guidance and control problems. Several scenarios have already been investigated: spacecrafts hovering or orbiting around small bodies, planetary landing and close-proximity operations [12–15]. According to [16], in the current state-of-the-art, DRL is a powerful tool when dealing with decision-making problems and its perfect compatibility with Artificial Neural Networks (ANN) as function approximators allows it to improve the generalising capabilities of the resulting policy, and to solve more and more complex problems characterised by high-dimensionality and continuous state and action spaces [17].

In this paper, we propose an innovative AI-based IP-GNC architecture, which combines a CNN image-based navigation tool with an autonomous DRL-based guidance and control (GC) agent, tested in a simulation environment with rendering-in-the-loop. In this architecture, the navigation block interfaces the AI-based pose estimation pipeline, based on the YOLOv8-pose-based architecture [11], and a relative navigation filter. The former estimates the chaser-target relative pose for each processed frame that is then passed to the navigation filter which outputs the overall estimation of the chaser-target relative state. This is then used by an autonomous guidance and control agent [18, 19], trained via Deep Reinforcement Learning to plan the chaser's relative trajectory around an uncooperative space object (i.e. the target) to optimise its inspection. In the following sections, the IP-GNC architecture is described, and its performance is investi-

*Corresponding author. E-Mail: andrea.brandonisio@polimi.it

gated in a TANGO relative dynamics scenario.

2 IP-GNC Architecture

This paper proposes a fully AI and image-based GNC algorithm working in a closed-loop environment employing an image generation step performed at each iteration. The testing environment is based on the five different blocks described in the following.

Dynamics. This block defines the chaser-target relative environment: the relative position and velocity are computed from the absolute position and velocity of both target and chaser integrated in the Earth-Centered Inertial (ECI) reference frame. The relative attitude dynamics is derived from a quaternion-based integration in the ECI reference frame.

Image Generation. The image generation (IG) block is entrusted with generating a synthetic representative frame for each GNC iteration. Each generated frame is fed to the IP block. The IG block comprises a synthetic image generation core based on POV-Ray [20] and a high-fidelity noise generator developed in [11]. Notably, the adopted image generator achieves an average rendering time of about 0.3 seconds per image, running only on the CPU of a general-purpose computer. Despite being low, the rendering time limits the possibilities of real-time simulation (i.e., simulated time equal to actual time in orbit) since, currently, it is not possible to ensure a rendering time equal to the image formation time or an actual spaceborne camera.

Image Processing. The IP block consists of an AI-aided keypoint-based pose estimation algorithm that leverages a single, multi-tasking, and efficient CNN that performs both target detection and keypoint regression in a single inference. Namely, the pipeline leverages the YOLOv8s-pose model due to its efficiency and top-level performances, while the relative pose is estimated by using the EPnP algorithm [21], leveraging the knowledge of the 3D keypoints from the CAD model, the 2D keypoint regressed by the YOLOv8s-pose model, and the 2D-to-3D correspondences retrievable from the regressed keypoint fixed order, without any need of depth estimation from images. This architecture is introduced in [11] to process each image generated by the IG block. Such architecture is capable of of handling the relative pose estimation tasks for several target geometries, by providing accurate estimates thanks to the outlier rejection scheme based on the confidence scores associated with the retrieved landmarks. The YOLOv8s-pose has been trained on the MINIMA dataset [22] using the stochastic gradient descent method. The criteria adopted to evaluate the performances of the YOLOv8s-pose are the Intersection-over-Union (IoU) and the average precision (AP) for the object detection performances. The trained model scored a mean AP_{50}^{95} of 97.4% and a mean IoU of 96.11% for the object detection task, while for the keypoint regression task, it achieved a mean AP_{50}^{95} of 99.1% and a mean Object Keypoint Similarity Index (OKS) of 97.99% on MINIMA images. Fig. 1 provides examples of detected ROI and keypoints extracted from sample images using the trained YOLOv8s-pose model. The outputs of

