SSC22-WKVI-07

Utilizing Deep Reinforcement Learning to Effect Autonomous Orbit Transfers and Intercepts on-orbit via Edge Compute

Gabriel Sutherland
Oregon State University
sutherlg@oregonstate.edu

Abstract

The goal of this project was to create a system capable of autonomous operation with minimal telemetry requirements, while operating within the limits of on-orbit compute and power reserves. Previous work has centered around the use of GPUs to train Deep Reinforcement Learning (DRL) agents for the purpose of autonomous space debris remediation. In the past, a DRL agent was fed orbital tracking data for both debris and active spacecraft to effect target intercepts. However, this approach proved problematic due to large network sizes and expensive computational and training costs. An updated approach utilized Nvidia GPUs to train a Double-Q network with a replay buffer, enabling autonomous orbit transfers and 1km intercepts in a simulated environment. Once within the intercept window, the spacecraft would switch control over to a Convolutional Neural Network (CNN), which relied on direct observational data to identify the target object. This data was supplied via simulated inputs for onboard lidar, infrared, and visible light sensors. Combined with the supplied ground tracking data for the target object, the spacecraft is able to identify the target before capture. While highly effective, the complete reliance on GPUs for inference precluded these solutions from being deployed to edge solutions in orbit due to the relatively high compute cost and cost of telemetry. To mitigate this issue, Deep-Q networks and CNNs were trained using traditional methods before being pruned to reduce both size and compute cost of the networks. To verify that the models had been successfully pruned while still maintaining performance, the models were uploaded to a cubesat model which was interfaced with the simulated environment. The physical cubesat model was configured with the intended operational limitations in mind: power generation and storage, compute power, telemetry capabilities, and sensor packages. The result was an autonomous spacecraft control system that can select the best candidate for a successful intercept, effect an orbit transfer, and capture the target with a relative velocity of less than 1m/s. After successful capture has been confirmed the spacecraft then deorbits the debris.

1 Introduction

The establishment of orbital infrastructure has undeniably been a great boon to humanity. Rapid and robust communications infrastructure, earth observation platforms, GPS navigation: each of these technologies has has positive impacts on society. Yet there remains the challenge of sustainably and responsibly utilizing the space in Earth-orbit. in 1 1978 paper Donald J Kessler first described a scenario in which a series of orbital collisions between spacecraft could lead to the disruption of safe navigation in orbit due to the sheer amount of debris.[1] While such a cascade has not yet happened, a solution to the rapid buildup of debris in orbit is needed. Early spaceflight programs had extremely polluting practices, such as explosive bolts for stage separation and leaving spent stages and derelict satellites in orbit after a mission completed its objectives and was decommissioned. Some programs and private spaceflight have put into place practices to mitigate debris creation, including hydraulic stage separation mechanisms and boosting decommissioned spacecraft to a higher "graveyard" orbit. Despite debris mitigating practices becoming more common, Low Earth Orbit (LEO) is becoming more and more crowded by the day, and several spacecraft collisions in the past few decades have contributed tens of thousands of trackable debris fragments. While the number of trackable objects which are greater than 10cm in diameter numbers in the tens of thousands,[2] it is estimated that there are over 100 million debris smaller than 1mm in diameter between LEO and Geosynchronous orbit. Debris mitigation technologies alone are unlikely to solve the issue of navigation hazards posed by orbital debris. Here we discuss a proposed solution utilizing ML and low-power inference hardware to autonomously capture and deorbit debris.

2 Challenges for ML aboard Spacecraft

Space is hard and manned space exploration is dangerous and unforgiving. Spacecraft are characteristically fragile, and debris strikes pose a serious operational hazard to manned and unmanned craft alike. There are several challenges to consider for any ML system deployed to a spacecraft:[3] a) limited live testing capabilities available, and testing in the form of payloads on missions are expensive; b) systems that are deployed need to be at a high technology readiness level (TRL);[4] c) environmental effects may influence deployed systems, such as ionizing radiation which can cause aberrant behavior in space capable hardware,[5] with behavioral inconsistencies that can present themselves in aberrant sensor behavior such as early or late sensor fusion, leading to potentially corrupted data to be processed; d) weight constraints make large orbital compute infrastructure infeasible; e) fully autonomous applications are almost exclusively deployed to controlled environments that provide information and environmental data; f) a lack of labeled data and elementary interpretability and explainability of ML systems makes ensuring reliability challenging. Unlike terrestrial applications, the orbital environment any space-faring vehicle will operate in is an extreme environment with conditions that can be difficult to predict, and are often challenging to scientifically quantify with current understanding. As a result of these myriad challenges, and ML systems deployed to orbital spacecraft must be robust and adaptable to weather the extreme environment and operate.

