remote sensing

Review

Review on Hardware Devices and Software Techniques Enabling Neural Network Inference Onboard Satellites

Lorenzo Diana [1,*,†] and Pierpaolo Dini [2,†]

[1] Independent Researcher, 56100 Pisa, Italy
[2] Department of Information Engineering, University of Pisa, Via G. Caruso 16, 56100 Pisa, Italy; pierpaolo.dini@ing.unipi.it
* Correspondence: ldiana.res@libero.it
† These authors contributed equally to this work.

Abstract: Neural networks (NNs) have proven their ability to deal with many computer vision tasks, including image-based remote sensing such as the identification and segmentation of hyperspectral images captured by satellites. Often, NNs run on a ground system upon receiving the data from the satellite. On the one hand, this approach introduces a considerable latency due to the time needed to transmit the satellite-borne images to the ground station. On the other hand, it allows the employment of computationally intensive NNs to analyze the received data. Low-budget missions, e.g., CubeSat missions, have computation capability and power consumption requirements that may prevent the deployment of complex NNs onboard satellites. These factors represent a limitation for applications that may benefit from a low-latency response, e.g., wildfire detection, oil spill identification, etc. To address this problem, in the last few years, some missions have started adopting NN accelerators to reduce the power consumption and the inference time of NNs deployed onboard satellites. Additionally, the harsh space environment, including radiation, poses significant challenges to the reliability and longevity of onboard hardware. In this review, we will show which hardware accelerators, both from industry and academia, have been found suitable for onboard NN acceleration and the main software techniques aimed at reducing the computational requirements of NNs when addressing low-power scenarios.

Keywords: TinyML; low power; AI hardware accelerators; onboard satellite; artificial intelligence; machine learning; deep learning; computer vision; industrial internet of things; embedded systems

check for updates

Citation: Diana, L.; Dini, P. Review on Hardware Devices and Software Techniques Enabling Neural Network Inference Onboard Satellites. *Remote Sens.* **2024**, *16*, 3957. https://doi.org/10.3390/rs16213957

Academic Editor: Silvia Liberata Ullo

Received: 30 September 2024
Revised: 21 October 2024
Accepted: 22 October 2024
Published: 24 October 2024

1. Introduction

1.1. Motivations and Contributions

The satellite market is experiencing significant growth due to several interconnected factors. The growing demand for faster and more reliable connectivity is one of the main drivers of this expansion. With the proliferation of digital technologies and the internet, there is a growing need for satellite-based solutions, especially in remote and unserved areas. This need is further amplified by the rise of autonomous vehicles, drones, and the Internet of Things (IoT), which require robust satellite communication capabilities [1,2]. Technological innovations have made satellites more efficient and affordable. The development of smaller and lighter satellites with advanced capabilities, such as electric propulsion and miniaturization, has reduced launch costs and improved operational efficiency. This has led to an increase in the deployment of low Earth orbit (LEO) satellites, which offer lower latency and faster data transmission compared to traditional geostationary satellites [3]. Satellite applications are diversifying, covering areas such as communication, Earth observation, military intelligence, and scientific research. The commercial communications sector, especially satellite internet services, has seen substantial growth due to the growing interest in satellite television and imaging. In addition, the growing number of space exploration missions and private investment in satellite technology are improving the market outlook. Governments around the world are also investing heavily in satellite infrastructure for various applications, including homeland security and disaster management. Integrating artificial intelligence (AI) models onboard satellites offers significant advantages that improve satellite operations, data management, and enable autonomous decisions in space. These advantages include the following:

- Improved Autonomy and Responsiveness: Onboard AI enables satellites to make real-time decisions by processing data and responding autonomously without waiting for instructions from ground control. This responsiveness is crucial for tasks such as collision avoidance, given the increase in space debris. Furthermore, autonomous systems can adapt to changing conditions in space, improving the reliability and efficiency of missions [4–8].
- Optimized Data Management: AI plays a vital role in data management, enabling onboard processing. By analyzing and filtering data before transmission, AI significantly reduces the volume of information sent to Earth, minimising bandwidth use and storage costs [9–13].
- Advanced Applications: AI improves Earth observation capabilities by improving the quality of collected data. For example, AI can improve image resolution and detect environmental changes, which are essential for monitoring natural disasters or the impacts of climate change. Furthermore, for planetary exploration or lunar rover missions, onboard AI facilitates autonomous navigation and obstacle detection, making these missions more efficient and less dependent on human intervention [14–17].
- Support for Complex Operations: AI enables the coordination of multiple satellites working together as a swarm, enabling complex operations that are difficult to manage manually from Earth. Furthermore, AI systems can monitor satellite health and predict potential failures, enabling proactive maintenance actions that extend the operational life of satellites [18–20].