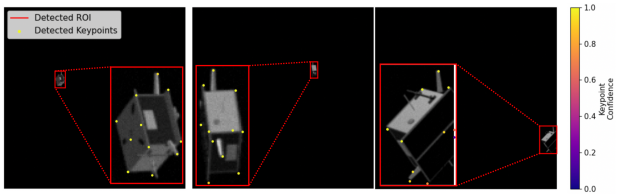

Figure 1: Detected ROI, keypoints, and confidence score using the trained YOLOv8s-pose model on sample images.

the IP block are the relative target position (r) and attitude (q) with respect to the chaser camera frame. Notice that, without any loss of generality, here it is assumed that the camera reference frame corresponds to the chaser body reference frame.

Navigation. The navigation (NAV) block is based on a loosely coupled architecture involving a relative translational filter and a relative rotational filter. The relative position and velocity estimation is entrusted to a H-∞ filter, which is a robust linear estimator. Regarding the relative rotational filter, the Invariant Extended Kalman Filter (IEKF) is exploited here. The input of the filter is the relative chaser-target angular velocity inferred from the chaser's absolute navigation, the earth-based observation of the target's rotational state, and the relative quaternion retrieved from the IP block. The output of the IEKF is a refined relative quaternion.

Guidance and Control. The Guidance and Control (GC) is the last block of the pipeline and is based on [23]. It takes as input the relative chaser-target position and velocity in the target LVLH reference frame and the relative attitude of the target in the chaser camera frame. Proximal Policy Optimization (PPO), a state-of-the-art DRL algorithm for continuous state-action spaces, is then employed to optimize the trajectory of the spacecraft around the target, according to the specified reward function, which shapes the decision-making of the agent. The DRL policy is

first trained on a simplified environment, where the camera is assumed to be always pointed towards the target and information coming from the navigation block are perfectly known. The stochastic policy is optimized throughout the training process by adjusting the parameters of the neural network that approximates the spacecraft's decision-making to maximize the reward function. The latter is designed to incentivize the spacecraft to remain in a bounded region of space around the target and to improve its mapping, as implemented in [23], in a limited time window and maximum propellant availability condition. The GC block outputs the control action of the chaser (shaped as an acceleration vector) that affects both the dynamics and navigation blocks. This methodology is strongly influenced by the target object (in this case, TANGO from PRISMA mission), on which the agent is trained, from the camera model with its FOV and the trajectory range which again shapes the training simulation. Fig. 2 shows the architecture pipeline described above.

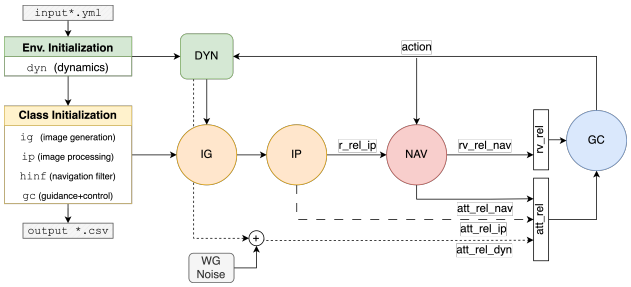

Figure 2: AI-based GNC pipeline architecture scheme.

3 Results on TANGO case study

The baseline simulation on which the AI and image-based GNC pipeline is based on some main assumptions that must be considered due to the input constraints of the different blocks:

- The **Earth** is never in the **background** of the image for two main reasons: the guidance reward model was not built considering Earth illumination and the IG block would have been excessively slow in rendering the images to maintain a feasible simulation time. Nevertheless, it is worth mentioning that the analyses related to the robustness of the IP step pointed out that the presence of backgrounds has negligible effects on the estimated relative pose accuracy [11].

- The guidance and control agent inserted in the GNC pipeline is trained on a **fixed target pointing** scenario, which at first, does not require any attitude controller.

- The **trajectory** is **limited** to remain between 1.5-35 m from the target to avoid collisions or over-reducing the target size in pixels, leading to unfeasible target detection and relative pose estimation steps within the IP block.

