ALLO: A Photorealistic Dataset and Data Generation Pipeline for Anomaly Detection During Proximity Operations in Lunar Orbit

Selina Leveugle[1], Chang Won Lee[2], Svetlana Stolpner[3], Chris Langley[3], Paul Grouchy[3], Steven Waslander[2], and Jonathan Kelly[1]

***Abstract*— NASA's Artemis missions are paving the way for the return of sustained human operations on and around the Moon. A key challenge for extended space missions is the unprecedented level of autonomy required for system control. This need arises from communication delays and limited human oversight. A critical component of autonomous operation is the ability of external robotic systems—such as future versions of the Canadarm2 currently operating on the International Space Station—to detect hazards from onboard imagery under extreme and variable lighting conditions. Traditional anomaly detection methods, which rely heavily on large labelled datasets, are not well suited for this setting due to the scarcity of representative imagery and the difficulty of obtaining ground-truth annotations. To address this gap, we introduce ALLO (Anomaly Detection in Lunar Orbit), a novel photorealistic synthetic dataset generated using Blender. The dataset includes 36,385 anomaly-free images and 15,024 anomalous images representing robotic-arm operations in lunar orbit, with realistic variations in illumination, pose, and background. We contribute (i) a photorealistic dataset for anomaly detection, (ii) an open-source Blender-based data generation pipeline for scalable dataset creation, and (iii) a qualitative validation of the rendered imagery against real photographs from the International Space Station (ISS). By releasing both ALLO and its data generation pipeline, we provide a tool for creating realistic space datasets and for developing and evaluating anomaly detection and localization methods for space applications. The dataset and code are available at: https://github.com/utiasSTARS/ALLO.git.**

I. Introduction

Over the past two decades, the international space community has increasingly focused on extending human space exploration beyond low-Earth orbit. NASA's Artemis program, for example, aims to establish a sustained human presence on the Moon and to test technologies required for future deep-space missions [1]. In support of this effort, new systems are being developed to operate in lunar orbit. Unlike the International Space Station (ISS), lunar-orbiting platforms are expected to function with a high degree of autonomy and without an onboard, human crew for extended periods during surface missions. The large distance between Earth and these spacecraft, combined with their orbital dynamics, limits the availability of fast and reliable communication for real-time control [2]. As a result, autonomy is particularly critical for systems such as Canadarm3, an advanced external robotic system being developed by MDA Space Ltd. in collaboration with the Canadian Space Agency. Canadarm3 is intended to support a range of on-orbit servicing, inspection, and vehicle capture operations for future lunar-orbiting platforms [3].

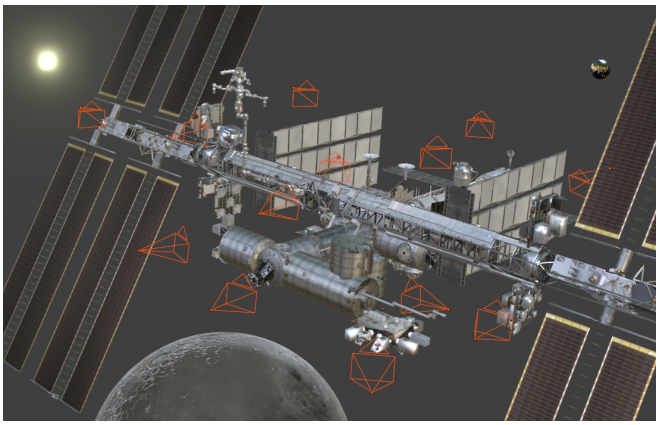


Fig. 1: Blender rendering of the ISS model used in the ALLO dataset. The Earth, the Moon, and the Sun are shown in the background; reference camera positions around the station are highlighted in orange.

Of particular interest are vision-based solutions that use the arm's cameras to autonomously detect potential collision hazards, such as loose tools or debris, thereby reducing the risk of damage during autonomous operations. Since the type and appearance of all possible hazards cannot be known in advance, it is desirable for the vision system to perform *anomaly detection*, that is, recognizing off-nominal situations that differ from expected operating conditions.

