SSC21-S1-46

Low-power boards enabling ML-based approaches to FDIR in space-based applications

James Murphy
Réaltra Space Systems Engineering, University College Dublin, School of Computer Science
Realtime Building, Clonshaugh Retail Park, Dublin, Ireland; +353 1 6875 723
jmurphy@realtra.space

John E Ward
Réaltra Space Systems Engineering
Realtime Building, Clonshaugh Retail Park, Dublin, Ireland; +353 1 6875 723
jward@realtra.space

Brian Mac Namee
University College Dublin, School of Computer Science
University College Dublin, School of Computer Science, Science Centre, Belfield, Dublin, Ireland; +353 1 716 2315
brian.macnamee@ucd.ie

ABSTRACT

Modern satellite complexity is increasing, thus requiring bespoke and expensive on-board solutions to provide a Failure Detection, Isolation and Recovery (FDIR) function. Although FDIR is vital in ensuring the safety, autonomy, and availability of satellite systems in flight, there is a clear need in the space industry for a more adaptable, scalable, and cost-effective solution. This paper explores the current state of the art for Machine Learning error detection and prognostic algorithms utilized by both the space sector and the commercial sector. Although work has previously been done in the commercial sector on error detection and prognostics, most commercial applications are not nearly as limited by the power, mass, and radiation tolerance constraints as for operation in a space environment. Therefore, this paper also discusses several Commercial Off-The-Shelf (COTS) multi-core micro-processors, small-footprint boards that will be explored as possible testbeds for future integration into a satellite in-orbit demonstrator.

INTRODUCTION

While traditional Failure Detection, Isolation and Recovery (FDIR) techniques are generally good at detecting single failures, they are limited in isolation capabilities, and struggle when multiple faults combine in unforeseen behaviors. Additionally, these systems offer limited capabilities for prognosis of future issues, reducing the opportunities to catch and correct potentially catastrophic problems. Most FDIR functions introduce automatic actions that are customized, bespoke, and complex. However, with the advance of space-based low-power, high-performance computing systems, more advanced FDIR functionality can be developed and deployed to greatly enhance the autonomous reaction of the spacecraft to immediate and foreseen failure modes. Specifically, the use of on-board Machine Learning algorithms that actively learn from in-flight data to diagnose and react rapidly to these current and future failures will minimize performance loss and thus provide an invaluable ability for the optimal performance of space-based assets.

One of the growing research topics in all major space agencies is the application of Machine Learning in both downstream (e.g., data analytics of Earth Observation data) and upstream (e.g., applying Machine Learning techniques in spacecraft on-board systems). There are new developments in many branches of space engineering including the emergence of expert systems. Due to the specific requirements for space hardware, the footprint of electronic devices carried must be as small as possible to reduce mass and volume for storage. Furthermore, due to the restricted power budgets of space missions, devices must also be low powered. This also aids the thermal properties of the spacecraft. The vacuum in space also presents a difficult problem to overcome as it raises thermal issues on circuit boards. Finally, radiation damages with electronic circuits and memory.

Detecting faults on circuit boards is difficult. Usually, the current and voltages are monitored using non-invasive (electromagnetic methods) and less accurate methods, or invasive (multimeter techniques) and more accurate methods. Additional boards could also be added to detect the system voltages and currents. These methods are normally considered the standard for FDIR and Prognostic error detection.

This paper reviews the state of the art in applying machine learning methods in space applications and describes and compares the leading currently available COTS boards for space-based machine learning.

STATE OF THE ART

The interest in these technologies is supported by many open-source tools which allow for rapid development of concepts. In addition, many low-cost cloud-based services supported by powerful computing hardware such as Google Colab[1] makes such services easily accessible. In the space domain, the use of these techniques is already being explored for Earth observation applications [5], sensor fusion for navigation [6] and satellite operations [3]. It is believed that these techniques also can benefit future space transportation systems, in applications such as avionics and system health monitoring [2]. This can also lead to the development of inexpensive electronic systems for space-based operations [2].