Starting from these assumptions on the different blocks, the testing campaign focused, at first, on the input of the chaser-target relative orientation, which could have been derived in three different ways: real quaternion value (computed by the dynamics block) with the addition of a proper level of noise, the IP estimation, or adding a rotational navigation filter to deal with this information. Once this was assessed, the second aspect was related to attitude control, removing the assumption of fixed target pointing. Here, only the most general and final case results will be presented, even if a testing workflow has been followed to increase the complexity of the simulations step by step, without rushing directly to the most general case. The stopping conditions defined for the simulation are taken from the GC agent's development: exceeding the trajectory limits (range between 1.5-35 m) or the time window (set to a maximum of 4 hours of trajectory length), and the achievement of map coverage (if reaches 100%). This small Montecarlo analysis divided in steps of increasing difficulty was able to assess the potential of the treated methodology. In the following the two most important aspects constraining the simulation are briefly discussed.

GC attitude input The attitude information needed by the GC agent can be derived directly by the dynamics, the IP block or by the rotational filter. In the first case, the estimation is less realistic because only noise on the real quaternion is being considered. In the second and third cases, the input is more proper since it comes from the IP or NAV estimation. In the following results, the simulations are based on the NAV estimation case, which is considered as the most realistic and difficult case.

Chaser attitude control In the most general case of unfixed target pointing, a PD control has been exploited to compute the target pointing with the quaternion estimator algorithm (QUEST) [24]. The proportional and derivative gains are not fine-tuned on the chaser to leave a small level of uncertainty on the attitude control to analyse the robustness of the GC; also the coupling between the trajectory and attitude control is not taken into account.

3.1 IP Analysis

The IP block robustness is assessed by collecting the outputs for all the episodes simulated, resulting in about 100000 samples retrieved for both the fixed target pointing and controlled pointing conditions. The performances have been evaluated by retrieving for each frame the normalized relative distance error and the relative quaternion error as defined in [25]. Some error metrics (on relative position and quaternion error) are shown in Fig. 3, where the peaks of maximum error between 10 m and 16 m are due to a weak attitude control action that makes the target exit the FOV (field-of-view). The high errors correspond to the transition from inside to outside of the FOV and vice versa. Overall, the normalized relative position error is below 5%. Similar results are obtained for the relative quaternion error which mostly grows when the relative distance is higher than 32 m (distance limit for target detection).

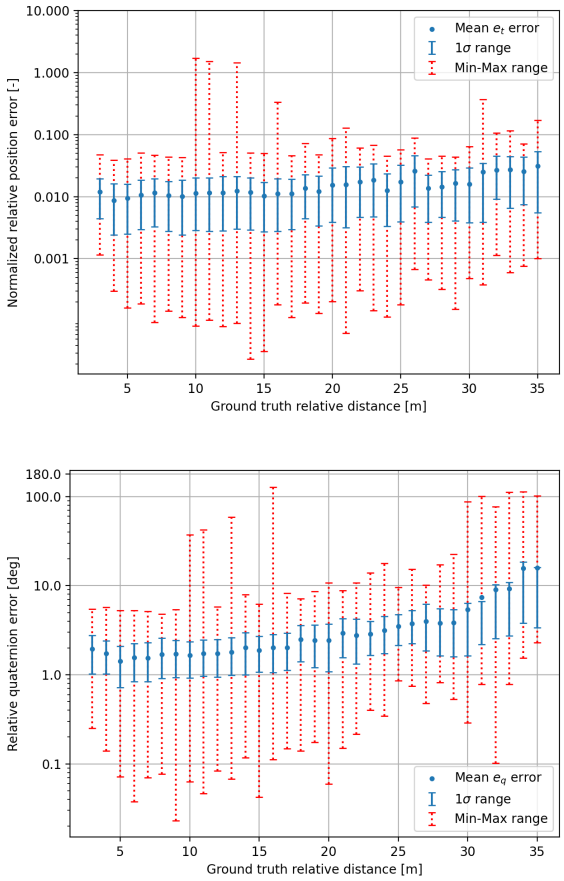

Figure 3: YOLOv8s-pose-based IP robustness analysis in non-target-pointing scenario.