Anomaly detection and localization involve identifying and pinpointing unexpected content in images that deviates from a set of expected inputs, such as debris, foreign objects, or structural irregularities. Such methods aim to detect pixels, features, or objects that differ statistically from an established baseline [4]. In the space domain, this task is particularly challenging due to high-contrast lighting conditions arising from the absence of atmospheric scattering, harsh direct solar illumination, and large variations in viewpoints during robotic operations. These factors can reduce visual contrast and make reliable detection difficult.

Although reliable anomaly detection would be highly valuable for reducing the risk of collisions involving Canadarm3, current methods and datasets are not tailored to this setting.

[1]Space and Terrestrial Autonomous Robotic Systems (STARS) Laboratory at the University of Toronto Institute for Aerospace Studies (UTIAS), Toronto, Canada. Email: <firstname>.<lastname>@robotics.utias.utoronto.ca

[2]Toronto Robotics and AI Laboratory (TRAIL) at the University of Toronto Institute for Aerospace Studies (UTIAS), Toronto, Canada. Email: john.lee@robotics.utias.utoronto.ca; steven.waslander@robotics.utias.utoronto.ca

[3]MDA Space Inc. Email: <firstname>.<lastname>@mda.space

Available space image datasets from on-orbit operations, such as imagery from Canadarm2 on the ISS, are limited and insufficient for developing anomaly detection algorithms for lunar-orbiting systems. In particular, images from past missions often lack ground-truth annotations and are not representative of operations in lunar orbit. To bridge this gap, we introduce an open-source dataset for anomaly detection and localization in robotic proximity operations in lunar orbit. Our main contributions are as follows.

- We develop a simulator and automated data generation pipeline using Blender's Cycles rendering engine.
- We introduce the ALLO (Anomaly Localization in Lunar Orbit) dataset, an open-source dataset for anomaly detection in vision-based proximity operations for space robotic systems. The dataset comprises 36,385 anomaly-free images and 15,024 anomalous images with pixel-level ground-truth maps.
- We validate the fidelity of the rendering process by reconstructing images captured in Earth orbit using ISS camera configurations and qualitatively comparing them to real imagery.
- We release both the ALLO dataset and the data generation pipeline to support the development of vision-based algorithms for space operations.

The remainder of this paper is structured as follows. Section II reviews existing space imagery and anomaly detection datasets. Section III details the Blender model used to simulate spacecraft operations in lunar orbit. Section IV describes the image rendering process used to generate the ALLO dataset and Section V evaluates the fidelity of the resulting images. The ALLO dataset and the data generation pipeline are available at https://github.com/utiasSTARS/ALLO.git.

II. Related Work

Although visual anomaly detection has not yet been applied to the space domain, there is substantial prior work on synthetic vision-based navigation datasets for space exploration and, separately, on anomaly detection for industrial inspection. In this section, we review existing visual datasets for space missions, followed by commonly used benchmarks for visual anomaly detection.

A. Datasets for Space Operations

Vision-based algorithms are increasingly used in space exploration missions, including planetary landings and on-orbit operations, because they enable autonomous navigation and perception. Developing and validating such algorithms requires accurate and representative images of the anticipated deployment scenarios. Images from past space missions are usually unsuitable for this purpose because they may not reflect the conditions of a new mission, may lack ground truth references, or may have incorrect sensor configurations [5], [6]. In some cases, acquiring new real images is impossible or inadequate because relevant environmental conditions cannot be reproduced on Earth [7].

As a result, computer-generated images are frequently used to prototype and evaluate vision-based algorithms. Images are typically created using rendering engines that apply path tracing to simulate light transport. Rendering programs rely on engines such as Cycles [8] or Unreal Engine 5 [9] to replicate anticipated mission environments.

NASA developed the DLES Unreal Simulation Tool (DUST) [10] to generate scenes of the lunar south pole in support of the Artemis missions. DUST uses lunar terrain information from the Digital Lunar Exploration Sites (DLES) dataset with Unreal Engine 5 rendering. Similarly, ESA developed the Planet and Asteroid Natural Scene Generation Utility (PANGU) [11] to create images of planetary surfaces. Unlike DUST, which focuses on supporting the development of lunar surface missions, PANGU emphasizes views of a planetary body during spacecraft descent, with particular attention to cratered terrain such as that found on the Moon and Mars.