3 Advantages and Use-Case

Two of the largest bottlenecks in space operations are limited ground stations and high latency, making quick decision making on remote missions difficult. Conversely, while incorporating autonomous processes into a spacecraft is a complex task, there are several key advantages to increased autonomy of spacecraft; a) reduced latency in decision making, b) adaptive maintenance - modern spacecraft both manned and unmanned are extremely complicated in hardware and software, both of which are prone to faults. ML systems can perform maintenance functions that transcend anomaly detection and fault prediction, enacting automated maintenance and resolving faults to continue mission operations,[3] c) reduce latency for ship functions which do not require a human operator, such as automated checks for data or models, or course corrections, d) recent advances in ML like bit quantization, pruning, and hardware approximations enable inference on resource constrained edge devices, some of which have already been demonstrated in space.[6] In the case of space debris remediation, the ability for a spacecraft to not only autonomously identify strong intercept candidates, but to also effect autonomous orbit transfers and captures of debris would allow for a large fleet of debris remediation spacecraft to operate continuously. Such a fleet of craft would only require monitoring of operations, rather than ground crews to calculate every orbital manoeuvre and manually effect intercepts and captures of debris targets.

4 Infrastructure Requirements for ML Systems on Spacecraft

Developing novel ML methods to provide mission critical capabilities is a difficult task. Even more difficult is developing such ML tools that can operate within the hardware restrictions of spacecraft, which are extremely limited in storage, processing power, and power storage and delivery capabilities. While these considerations are also very much relevant for terrestrial applications, they gain increased importance in the extreme operating environment of space.

4.1 Data Collection

Data collection is of extreme importance, and one of the most challenging aspects of ML on edge hardware. Unlike industrial or lab environments where pre-labeled and processed data sets can be fed to the ML agent, in orbit the data is unlabeled and noisy, which necessitates that any ML model deployed to orbit has been trained with a dynamic and representative data set that reflects the orbital environment well. Domain shift robustness is an open topic of ML research that centres on adapting existing methods to increase resistance to distributional shifts from the underlying training data domain. In the context of space missions, the training data can be rendered fairly inaccurate compared to what the onboard spacecraft sensors detect once in orbit, as sensors decay and are exposed to various radiation sources.

4.2 Anomalous phenomena

The long-term exposure of electronics to radiation is disastrous, as the materials slowly break down and begin to malfunction. While some of the malfunctions that occur due to radiation exposure and cosmic rays can be predicted, and to an extent mitigated, there are also completely unpredictable incidents that can occur. As a result, ML systems cannot be expected to cover the full range of malfunction scenarios, and should be expected to fail when presented with an anomalous event not in its programming. In order to mitigate potential catastrophic failures of ML systems onboard a spacecraft, any ML system should be accompanied with an anomaly detection mechanism that flags anomalous data for review by a consensus checking mechanism. This mechanism, paired with a feedback pipeline such that ML systems can learn from anomalous events, should allow for the evolution of onboard ML to better cope with and respond to anomalous events.

4.3 Computational Limitations

High latency between ground stations and spacecraft, and increasingly complex onboard sensor suites, necessitates on-board computation that has a degree of autonomy. The computational requirements of a spacecraft must be balanced with the radiation tolerance required for the operational mission as well as weight and power consumption of the computational modules. While onboard computers are vulnerable to all kinds of radiation, cosmic rays have a tendency to produce hard to predict errors in modern computing units. With these limitations in mind, any ML system deployed to a spacecraft should be able to operate with severely limited computational power. Nvidia's Jetson Xavier Modules have been shown to be able to withstand a significant level of proton based radiation,[7] making them optimal candidates for on-board inference.