The adoption of onboard data processing (OBDP) technologies represents a paradigm shift, allowing satellites to process data in situ, minimizing the volume of data transmitted and enabling near-real-time analysis. This approach mitigates the inefficiencies related to the transmission of large volumes of raw data to Earth for back-end processing [21,22]. Key technologies enabling OBDP include digital signal processors (DSPs) and field-programmable gate arrays (FPGAs), which offer parallel processing capabilities and optimizations for the balance between performance and power consumption. These devices are particularly suited to applications such as synthetic aperture radar (SAR) image processing, where large volumes of data must be processed in real time [23]. Examples of onboard data processing systems include AIKO's data processing suite, Fraunhofer EMI's data processing unit, and KP Labs' Leopard technology, which integrate FPGAs and AI processors for real-time analytics of hyperspectral images on-board satellites [24].

1.2. Limitations for Adopting AI Onboard Satellites

The adoption of Artificial Intelligence (AI) onboard satellites faces several significant limitations. These challenges stem from technical, environmental, and operational constraints that impact the feasibility and effectiveness of AI implementations in space. We can group these limiting factors in three main categories:

- Hardware-related limitations.
  Ensuring the reliability of AI models in the face of errors caused by radiation is paramount. Satellites typically use radiation-hardened (rad-hard) processors to withstand the harsh space environment [25–27]. However, these processors are often significantly less powerful than contemporary commercial processors, limiting their ability to handle complex AI models effectively [28]. The performance gap makes it challenging to deploy state-of-the-art AI frameworks, which require substantial computational resources. Testing and validation processes must ensure that the AI models function correctly under extreme space conditions. This often involves cre-

ating simulations of these conditions, which can be both complex and costly [29]. Designing fault-tolerant systems capable of detecting and correcting errors in real time is a complex task and can introduce computational overhead. Balancing reliability with the available resources becomes necessary. Implementing radiation mitigation and fault tolerance solutions can increase costs and resource requirements, making it important to balance cost, performance, and resilience [30].
Satellites' onboard computational and hardware resources are limited. Onboard AI models may need considerable working memory to store model parameters and intermediate results during computations. Many satellite systems are not equipped with the necessary memory capacity, which restricts the complexity of the AI models that can be deployed, creating the need for optimal resource allocation techniques [31]. Moreover, the power available on satellites is limited, which restricts the use of high-performance chips that consume more energy. This results in a trade-off where lighter and smaller satellites may not be able to support the power demands of advanced AI processing units. Finally, the space environment presents unique challenges such as extreme temperatures and radiation exposure, which can affect the reliability and longevity of electronic components used for AI processing. Additionally, these conditions complicate the design of AI systems that must operate autonomously without human intervention.

- Model-related limitations.
  Lack of Large Datasets: Effective AI models, especially those based on deep learning, require large amounts of labeled training data to perform well. In many cases, particularly for novel instruments or missions to unexplored environments, such datasets are not available. This scarcity can hinder the model's ability to generalize and perform accurately in real-world conditions [32]. Model Drift and Validation: Continuous validation of AI models is essential, especially for mission-critical applications [33]. This involves downlinking raw data for performance assessment and potentially retraining models onboard, a process complicated by limited communication bandwidth and high latency in space. Moreover, post-launch updates to models need to be carefully managed. Limitations in communication [34] and the risks associated with system malfunctions make it challenging to perform updates during a mission. It is crucial to monitor models continuously and, if necessary, implement updates to ensure ongoing reliability.
- Other limitations.
  There are also limitations that are broad and that may affect not only AI applications, such as unauthorized access risks. The integration of AI into satellite systems increases vulnerability to hacking and unauthorized control. Ensuring cybersecurity [35–37] is critical but adds another layer of complexity to the deployment of AI systems in space. Cybersecurity measures such as encryption and authentication are vital to protect AI models and data from potential threats. However, implementing these mechanisms can be complex and resource-intensive. Care must be taken to ensure that these security measures do not compromise the system's overall performance. Continuous monitoring for potential threats is also necessary, but this requirement adds to the system's workload.