For on-orbit applications, Airbus designed SurRender [5], a rendering framework for vision-based guidance, navigation, and control tasks such as satellite servicing and debris-removal operations. SurRender uses a domain-specific language derived from C to model projection, geometry, material properties, and reflectance. The Space Imaging Simulator for Proximity Operations (SISPO) [7] was developed to simulate asteroid imagery for proximity and fly-by missions. SISPO leverages Blender's Cycles engine to calculate the illumination of the simulated scene. Specifically, the software aims to model challenging phenomena such as gas and dust around asteroids.

Some rendering programs incorporate additional sensing modalities with the rendered images to extend possible applications. The Navigation and Rendering Pipeline for Astronautics (NaRPA) [12] was devised for spacecraft navigation around non-Earth celestial bodies and includes modalities such as LiDAR point clouds and stereo depth. The TANGO dataset [6] uses a CAD model of the TANGO spacecraft from the 2010 PRISMA satellite project for autonomous formation flying to generate a synthetic dataset for monocular spacecraft pose estimation, together with ground truth pose labels.

Overall, current image generation methods for vision-based algorithms in the space domain are largely bespoke, mission-specific, and tailored to tasks such as landing, pose estimation, or asteroid proximity operations. As a result, the generated imagery is often are not readily transferable to robotic inspection tasks. In addition, many of these existing tools are not open source, which limits adaptation to new mission scenarios. These limitations motivate the development of a rendering pipeline and dataset tailored to anomaly detection for worksite surveying by the Canadarm3.

B. Datasets for Industrial Defect Inspection

Most recent progress in visual anomaly detection has been driven by industrial inspection benchmarks. The 2D MVTec anomaly detection dataset [13] is the most widely used benchmark for this application due to its pixel-level

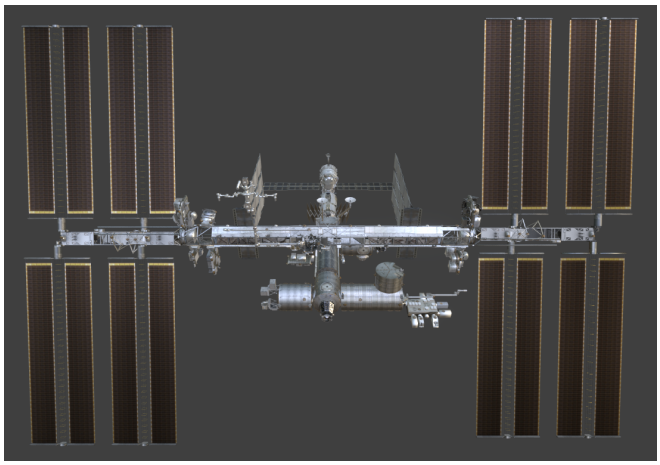


Fig. 2: Blender model of the ISS used in the ALLO dataset.

annotations and diverse range of objects and textures [14]. Other commonly used datasets, such as BTech [15] and Kolektor [16], also focus on defect detection in controlled inspection settings. However, these datasets typically involve fixed viewpoints, relatively consistent lighting, and simple backgrounds. In contrast, anomaly detection for space-based robotic inspection must contend with substantial viewpoint variation, highly nonuniform illumination, and complex backgrounds. These characteristics are not well represented in existing anomaly detection benchmarks, which motivates a dedicated dataset for anomaly detection in space operations.

III. Blender Model Setup

Images acquired by robotic manipulators such as Canadarm3, whether on future lunar-orbiting platforms or other spacecraft, are expected to exhibit significant variability in lighting conditions. The variability arises from several sources, including the dark background of space, direct solar illumination, indirect reflections from the lunar surface and the station structure, and artificial illumination provided by the arm's inspection lights. Accurately reproducing these lighting conditions is essential for generating realistic training data for visual anomaly detection. Blender was chosen to create the ALLO dataset and replicate these lighting conditions because its Cycles rendering engine uses path tracing to calculate highly realistic lighting.