NAV Analysis The output of the NAV functional block is tightly related to the IP one, and the specific estimation error values depend on the episode configuration. Also for the NAV block, it is possible to notice that the percentage of image-based outlier measurements is extremely low, and there are no failures within the IP block. It can be noticed that the filter is beneficial in both cases, since it reduces the estimation error of the IP functional block. It can also be noticed that the presence of the filter is effective in mitigating the error peaks that occur during the episode. The overall position and attitude estimation decrease by 30% and 7%, respectively. Even if the position error reduction is higher in percentage, it is important to stress the fact that its magnitude is always extremely low, always below 5% of the relative inter-satellite range. The relative position and attitude estimation errors are shown in Fig. 4 and Fig. 5, respectively.

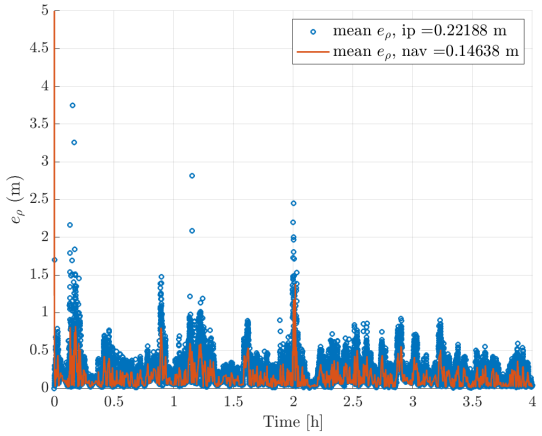

Figure 4: Relative translation filter estimation errors in non-target-pointing scenario, off-nominal IP behavior.

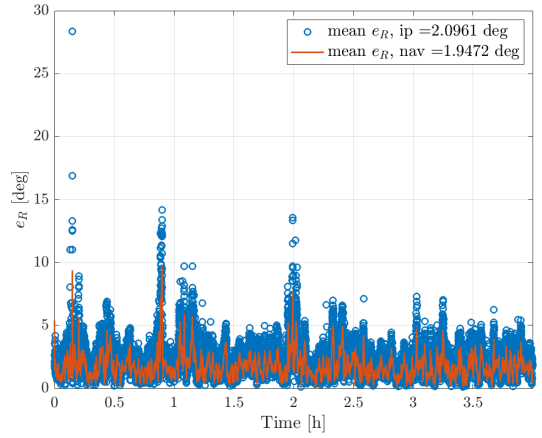

Figure 5: Relative rotation filter estimation errors in non-target-pointing scenario, off-nominal IP behavior.

3.2 GC Analysis

The performance results of the GC agent remain constant during all the tests, proving the great robustness of the trained model. The resulting evaluation metrics for the most general case (input from NAV block and chaser attitude control, as defined before) are shown in Tab. 1 and respect the standards of the nominal performance detected in [23].

Table 1: GNC pipeline results.

Reward Scr.	Map	Traj. time	Thrust level
574.2	96.5 %	4.0 h	53.2 %

In this case, all episodes end exceeding the time constraint of 4h, even if achieving a very high mapping level. But, this result shows the benefits of the rotational filter which reduces the average error of the GC pointing information input, in Tab. 2.

Table 2: Average pointing error in GC input throughout the different step of the GNC pipeline tests.

Pointing	DYN+noise	IP+noise	NAV
Avg. Point. Err.	10-20°	10-20°	5-10°

An example of the results acquired throughout the test analysis is shown in Fig. 6, in terms of chaser-target relative position (on top) and velocity (in the middle). Looking at the trajectory plot, it is observed how the agent tries to keep the spacecraft at an average relative position and velocity value, actively controlling it (red dots) only when these two parameters start to drift away.