5 Experimental Setup

The experimental phase was conducted almost entirely in simulations, with a Deep Reinforcement Learning (DRL) agent operating on physical edge hardware to validate power consumption and program speed. The simulation environment was established with the following information: an orbital dynamics environment representative of the Earth-Moon-Sun system was created in python, gravitational anomalies were added, atmospheric densities and drag were modeled, and the earths magnetosphere was modeled using data from the World Magnetic Model. Two catalogs of orbital objects were imported, one representing all active spacecraft in orbit, and another from NORAD approximating all tracked space debris as spheres of various sizes. Finally, a spacecraft model was imported which represents ongoing work to develop a spacecraft capable of capturing smaller pieces of orbital debris, as shown in Figure 1. This spacecraft was equipped with an ion engine and large photovoltaic panels to provide plenty of power.

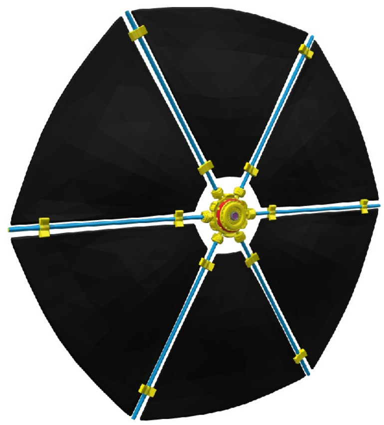

Figure 1: Debris Remediation Spacecraft Concept

Next, the ML agent architecture was designed. It was decided that for the purpose of training a ML agent to effect orbit transfers in a simulated environment, a Double Deep Q Network was the best architecture. Figure 2 shows a single Deep Q Network.

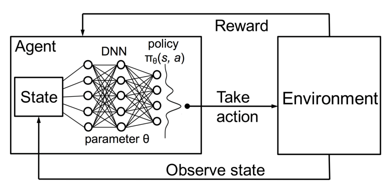

Figure 2: Deep Q Network

The ML agent was put in control of the spacecraft and given the task of intercepting and capturing a debris candidate with as little ΔV as possible. However, this was not the end. The agent attempted to find the most efficient orbital manoeuvre to successfully intercept the target, and several updates to the simulation needed to be implemented - a minimum altitude the spacecraft was not allowed to descend below, relative speed at intercept ¡1m/s, and refining the reward mechanism to be a function of intercept time and fuel consumption. Once the most blatant oversights were patched, the ML agent began to go through intercept after intercept, learning how to effect orbit transfers more efficiently with each epoch. The Double Deep-Q-Network was not given its actual orbital coordinates, only the velocity vector. Periodically, the ML agent was told its actual orbit vs its target position. As each epoch completed, the ML agent learned from it's errors and became better and better at maneuvering where it intended to.

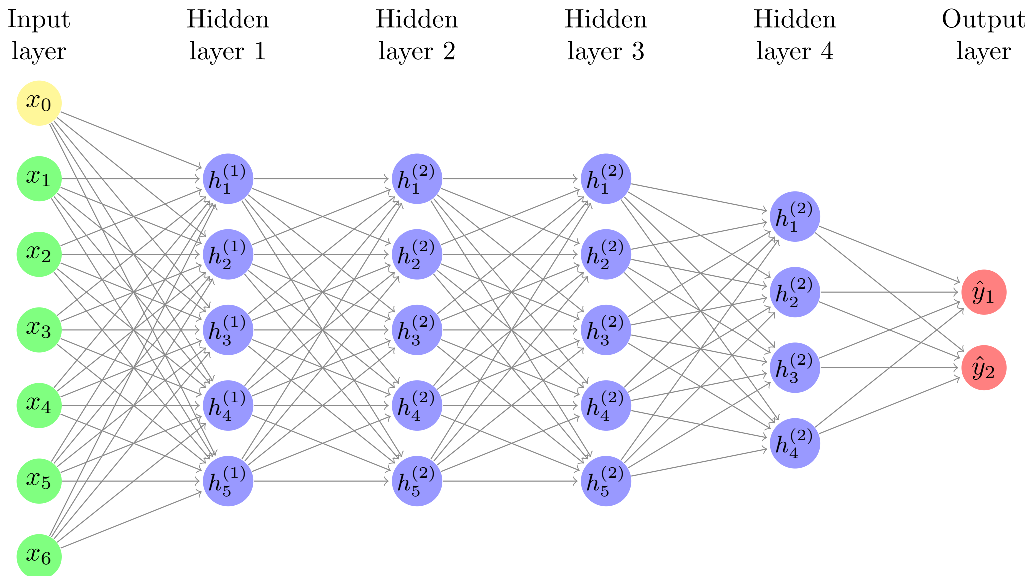

Figure 3: CNN for Debris Sighting during Intercept

Subsequently a Convolutional Neural Network, as shown in figure 3, would activate once the spacecraft was within 1km of its target, and provided estimates of target distance and orientation. At this stage in the intercept process, a simple PID controller was utilized to steer the spacecraft gently into the debris target.