In this paper, we review the most interesting hardware and software solutions aimed to enable the use of AI, in particular neural networks (NNs), onboard satellites. The rest of this paper is organized as follows. In Section 2, we introduce the most relevant hardware solution presenting both commercially available hardware accelerators and the most recent Field-Programmable Gate Array (FPGA)-based designs from academia. In Section 3, we introduce the most interesting software methodologies and solutions that can help in porting Neural Network (NN)-based algorithms onboard satellites. In Section 4, we provide an overview of real case applications that can benefit from using onboard AI algorithms that have been addressed by the research community. Finally, in Section 5 we provide the conclusion.

2. Hardware Solutions

NNs usually require a considerable amount of hardware resources, i.e., memory and computation capability. Small satellites, e.g., CubeSat, cannot satisfy these needs due to cost and technical requirements. In particular, power consumption can be a stringent limiting factor when running inference of NNs onboard this kind of satellite [38,39]. A possible solution to reduce the power consumption when running NNs is to leverage a hardware accelerator. Thanks to the higher efficiency related to NN workload computation, hardware accelerators represent a valuable solution to enable the use of NNs onboard small satellites. On the one hand, there are numerous commercial off-the-shelf hardware accelerators for NNs available on the market [40]. Despite none of those accelerators having been specifically designed to meet space-grade requirements, some of them have already been tested for and used in short-term low Earth orbit satellite missions [41–43]. On the other hand, many NN accelerators have been developed to be deployed on FPGAs. These accelerators can be effectively used onboard satellites when deployed on space-grade FPGAs. Deploying AI models on dedicated hardware, such as Vision Processing Units (VPUs) or FPGAs, can deliver high-performance acceleration while reducing power consumption [44]. However, implementing and configuring AI models on these types of hardware requires specialized knowledge, and the process can be complex. Customizing the hardware and programming introduces additional challenges, particularly when it comes to updating models. For example, updating models on FPGAs often requires hardware re-programming, a process that must be managed carefully to avoid errors or malfunctions. To address these limitations, some hardware solutions are better suited for specific applications. For instance, Application-Specific Integrated Circuits (ASICs) like the Google Edge Tensor Processing Unit (TPU) are designed for low-power, high-efficiency inference and have been tested for radiation tolerance, making them suitable for space applications. Similarly, Nvidia Jetson boards, such as the Jetson Orin Nano and Jetson TX2i, offer robust performance and have been used in short-duration satellite missions, demonstrating their adaptability in harsh environments. These solutions provide a balance between performance, power consumption, and reliability, addressing the drawbacks of VPUs and FPGAs in space-grade applications.

2.1. ASIC

As already said, to the best of our knowledge, there are no ASIC AI accelerators specifically developed for space applications that are commercially available. Nonetheless, in the last few years, some AI accelerators have been tested for radiation tolerance capability and used in satellite missions. In this subsection, we will introduce these accelerators, explaining how they can enable onboard inference of NNs for satellite missions.

2.1.1. Google Edge TPU

The Coral Edge TPU, developed by Google, is an Application-Specific Integrated Circuit (ASIC) specifically engineered to accelerate TensorFlow Lite models while maintaining exceptionally low power consumption [45]. This hardware is optimized for the execution of quantized 8-bit neural network (NN) models, which are compiled for the Edge TPU, allowing for highly efficient inference processing. It is essential to acknowledge, however, that the Edge TPU does not support all operations available in TensorFlow [46]. In instances where unsupported operations are identified, the Edge TPU compiler automatically partitions the neural network into two distinct sections: the initial segment is executed on the Edge TPU, while the remaining operations are transferred to the central processing unit (CPU). To further streamline the process of neural network inference on the Edge TPU, Google offers the PyCoral API, which enables users to perform sophisticated tasks with minimal Python code [47]. In addition, Google provides a suite of pre-trained neural network models specifically tailored for the Edge TPU. These models cover a broad range of applications, including audio classification, object detection, semantic segmentation, and pose estimation [48]. The Google Coral Edge TPU is available in multiple form factors, suitable for both development and production environments, and has garnered significant attention for its potential in low Earth orbit (LEO) missions [43]. Moreover, it has undergone performance and radiation testing to evaluate its suitability for onboard satellite applications [49,50]. Given the increasing demand for low-power AI accelerators in the aerospace industry, the Edge TPU stands out as a highly promising and efficient solution [45]. Figure 1 shows the main products featuring the Google Coral TPU.