Current FDIR space systems are considered crude but effective. Prognosis is currently non-existent in space systems. However, commercial companies such as Deutsche Bahn for rail and Boeing for aircraft [11] are currently researching prognostics for future applications. Most FDIR systems have physical circuit monitors such as latch-up protection or voltage/current monitoring systems. These add heavy and expensive components to a board to give the ability to recover. The requirement not to fail in general is also on the individual components as current FDIR systems cannot account for component level failures, increasing the cost of the boards by factors of hundreds or even thousands. Adding a system that can compensate for unexpected inputs may reduce potential fail points, thereby reducing overall costs.

Research into anomaly detection has also been conducted around time-series data with regards to live data streaming. The scenario in space is even more challenging than in terrestrial applications due to the extremely harsh environment. The requirement on boards to survive the massive vibrations of a rocket launch to the extreme radiation and thermal environment of space, requires hardware to be robust and tested to survive in these environments. This is one of the largest factors contributing to the cost of these products. Creating a system that reduces the need for these intensive tests is the next step of space-rated computer systems. That is where an opportunity exists to utilize ML techniques to reduce the reliance on testing.

The options for a small footprint board are currently limited for terrestrial applications due to the required processing power to perform machine learning algorithms. The number of options for radiation-hardened, space grade boards are currently even lower as most space quality hardware are several years behind the terrestrial market. Though, these electronics and boards are currently under development by large international companies such as Texas Instruments, Irish branches of international companies such as the Movidus group at Intel and Xilinx Dublin and start-up companies such as Ubotica, also located in Dublin.

Phi-Sat-1 is the European Space Agency's (ESA) first attempt at putting an Edge AI board in space. It successfully launched on 3rd of September 2020 on board a European Vega rocket [1]. This is the first in-orbit demonstration of an Edge AI board. Phi-Sat-1 is a cube sat focused on Earth observation and on-board image analysis. Its primary payload was a hyperspectral imager and the Machine Learning board. It is operated by ESA's Phi Lab which focuses on machine learning applications in space. The revolutionary idea of Phi-Sat-1 was that if an Edge AI board could be put on-board a satellite and an image analysis algorithm deployed on it, the link budget could be reduced, saving precious bandwidth for the mission. To accomplish this, ESA chose the Intel Movidius Myriad 2 chip as their hardware accelerator due to the low mass and power requirements. Ubotica was contracted to develop the algorithm and to qualify the chipset for space-based operations. This led to an intensive qualification campaign as the Myriad 2 would be the first Machine Learning board qualified for in-orbit operations. At the time of writing this paper, initial results from the Phi-Sat-1 mission are promising. Conducting the image analysis on board has saved up to 90% of the bandwidth for a similar outcome when compared to a ground-based analysis [1].

Machine Learning applications for space can be broken down into two categories, Space Based and Ground Based. These have vastly different requirements when it comes to Size Weight and Power (SWaP) constraints. For example, a board in Space must deal with harsh environments with regards to temperature and radiation. This puts limitations on the board, which in turn, limits the capabilities of any deployed algorithm. A ground-based system may not be as useful as an in-orbit system

[1] https://colab.research.google.com

due to the substantially smaller amount of data the system will receive due to mission link budgets. This section explores two examples of each system.

There are a multitude of techniques utilized in Machine Learning. These include both Supervised and Unsupervised approaches. The methods used by each of the examples in this section are also explored.

Space Based Applications

Image Analysis - Earth Observation: Phi-Sat-1 is a CubeSat designed by ESA's Phi Lab for use on in-orbit Earth observation research. Phi-Sat-1's primary mission is to determine whether an in-orbit solution to image analysis could be deployed. To accomplish this, a powerful, but low powered board was required. It was also required to be qualified for long term space operations, meaning, survive the thermal environment in space i.e., large gradients, vacuum, the vibrations encountered during launch, and the radiation environment of low-Earth orbit. Phi-Sat-1 was designed to analyze the images taken by its hyperspectral imager and analyze cloud coverage, only transmitting the processed and cloudless images to ground, saving on downlink budget. Phi-Sat-1 was launched on August 17th, 2020, and its initial results have been promising. The processed images have cut down on downlink budgets as expected and the system thus far has survived the space environment [1].