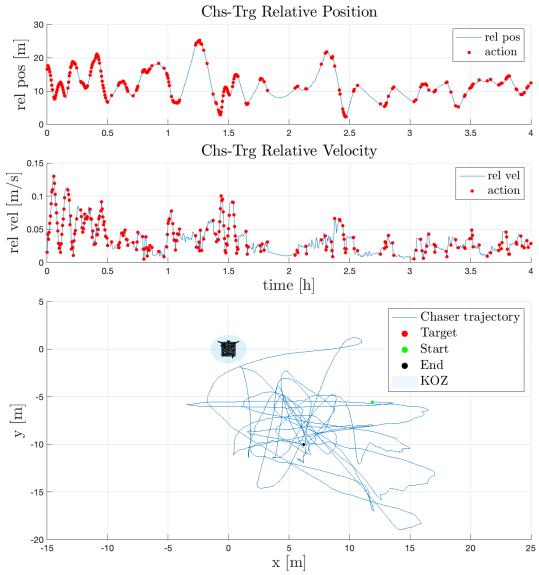

Figure 6: Example of Chaser-Target trajectory and evaluation metrics in the *controlled* target pointing case.

4 Discussion

This paper proposes an innovative GNC algorithm architecture which has been shortly presented in all of its components and their relative performances. In conclusion, the GC-trained agent exhibits a stable and unaltered level of performance throughout all the steps of the testing; moreover, the policy learned is robust to all the uncertainties and model differences introduced in the new GNC framework. The IP block, based on YOLOv8s-pose-based relative pose estimation algorithm, is recognised to be highly accurate. Therefore, in general, the outcomes confirmed the strong robustness of the proposed approach, where the pose estimation pipeline is not affected by the position of the target within the FOV. The navigation filters downstream the IP functional block enhance the robustness of the GNC chain, being beneficial to the accuracy of the pose estimate and to enabling operations even in case of faulty IP measurements, for example when the target exits the camera FOV. Anyway, some points still remain open, e.g. generating high-fidelity images with calibrated noise components in real time, coupling compatibility between trajectory and attitude control, absolute attitude knowledge assumption, etc., but these results have strengthened the awareness of having a powerful tool, that, at this point, feels the need to be tested also in a realistic or hardware-in-the-loop environment.

References

1. Izzo, D., Märtens, M. & Pan, B. A survey on artificial intelligence trends in spacecraft guidance dynamics and control. *Astrodynamics* **3,** 287–299 (2019).
2. Tipaldi, M., Iervolino, R. & Massenio, P. R. Reinforcement learning in spacecraft control applications: Advances, prospects, and challenges. *Annual Reviews in Control* **54,** 1–23. ISSN: 1367-5788. https://www.sciencedirect.com/science/article/pii/S136757882200089X (2022).
3. Hovell, K. & Ulrich, S. Deep Reinforcement Learning for Spacecraft Proximity Operations Guidance. *Journal of Spacecraft and Rockets* **58,** 254–264 (2021).
4. Silvestrini, S. *et al.* in *Modern Spacecraft Guidance, Navigation, and Control* (eds Pesce, V., Colagrossi, A. & Silvestrini, S.) 819–981 (Elsevier, 2023). ISBN: 978-0-323-90916-7.
5. Izzo, D., Meoni, G., Gómez, P., Dold, D. & Zoechbauer, A. *Selected Trends in Artificial Intelligence for Space Applications* 2022. arXiv: 2212.06662 [cs.LG].