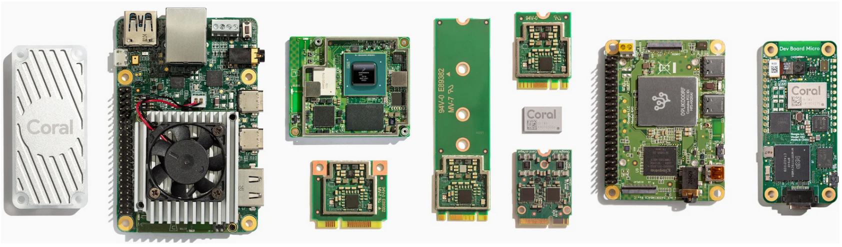

Figure 1. The main products featuring the Google Coral TPU. Image taken from [45].

In [51], the authors present the design and capabilities of a CubeSat-sized co-processor card, known as the SpaceCube Low-power Edge Artificial Intelligence Resilient Node (SC-LEARN). This work aims to facilitate the use of advanced AI frameworks in space missions, addressing challenges associated with traditional spacecraft systems that limit onboard AI capabilities due to their reliance on radiation-hardened components. The motivation behind this work resides in the fact that in the last few years, there has been a growing importance of AI in different domains, including autonomous systems and remote sensing. This highlights the need for specialized, low-power AI chips that can handle complex tasks in space, particularly for Earth science applications such as vegetation classification and disaster monitoring. The SC-LEARN card integrates the Google Coral Edge TPU to provide high-performance, power-efficient AI applications tailored for space environments. It complies with NASA's CubeSat Card Specification (CS2), allowing for seamless integration into existing SmallSat systems. The SC-LEARN card operates in three distinct modes:

- High-performance parallel processing mode for demanding computational tasks;
- Fault-tolerant mode designed for resilience against operational failures;
- Power-saving mode to conserve energy during less intensive tasks.

Moreover, the authors discuss the training and quantization of TensorFlow models specifically for the SC-LEARN, utilizing representative open-source datasets to enhance on-board data analysis capabilities. Finally, some future research plans are outlined, including the following:

- Radiation beam testing to assess the performance of the SC-LEARN in space conditions;
- Flight demonstrations to validate the effectiveness of the SC-LEARN in real mission scenarios.

Overall, this work represents a significant step toward integrating advanced AI technologies into space missions, enabling more autonomous operations and efficient data analysis directly onboard spacecraft.

2.1.2. Nvidia Jetson Orin Nano

The Nvidia Jetson Orin Nano [52], a significant advancement over the original Jetson Nano board, integrates a powerful combination of a multi-core CPU and an Nvidia Ampere-based GPU. This advanced architecture offers flexible power consumption settings, adjustable between 7 and 15 watts, making it suitable for a wide range of applications. The Orin Nano is available in a convenient development kit, which facilitates rapid prototyping and accelerates the design and testing of machine learning models and AI applications [53]. Running on a Linux-based operating system, the Jetson Orin Nano capitalizes on the robust computational capabilities of its GPU to efficiently execute diverse neural network (NN) models. This platform supports a broad array of high-level machine learning frameworks, enabling seamless neural network inference across multiple use cases. A key component of its performance is Nvidia's TensorRT, a deep learning inference optimizer that enhances the speed and efficiency of NN models by optimizing them for the underlying hardware architecture [54]. Of particular interest, certain devices in the Nvidia Jetson family, including the Orin Nano, have proven to be well suited for short-duration satellite missions [55,56]. These applications highlight the Jetson Orin Nano's durability and adaptability in harsh and resource-constrained environments, demonstrating its potential for use in aerospace and other challenging fields.

The SpIRIT satellite aims to utilize onboard AI to enhance data processing capabilities while addressing constraints such as limited computational resources, cosmic radiation resilience, extreme temperature variations, and low transmission bandwidths. In [57], the authors introduce Loris, an imaging payload employed onboard the SpIRIT mission, consisting of the following:

- Six visible light cameras (Sony IMX219);
- Three infrared cameras (FLIR Lepton 3.5);
- A camera control board;
- An NVIDIA Jetson Nano for processing.

Loris allows for advanced onboard computer vision experiments and supports innovative image compression techniques, including progressive coding. Such capabilities are crucial for optimizing data handling in space. The authors adopted several design considerations to enhance the robustness of the system:

- Multiplexing Strategy: this approach mitigates the risk of individual sensor failures by integrating multiple sensors into the system;
- Thermal Management: the payload is designed to operate effectively within a sinusoidal thermal profile experienced in orbit, ensuring that components like the Jetson Nano remain operational despite temperature fluctuations;
- Cost-Effectiveness: the use of commercial off-the-shelf (COTS) components aligns with budget constraints typical of nanosatellite missions, although it introduces higher risks regarding component reliability.