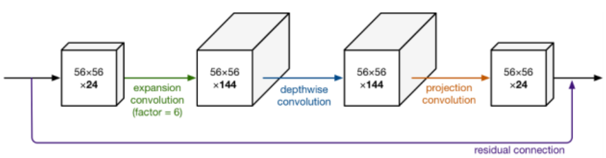

Figure 1: Convolutional Neural Network

The method used by Phi-Sat-1 to detect clouds was a convolutional neural network (CNN) [1]. A CNN is an artificial neural network designed to recognize patterns efficiently and accurately within structured arrays of data such as images and have become the standard approach for computer vision problems. This makes it ideal for use in Earth observation scenarios such as the one used by Phi-Sat-1. These models tend to be quite large due to the size of the images being analyzed, especially in Earth observation where there are TB of raw image data per orbit. The success of this method in Phi-Sat-1 has proven the usability of powerful Machine Learning Edge boards in an in-orbit environment.

Anomaly Detection - ESA's Future Launcher Preparatory Program FLPP: ESA's FLPP program is currently investigating Commercial Off the Shelf (COTS) avionics solutions for launchers employing Machine Learning techniques. The primary idea of this is to detect anomalies during flight and potentially rectify the issue. The study was to identify the most promising boards and algorithms to time-series datasets for a launcher environment. This also imposes certain limitations on a potential system due to the harsh environment of a launcher. Benefits vs risks were also explored in the study based on different Machine Learning method and board combinations. The goal was to develop a generalized building block to protect avionics from the environment experienced by a launcher. This resulted in the development & in the development and prototyping of several proofs of concepts [2]. The most promising result from this paper was found to be a Long Short-Term Memory (LSTM) based Autoencoder.

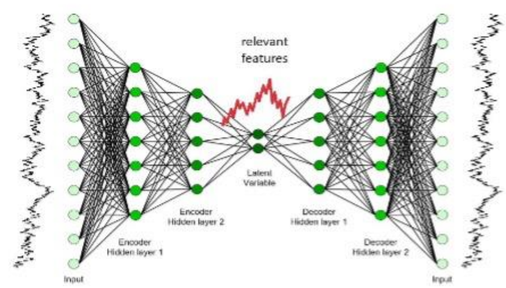

Figure 2: Conceptual example of an autoencoder

An autoencoder [12] is a type of artificial neural network used to learn efficient data encodings for unsupervised data. The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically dimensionality reduction, by training the network to learn signal noise. The key to autoencoders is not only that there is a reduction, but also a reconstructing side, where the autoencoder tries to re-generate the data from the reduced encoding as close as possible to its original input. Autoencoders are often trained with only a single layer encoder and a single layer decoder but using many or deep encoders and decoders offers many advantages. It is possible to build autoencoder based on feedforward neural networks. However, to consider the temporal data, the autoencoder can be based on long short-term memory (LSTM) layers. Unlike feedforward neural network, we put information into the LSTM sequentially, one number at a time. These are explained in more detail below.

Ground Based Applications

Anomaly Detection - Downstream Anomaly Detection: Hundman et. al. [3] researches the possibility of replacing the satellite operator with a Machine Learning replacement. However, the ability to reduce the workload on satellite operators is nevertheless sought after. This work defined an important and growing challenge within the satellite telemetry sector. Spacecraft operations stand to benefit greatly from Machine Learning anomaly detection approaches. LSTMs were found to be the most applicable method for predicting spacecraft telemetry anomalies while addressing key challenges around interpretability and complexity. This work has been deployed on the SMAP satellite ground segment where over 700 channels are monitored in real-time. There have been several correctly identified anomalies thus far. However, there have also been multiple false positives, showing the need for further refinements in the model [3].