In Blender, the operation of a space station around the Moon was simulated, with the Earth and Sun in the background, using NASA's Blender model of the ISS [17], shown in Figure 2. The Lunar Gateway's orbit around the moon was approximated in Blender using an elliptical trajectory with parameters corresponding to the station's expected perilune and apolune [18]. To simulate the evolving lighting conditions throughout the mission, the relative positions of the Earth and sun with respect to the moon were computed using the Skyfield astronomical library [19]. Ephemeris data from the year 2030 was used to generate physically accurate positions for these celestial bodies. The positions of the Moon, Earth, and Sun relative to the station were then simulated over a full 365-day period, allowing the dataset to capture the wide range of lighting and background configurations expected during operation.

TABLE I: Blender camera parameters used to render the images in the ALLO dataset.

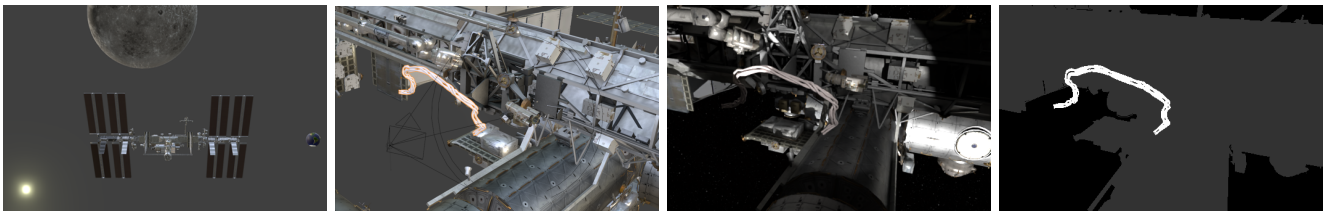
Camera Parameter	Value
Camera Type	Perspective
Focal Length	25 mm
Horizontal Sensor Size	36 mm
Vertical Sensor Size	24 mm

To simulate the viewpoints of the arm-mounted cameras, 50 unique base camera poses were manually defined around the station model. These poses correspond to representative locations along potential trajectories of the Canadarm3 near the arm's external mounting fixtures. To introduce realistic variability in camera placement and improve viewpoint diversity, each base pose was randomly perturbed during rendering. Gaussian noise was applied to both the camera position and orientation, with standard deviations of 1 m for translational offsets and 0.2 rad for rotational offsets (roll, pitch, and yaw). This perturbation strategy increases the diversity of viewpoints while maintaining coverage of key station regions relevant to inspection tasks. The station model and some of the reference camera positions are shown in Fig. 1; the Blender camera parameters are listed in Table I.

IV. Image Rendering

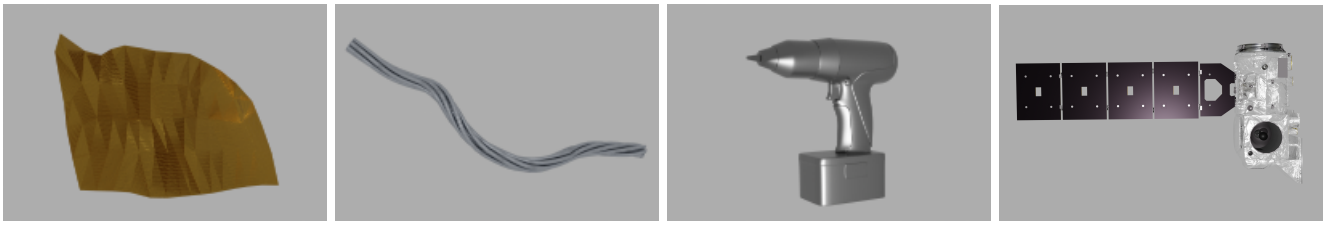
Different camera poses are used for the training and test sets to prevent overlap between viewpoints and reduce the likelihood of a learned model memorizing specific station perspectives. Only the test set includes scenes containing anomalous objects that represent potential collision hazards for the robotic manipulator. A total of 16 distinct anomalous object models are used (see Fig. 4). These objects were selected to reflect plausible hazards, including dislodged station components (e.g., thermal blankets or loose cables), items that may be lost during extravehicular activities (e.g., astronaut tools or gloves), and objects resembling experimental hardware from prior missions.