5.1 Debris Capture Results

In simulations, the spacecraft controlled by the DDPG agent was able to effect orbit transfers and intercept debris targets while avoiding collisions with active spacecraft. While in the simulation there were several simplifications made, this is a promising step towards an autonomous debris remediation system.

6 Conclusion

Space debris are currently one of the largest threats to our orbital ecosystem. While steps to reduce the number of new debris created will help in minimizing the addition of new debris to orbit, it is likely that we will need to remove some of the objects in orbit to keep LEO safe for navigation. While this proposal is not a ready solution, we believe it is a step towards the development of autonomous debris remediation craft.

7 Acknowledgements

We would like to thank NASA, European Space Agency (ESA), NVIDIA, the Oregon Space Grant Consortium, Intel, the SETI Institute, and the Frontier Development Lab (FDL) for their continuous support.

References

[1] Kessler, D.J. and N.L. Johnson and J.-C. Liou and M. Matney, "The Kessler Syndrome: Implications to Future Space Operations," Preprint from n Astronautical Society Advances in the Astronautical Sciences Volume 137, 2010.

[2] Garcia, M., "Space Debris and Human Spacecraft," NASA.gov, May 26, 2021.

[3] Vlontzos, A. and G. Sutherland. and S. Canju and F. Soboczenski, "Next-Gen Machine Learning Supported Diagnostic Systems for Spacecraft," AI for Spacecraft Longevity Workshop at IJCAI, 2021.

[4] Lavin, A. and C.M. Gilligan-Lee and A. Visnjic and S. Ganju and D. Newman and S. Ganguly and D. Lange and A.G. Baydin and A. Sharma and A. Gibson and Y. Gal and E.P. Xing and C.

Mattmann and J. Parr, "Technology Readiness Levels for Machine Learning Systems," arXiv, 2021.

[5] Page, T.E. and J.M. Benedetto, "Extreme latchup sisceptibility in modern commercial-off-the-shelf (COTS) monolithic 1M and 4M CMOS static random-access memory (SRAM) devices," IEEE Radiation Data Workshop, 2005.

[6] "PhiSat-1 Nanosatellite Mission," Earth Observation Portal, 2020.

[7] Hiemstra, D.M. and C. Jin and Z. Li and R. Chen and S. Shi and L. Chen, "Single Event Effect Evaluation of the Jetson AGX Xavier Module Using Proton Irradiation," IEEE Radiation Effects Data Workshop in conjuction with 2020 NSREC, 2020.

[8] Wang, E. and J.J. Davis and R. Zhao and H. Ng and X. Niu and W. Luk and P.Y.K. Cheung and G.A. Constantinides, "Deep Neural Network Approximation for Custom Hardware," ACM Computing Surveys, May, 2019.

[9] Long. M. and H. Zhu and J. Wang and M.I. Jordan, "Unsupervised domain adaptation with residual transfer networks," arXiv preprint, 2016.

[10] Nilesh, G., "DNN Radiation hardened Coprocessor companion chip to NASA's upcoming High-Performance Spaceflight Computing processor," SBIR, 2020.

[11] Louizos, C. and U. Shalit and J.M. Mooij and D. Sontag and R. Zemel and M. Wellington, "Causal effect inference with deep latent-variable models,", Advances in Neural Information Processing Systems, 2017.

[12] Buesing. L. and T. Weber and Y. Swols and S. Ravaniere and A. Guez and J.B. Lespiau and N. Heess, "Woulda, coulda, shoulda: Counterfactually-guided policy search," arXiv preprint, 2018.

[13] Dietterich, T.G., "Hierarchical reinforcement learning with the MAXQ value function decomposition," Journal of artificial intelligence research, 2000.

[14] Vlontzos, A. and A. Alansary and K. Kamnitsas and D. Rueckert and B. Kainz, "Multiple landmark detection using multi-agent reinforcement learning," International Conference on Medical Computing and Computer-Assisted Intervention, 2019.

[15] Vlontzos, A. and H.B. Rocha and D. Rueckert and B. Kainz, "Causal Future Prediction in a Minkowski Space-Time," arXiv preprint, 2020.