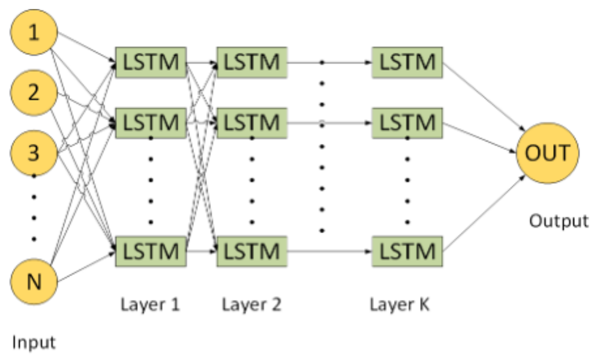

Figure 3: Neural network architecture based on LSTM layer

A Long Short-Term Memory (LSTM) network is a type of recurrent neural network (RNN). LSTMs have feedback connections unlike regular RNNs and preserves errors that can be backpropagated. This allows LSTMs to continue to learn for many steps. They can process single points and sequences of data, composed of a cell, an input, an output and a forget gate. LSTMs contain information outside the normal operations of a recurrent neural network in a gated cell. This allows cells to be treated like computer memory through reading, writing and storage. This makes LSTMs suited for working with time series data.

EDGE AI HARDWARE

This section reviews the Edge AI boards most applicable to space-based systems. The power draw is considered the most important factor due to limitations of power generation capabilities on board satellite subsystems. The boards investigated in this paper have a broad selection of power draws and Trillions of Operations Per Second (TOPS) rates allowing a wide range of potential results when used with deployed Machine Learning algorithms.

The list of boards investigated here are:

- Nvidia Jetson Xavier NX
- Huawei Atlas 200
- Google Coral
- Intel Movidius Myriad 2

Nvidia Jetson Xavier NX

The Nvidia Jetson Xavier NX is a high-power small footprint edge AI board using Nvidia 12nm architecture. It is capable of up to 32 TOPS of computing power and drawing a minimum of 10W of power. The Jetson Xavier also uses Nvidia's software development suite JetPack[2] allowing cross compatibility between the entire Jetson family of boards [7]. The Xavier NX model is used for more intensive operations than intended in this paper. However, this gives a good baseline for more powerful boards.

Figure 4: Nvidia Jetson Xavier NX

Huawei Atlas 200

The Huawei Atlas 200 is one of the closest competitors to the Nvidia Jetson Xavier in terms of Edge AI computing. The Ascend 310 chip is designed for image processing and other Machine Learning applications. This gives the Atlas 200 up to 22 TOPS of Machine Learning power at a maximum of 20W [8]. The Atlas is comparatively expensive and low powered, but it is nevertheless a good comparison to the Jetson Xavier.

[2] https://developer.nvidia.com/embedded/jetpack

Figure 5: Huawei Atlas 200

Google Coral

The Google Coral is powered by a quad Cortex-A53 and uses a Google Edge TPU as a coprocessor to provide 4 TOPS at only 2W. The Google Coral is tied as the most efficient board in this paper at 2 TOPS/Watt [9]. The Intel Myriad X also supplies this efficiency. The Google Coral also has a larger development board model and small USB style accelerator. The large development board assists software development and debugging before being deployed on the accelerator unit.

Figure 6: Google Coral

Intel Movidius Myriad 2

The Intel Neural Compute Stick is powered by an Intel Movidius Myriad 2 chipset. The Myriad 2 supplies the board with 1 TOPS at 1W [10]. Intel has already released the Myriad X powered Neural Compute Stick 2 which gives 2 TOPS at 1W, making it a much more powerful board [4]. However, the Myriad 2 chip is the only chip in this paper which also has space heritage and has been qualified for the space environment. The Myriad 2 VPU was integrated into the Phi-Sat-1 mission [1] as its primary inference device for image analysis. The Myriad 2 was also the first Edge AI board to fly on a space mission. For this reason, it is used in this paper.

Figure 7: Intel Movidius Neural Compute Stick

Board Comparison

This section compares the specifications of each board. The most important aspects of a board in this paper are Power and TOPS as space-based applications have a hard limit on power inputs. However, price in USD is also used in this analysis. Table 1 lists the specifications of all boards used in this paper.

Table 1: Board Comparison Table

Board	Power (W)	TOPS	Price USD	Space Heritage
Xavier	15	35	400	No
Atlas	20	22	950	No
Coral	2	4	100	No
Myriad X	2	4	80	No
Myriad 2	2	2	60	Yes

Figure 5 is a plot of TOPS/Watt and USD/TOPS. This gives an overview of the wide array of options available in the commercial market and to try and find which board offers the best value per Watt and USD. In terms of TOPS/Watt and USD/TOPS, the boards are quite similar. This means that the potential applications of the board will be the determining factor of which board could be used.