For each anomalous scene, only one such object is inserted into the camera's view frustum to simplify evaluation. The object is randomly positioned within the frustum in the vicinity of the station structure. To introduce variation in apparent object size and depth, the object is rendered at five different distances from the camera: an initial randomly sampled depth and four additional depths obtained by displacing the object by ± 1 m and ± 2 m along the viewing direction. To ensure that anomalies remain visible and contribute meaningful supervision during evaluation, additional placement constraints are applied. If the inserted anomaly becomes totally occluded by the station geometry or occupies less than 0.1% of the rendered image area after projection, the object is repositioned until these visibility criteria are satisfied. This procedure produces a diverse range of anomaly appearances while maintaining physically plausible object placements near the station.



(a) Scene setup. (b) Adding camera and anomaly. (c) Anomalous image render. (d) Segmentation mask.

Fig. 3: Rendering process for an anomalous image in the ALLO dataset. The setup of the station and celestial bodies are shown in 3a, the placement of the camera and anomalous cable are shown in 3b, and the rendered anomalous image and the corresponding segmentation mask are presented in 3c and 3d respectively.



(a) Thermal blanket (b) Cable (c) Drill (d) Experimental Satellite

Fig. 4: Examples of models of anomalous objects used in the ALLO dataset.

All rendered images have a resolution of 1,920 × 1,080 pixels. Each image is accompanied by a three-class segmentation mask that labels pixels as background, non-anomalous foreground (including station structures and celestial bodies), or anomalous regions corresponding to the inserted anomalous objects. The complete rendering pipeline, illustrated in Fig. 3, is executed twice using different random seeds to increase dataset diversity. Example images from the resulting dataset are shown in Fig. 5 and Fig. 6.

In summary, scenes are generated for each day represented in the ephemeris dataset by first positioning the Moon, Earth, and Sun according to their computed celestial coordinates. Camera poses are then selected from the predefined set, with indices 0–39 reserved for normal scenes and indices 40–49 used for anomalous scenes. For each camera pose and anomaly configuration, the following procedure is executed:

1) Place the camera at the selected pose and apply Gaussian perturbations to its position and orientation.
2) Add a nearby spotlight to simulate the arm-mounted inspection light.
3) If the scene is anomalous:
    a) Insert a randomly selected anomalous object into the camera view frustum.
    b) Reposition the object if necessary to ensure it remains visible and sufficiently large in the image.
4) Apply simulated image artifacts, including sensor noise and lens glare.
5) Render the image and the corresponding segmentation mask using the Blender Cycles rendering engine.

A total of 51,409 images were rendered and the breakdown of normal and anomalous images in the training and test sets is shown in Table II

TABLE II: Composition of the ALLO dataset (camera poses and images) across the training and test sets. The training set is composed exclusively of normal images, while the test set contains both normal and anomalous images rendered from disjoint sets of camera poses.

Dataset Set	Training	Test	Total
Camera Poses	1–40	41–50	50
Normal Images	29,085	7,300	36,385
Anomalous Images	0	15,024	15,024

V. Dataset Validation

To assess the realism of the synthetic images in the ALLO dataset, we compared them with real photographs of the International Space Station (ISS) captured during previous missions. For this comparison, a set of real ISS images was selected and corresponding camera poses were manually approximated in Blender. These poses were adjusted to align the synthetic camera viewpoint with the perspective of the real photographs, enabling a direct visual comparison of station geometry, lighting conditions, and surface appearance. Example real–synthetic image pairs are shown in Fig. 7. This qualitative comparison demonstrates that the rendered images reproduce key visual characteristics observed in real space imagery, including shadowing patterns, surface reflections, and large-scale illumination effects.