Loris enables on-orbit fine-tuning of AI models and enhances remote sensing capabilities. The imaging system is expected to facilitate a wide range of applications, including Earth observation and environmental monitoring, thereby broadening the potential uses of nanosatellites that generate vast amounts of data in scientific research and industry. Figure 2 shows the Loris architecture and electronic sub-module.

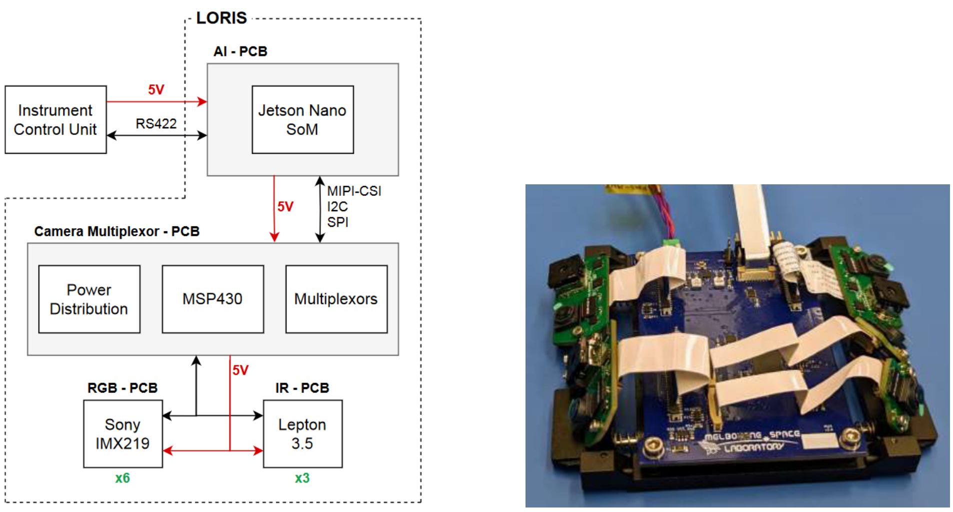

Figure 2. On the left, Loris architecture. On the right, the camera and multiplexing electronic sub-module. Images taken from [57].

2.1.3. Other NVIDIA Jetson Boards

- NVIDIA Jetson TX2i. The Jetson TX2i is another notable component used in space applications:
  - Performance: it provides up to 1.3 TFLOPS of AI performance, making it suitable for demanding imaging and data processing tasks.
  - Applications: Aitech's S-A1760 Venus system incorporates the TX2i, specifically designed for small satellite constellations operating in low Earth orbit (LEO) and near-Earth orbit (NEO). This system is characterized by its compact size and rugged design, making it ideal for harsh space environments.
- NVIDIA Jetson AGX Xavier Industrial. The Jetson AGX Xavier Industrial module offers advanced capabilities for more complex satellite missions:
  - High Performance: it delivers server-class performance with enhanced AI capabilities, making it suitable for sophisticated tasks like sensor fusion and real-time image processing.
  - Radiation Resistance: studies have indicated that the AGX Xavier can withstand radiation effects when properly enclosed, making it a viable option for satellites operating in challenging environments.
- Planet Labs' Pelican-2 Satellite. The upcoming Pelican-2 satellite, developed by Planet Labs, will utilize the NVIDIA Jetson edge AI platform:
  - Intelligent Imaging: this integration aims to enhance imaging capabilities and provide faster insights through real-time data processing onboard.
  - AI Applications: the collaboration with NVIDIA will enable the satellite to leverage AI for improved data analytics, supporting rapid decision-making processes in various applications.