6. Kisantal, M. *et al.* Satellite pose estimation challenge: Dataset, competition design, and results. *IEEE Transactions on Aerospace and Electronic Systems* **56,** 4083–4098 (2020).
7. Park, T. H. *et al.* Satellite Pose Estimation Competition 2021: Results and Analyses. *Acta Astronautica* **204,** 640–665. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576523000048 (2023).
8. Chen, B., Cao, J., Parra, A. & Chin, T.-J. *Satellite pose estimation with deep landmark regression and nonlinear pose refinement* in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops* (2019), 0–0.
9. Pauly, L. *et al.* A survey on deep learning-based monocular spacecraft pose estimation: Current state, limitations and prospects. *Acta Astronautica* **212,** 339–360. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576523003995 (2023).
10. Bechini, M., Gu, G., Lunghi, P. & Lavagna, M. Robust spacecraft relative pose estimation via CNN-aided line segments detection in monocular images. *Acta Astronautica* **215,** 20–43. ISSN: 0094-5765 (2024).
11. Bechini, M. *Monocular Vision for Uncooperative Targets through AI-Based Methods and Sensors Fusion* PhD thesis (Politecnico di Milano, 2024).
12. Gaudet, B., Linares, R. & Furfaro, R. Adaptive guidance and integrated navigation with reinforcement meta-learning. *Acta Astronautica* **169,** 180–190. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576520300072 (2020).
13. Gaudet, B., Linares, R. & Furfaro, R. Six degree-of-freedom body-fixed hovering over unmapped asteroids via LIDAR altimetry and reinforcement meta-learning. *Acta Astronautica* **172,** 90–99. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576520301545 (2020).
14. Gaudet, B., Linares, R. & Furfaro, R. Deep reinforcement learning for six degree-of-freedom planetary landing. *Advances in Space Research* **65,** 1723–1741. ISSN: 0273-1177 (2020).
15. Scorsoglio, A., Furfaro, R., Linares, R. & Massari, M. Relative motion guidance for near-rectilinear lunar orbits with path constraints via actor-critic reinforcement learning. *Advances in Space Research* **71,** 316–335. ISSN: 0273-1177 (2023).
16. Silvestrini, S. & Lavagna, M. Deep learning and artificial neural networks for spacecraft dynamics, navigation and control. *Drones* **6,** 270 (2022).
17. Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction* 2nd ed. (MIT press, Cambridge, Massachusetts, USA, 2018).
18. Brandonisio, A., Capra, L. & Lavagna, M. Deep reinforcement learning spacecraft guidance with state uncertainty for autonomous shape reconstruction of uncooperative target. *Advances in Space Research.* ISSN: 0273-1177 (2023).
19. Capra, L., Brandonisio, A. & Lavagna, M. Network architecture and action space analysis for deep reinforcement learning towards spacecraft autonomous guidance. *Advances in Space Research* **71,** 3787–3802. ISSN: 0273-1177 (2023).
20. Plachetka, T. *POV Ray: Persistence of Vision parallel raytracer* in *Proc. of Spring Conf. on Computer Graphics, Budmerice, Slovakia* **123** (1998), 129.
21. Lepetit, V., Moreno-Noguer, F. & Fua, P. EPnP: An Accurate O(n) Solution to the PnP Problem. *International Journal of Computer Vision* **81,** 155–166. ISSN: 1573-1405 (Feb. 2009).
22. Bechini, M., Lavagna, M. & Lunghi, P. Dataset generation and validation for spacecraft pose estimation via monocular images processing. *Acta Astronautica* **204,** 358–369. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576523000127 (2023).
23. Brandonisio, A. *AI-based Guidance for Spacecraft Proximity Operations around Uncooperative Targets* PhD thesis (Politecnico di Milano, 2024).
24. Crassidis, J. L., Markley, F. L. & Cheng, Y. Survey of Nonlinear Attitude Estimation Methods. *Journal of Guidance, Control, and Dynamics* **30,** 12–28. https://doi.org/10.2514/1.22452 (2007).
25. Sharma, S., Ventura, J. & D'Amico, S. Robust Model-Based Monocular Pose Initialization for Noncooperative Spacecraft Rendezvous. *Journal of Spacecraft and Rockets* **55,** 1414–1429. https://doi.org/10.2514/1.A34124 (2018).