To further quantify the similarity between real and synthetic imagery, we evaluated two commonly used image similarity metrics: the Learned Perceptual Image Patch Similarity (LPIPS) metric [20] and the Fréchet Inception Distance (FID) [21]. Using ten manually aligned real–synthetic image pairs, we obtained an average LPIPS score of 0.58. This value is comparable to those reported for several modern image synthesis models, including GILL (∼0.70 on the VIST

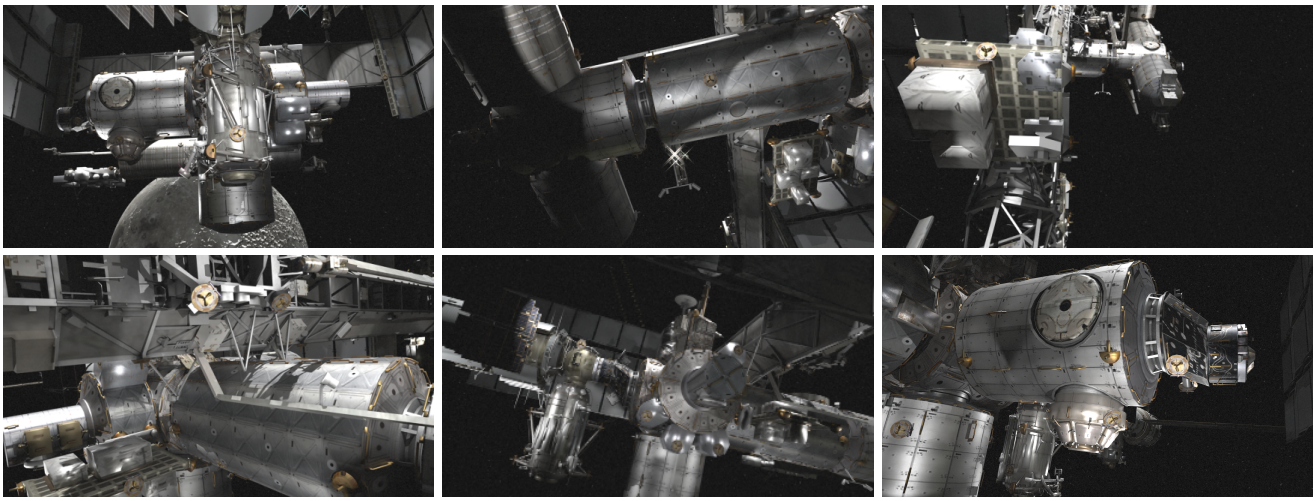


Fig. 5: Normal (anomaly-free) images from the ALLO dataset.

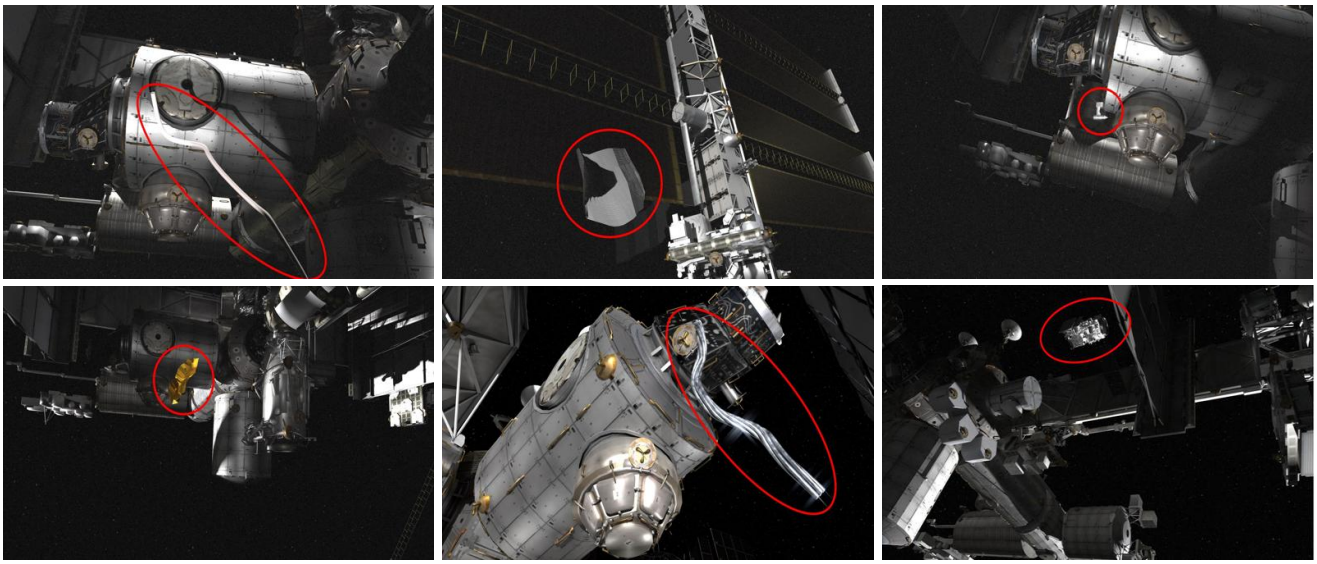


Fig. 6: Anomalous images from the ALLO dataset. Anomalies are circled in red.