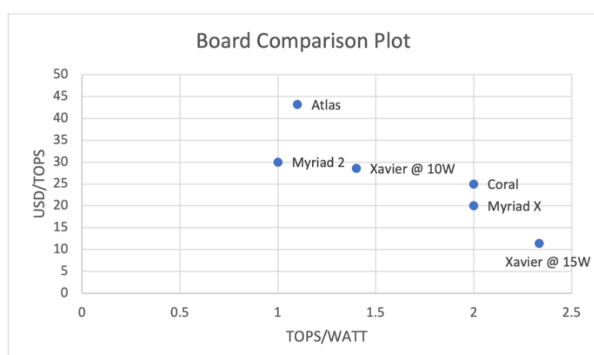

Figure 8: Board Comparison Plot

CONCLUSION

There are many opportunities for the space sector to take COTS modules from the commercial sector for use in space flight. Work has already been done on several systems to qualify them for either aeronautical or space environments. The variance in the computing performance and the power consumption between these boards also allows for a wide range of applications. Low power consumption boards are generally suited for missions with low power budgets, but still have enough computing performance to deploy most Machine Learning methods. Higher powered boards are less suited for small missions such as CubeSats due to their large power consumption. They are also more susceptible to radiation due to their generally higher density of components, which reduces their applicability to deep space missions. However, ground segment development is ideal for these boards as shown. Due to the multitude of applications of Machine Learning in the space sector, there are also many different Machine Learning methods that may be used. In summary, there are a wide range of platforms available to the space sector that can be either used directly or modified for use in-orbit or for ground segment missions. However, the mission requirements will decide which board and which Machine Learning method should be used.

Acknowledgments

This work was supported by Réaltra Space Systems Engineering. This work was also supported by the Irish Research Council under the Employment-based postgraduate Scheme (IRC-EBPPG/2020/11).

References

1. Furano, G., Meoni, G., Dunne, A., Moloney, D., Ferlet-Cavrois, V., Tavoularis, A., ... & Fanucci, L. (2020). Towards the Use of Artificial Intelligence on the Edge in Space Systems: Challenges and Opportunities. *IEEE Aerospace and Electronic Systems Magazine*, *35*(12), 44-56.
2. FLPP preparing for Europe's next-generation launcher, Publication available: https://www.esa.int/Enabling_Support/Space_Transportation/New_Technologies/FLPP_preparing_for_Europe_s_next-generation_launcher (Accessed 27-May-2020).
3. Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018, July). Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 387-395).
4. Barry, B., Brick, C., Connor, F., Donohoe, D., Moloney, D., Richmond, R., ... & Toma, V. (2015). Always-on vision processing unit for mobile applications. *IEEE Micro*, *35*(2), 56-66.
5. Vojtekova, A., Lieu, M., Valtchanov, I., Altieri, B., Old, L., Chen, Q., & Hroch, F. (2021). Learning to denoise astronomical images with U-nets. Monthly Notices of the Royal Astronomical Society, 503(3), 3204-3215.
6. Bagnell, J. A., Bradley, D., Silver, D., Sofman, B., & Stentz, A. (2010). Learning for autonomous navigation. *IEEE Robotics & Automation Magazine*, *17*(2), 74-84.
7. Nvidia Jetson Xavier NX Datasheet. Available at: https://developer.nvidia.com/embedded/jetson-xavier-nx (Accessed: 18-Feb-2021).
8. Huawei Atlas 200 Datasheet. Available at: https://e.huawei.com/en/products/cloud-computing-dc/atlas/atlas-200 (Accessed: 18-Feb-2021).
9. Google Coral DK Datasheet. Available at: https://coral.ai/docs/dev-board/datasheet/ (Accessed: 18-Feb-2021).
10. Intel Movidius Myriad 2 Datasheet. Available at: https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/myriad-x-product-brief.pdf (Accessed: 18-Feb-2021).
11. Deutsche Bahn Machine Learning: https://www.deutschebahn.com/en/Digitalization/technology/New-Technology/artificial_intelligence-3520346 (Accessed 27-May-2020).
12. Kelleher, J. D., Mac Namee, B., & D'arcy, A. (2020). Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. MIT press.