dataset) [22] and Stable Diffusion XL (0.88 on COCO2017) [23]. The FID score was calculated between 319 ALLO training images and a set of historical ISS photographs, resulting in a score of 3.61. This value compares favourably with those reported for several recent generative models, including DiGIT [24], which reports an FID of 4.59 on ImageNet, and LC-GAN [25], which achieves an FID of 5.77 on AFHQ-v2.

The qualitative comparisons and quantitative metrics indicate that the ALLO synthetic images capture many of the visual characteristics of real space imagery. This suggests that the dataset provides a sufficiently realistic visual domain for evaluating anomaly detection methods intended for space-based robotic inspection tasks.

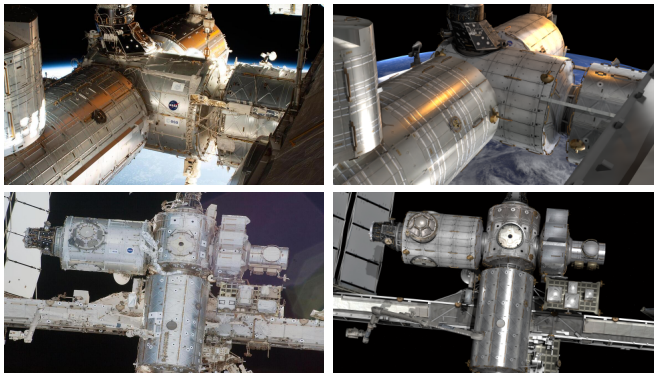


Fig. 7: Real images of the ISS (left) and their synthetic recreations (right).

VI. Conclusion

Space imagery presents challenges for autonomous navigation due to varying lighting conditions and complex scene geometry. These challenges must be addressed before a robust algorithm can be deployed on a space station. We introduce the ALLO dataset and its automated data generation pipeline, marking the first open-source anomaly detection dataset featuring images from lunar orbit. We demonstrate that the image in the ALLO dataset are realistic and reflective of the those that are expected to be captured by Canadarm3 on

lunar-orbiting vehicles. As autonomy becomes increasingly important in future missions, we aim for this dataset and data generation pipeline to be foundational in developing robust anomaly detection algorithms for space exploration.

REFERENCES

[1] J. C. Crusan, R. M. Smith, D. A. Craig, J. M. Caram, J. Guidi, M. Gates, J. M. Krezel, and N. B. Herrmann, "Deep space gateway concept: Extending human presence into cislunar space," in *2018 IEEE Aerospace Conference*, pp. 1–10. [Online]. Available: https://ieeexplore.ieee.org/document/8396541

[2] NASA's gateway program - NASA. Section: Humans in Space. [Online]. Available: https://www.nasa.gov/reference/nasas-gateway-program/

[3] (2020, June) About Canadarm3. [Online]. Available: https://www.asc-csa.gc.ca/eng/canadarm3/about.asp

[4] T. Ehret, A. Davy, J.-M. Morel, and M. Delbracio, "Image anomalies: a review and synthesis of detection methods," vol. 61, no. 5, pp. 710–743. [Online]. Available: http://arxiv.org/abs/1808.02564

[5] R. Brochard, J. Lebreton, C. Robin, K. Kanani, G. Jonniaux, A. Masson, N. Despré, and A. Berjaoui, "Scientific image rendering for space scenes with the SurRender software." [Online]. Available: http://arxiv.org/abs/1810.01423

[6] M. Bechini, M. Lavagna, and P. Lunghi, "Dataset generation and validation for spacecraft pose estimation via monocular images processing," *Acta Astronautica*, vol. 204, pp. 358–369, March 2023. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2023AcAau.204..358B

[7] M. Pajusalu, I. Iakubivskyi, G. J. Schwarzkopf, O. Knuuttila, T. Väisänen, M. Bührer, H. Teras, G. L. Bonhomme, M. F. Palos, J. Praks, and A. Slavinskis, "SISPO: Space imaging simulator for proximity operations," vol. 17, no. 3, p. e0263882. [Online]. Available: http://arxiv.org/abs/2105.06771

[8] B. Foundation. blender.org - home of the blender project - free and open 3d creation software. [Online]. Available: https://www.blender.org/

[9] Unreal engine 5. [Online]. Available: https://www.unrealengine.com/en-US/unreal-engine-5

[10] L. Bingham, J. Kincaid, B. Weno, N. Davis, E. Paddock, and C. Foreman, "Digital lunar exploration sites unreal simulation tool (DUST)," in *2023 IEEE Aerospace Conference*. IEEE, pp. 1–12. [Online]. Available: https://ieeexplore.ieee.org/document/10115607/

[11] S. Parkes, I. Martin, M. Dunstan, and D. Matthews, "Planet surface simulation with PANGU."

[12] R. T. Eapen, R. R. Bhaskara, and M. Majji, "NaRPA: Navigation and rendering pipeline for astronautics." [Online]. Available: http://arxiv.org/abs/2211.01566

[13] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD — a comprehensive real-world dataset for unsupervised anomaly detection," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9584–9592, ISSN: 2575-7075. [Online]. Available: https://ieeexplore.ieee.org/document/8954181

[14] Y. Zheng, X. Wang, Y. Qi, W. Li, and L. Wu, "Benchmarking unsupervised anomaly detection and localization," publication Title: arXiv e-prints ADS Bibcode: 2022arXiv220514852Z. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2022arXiv220514852Z

[15] MyNew, "Btech dataset," https://universe.roboflow.com/mynew/btech, feb 2024, visited on 2024-09-11. [Online]. Available: https://universe.roboflow.com/mynew/btech

[16] D. Tabernik, S. Šela, J. Skvarč, and D. Skočaj, "Segmentation-Based Deep-Learning Approach for Surface-Defect Detection," *Journal of Intelligent Manufacturing*, May 2019.

[17] International space station 3d model - NASA science. [Online]. Available: https://science.nasa.gov/resource/international-space-station-3d-model/

[18] D. E. Lee, "White paper: Gateway destination orbit model: A continuous 15 year NRHO reference trajectory," NTRS Author Affiliations: NASA Johnson Space Center NTRS Report/Patent Number: JSC-E-DAA-TN72594 NTRS Document ID: 20190030294 NTRS Research Center: Johnson Space Center (JSC). [Online]. Available: https://ntrs.nasa.gov/citations/20190030294

[19] ASCL.net - skyfield: High precision research-grade positions for planets and earth satellites generator. [Online]. Available: https://ascl.net/1907.024

[20] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 586–595. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_The_Unreasonable_Effectiveness_CVPR_2018_paper.html

[21] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.

[22] J. Y. Koh, D. Fried, and R. R. Salakhutdinov, "Generating images with multimodal language models," *Adv. Neural Inf. Process. Syst.*, vol. 36, pp. 21 487–21 506, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/hash/43a69d143273bd8215578bde887bb552-Abstract-Conference.html

[23] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Müller, J. Penna, and R. Rombach, "SDXL: Improving latent diffusion models for high-resolution image synthesis," in *The 12th Intl. Conf. Learn. Represent.*, 2023. [Online]. Available: https://openreview.net/forum?id=di52zR8xgf

[24] Y. Zhu, B. Li, H. Zhang, X. Li, L. Xu, and L. Bing, "Stabilize the latent space for image autoregressive modeling: A unified perspective," *Adv. Neural Inf. Process. Syst.*, vol. 37, pp. 28 636–28 661, 2024.

[25] S. Lee, M. Kim, Y. Chae, and B. Stenger, "Linearly controllable GAN: Unsupervised feature categorization and decomposition for image generation and manipulation," in *European Conf. Comput. Vis.* Springer, 2024, pp. 229–245.