2.1.4. Intel Movidius Myriad X VPU

The Intel Movidius Myriad X VPU is a specialized hardware accelerator known for its advanced capabilities in neural network (NN) inference, powered by a cluster of processors referred to as SHAVE cores. One of its key features is the flexibility it offers developers, allowing the number of SHAVE cores utilized for NN inference to be customized. With a total of 16 SHAVE processors, the Myriad X provides users with the ability to fine-tune the performance-to-power-consumption ratio, enabling them to optimize the device's operation based on their specific use cases and energy requirements. The Myriad X particularly excels in accelerating neural networks that incorporate convolutional layers, such as Fully Convolutional Networks (FCNs) and Convolutional Neural Networks (CNNs). Its ability to efficiently handle these complex architectures makes it a preferred solution for applications that demand high-speed, low-power processing of computationally intensive models. In addition to its use in terrestrial applications, the Myriad VPU family, including both the Myriad X and its predecessor, the Myriad 2, has established a foothold in the aerospace sector. Like the Google Edge TPU and Nvidia Jetson platforms, the Myriad VPU has been recognized as a dependable accelerator for NN inference in space. These devices have played critical roles in various space missions, having been deployed on satellites and aboard the International Space Station (ISS) [42,58,59]. Their combination of adaptability, resilience, and performance under challenging conditions makes them particularly well suited for use in low Earth orbit missions [41,43,60]. These characteristics underscore Myriad X's versatility as a robust solution for AI-driven tasks in both space exploration and other resource-constrained environments.

2.2. FPGA-Based Designs

FPGAs are the standard devices where hardware designs are prototyped. In some cases, they can also be used as deployment devices; satellites are one of these cases. FPGAs play a crucial role in satellite missions due to their versatility, re-programmability, and ability to handle complex data processing tasks. Space environments expose electronics to radiation, which can cause Single-Event Upsets (SEUs) that disrupt normal operation. Certain FPGAs are designed with radiation tolerance or hardness, minimizing the risks associated with space radiation. This makes them suitable for critical applications where reliability is paramount. Examples of radiation-hardened FPGAs are as follows:

- Xilinx Virtex-5QV [61]: it is designed specifically for space applications and is known for its high performance and re-programmability. It features enhanced radiation tolerance, making it suitable for use in satellites and other space systems where reliability is critical. This FPGA has been used in various missions, including those conducted by NASA and the European Space Agency (ESA).
- Actel/Microsemi ProASIC3 [62]: these FPGAs are anti-fuse-based devices that provide a high degree of immunity against radiation-induced faults [63]. These FPGAs are one-time programmable and are often used in applications where re-programmability is less critical but where robustness against radiation is essential.
- Microsemi RTG4 [64]: it is engineered to withstand radiation-induced Single-Event Upsets (SEUs) and Total Ionizing Dose (TID) effects [65,66], making it suitable for use in high-radiation environments such as space missions.

In this subsection, we will see the most recent and relevant designs aimed to accelerate inference of NNs onboard satellites.

In [67], the authors address the problem of accelerating Deep reinforcement learning (DRL) models onboard satellites. The application addressed concerns about the onboard real-time routing for dynamic low Earth orbit (LEO) satellite networks. The authors propose to reduce the inference time of DRL models by parallelizing and accelerating part of the convolutional layer's operation using an onboard FPGA. They propose a co-design method to improve the onboard inference time of the Dueling-DQN-based routing algorithm and they tested the proposed solution using an Onboard Computer (OBC) featuring both a CPU and an FPGA. The parallelization focuses on the ReLU activation function and on the 2D convolutional layer of the NN exploiting the FPGA, while all the other algorithm's operation is executed on the CPU. In particular, they modify the way the sum-pooling operation is performed inside the convolutional layer to enable parallelization. The implementation is carried out using Vivado HLS. They tested the proposed solution on a PYNQ-Z2 board, achieving a 3.1× inference time speedup compared to the CPU-only deployment.

In [68], the authors propose an FPGA-based hardware accelerator to perform inference of NNs onboard satellites. The authors also provide a comparison between their hardware accelerator and the Intel Myriad-2 VPU. To perform this comparison, they took into consideration the CloudScout case study showing that the proposed solution can achieve a lower inference time and better customization capability at the expense of a higher time to market and power consumption. The authors applied a custom quantization method to reduce the Convolutional Neural Network (CNN)'s size while maintaining a comparable accuracy. One of the main trade-offs highlighted by the authors is the one between performance and FPGA resource consumption. The proposed solution was deployed and tested on a Zynq Ultrascale+ ZCU106 Development Board. The authors noted that despite the CloudScout's CNN having been quantized, the CNN's size exceeded the on-chip memory capability of most commercially available FPGAs, requiring the integration of the off-chip DDR memory featured by the board into the accelerator's design. The proposed solution achieved a 2.4× lower inference time and 1.8× higher power consumption compared to the Intel Myriad 2 VPU. At the same time, the energy per inference was reduced by 24%. Finally, the proposed solution was deployed on the rad-hard Xilinx Kintex Ultrascale XQRKU060 to prove the proposed design could fit space-grade devices to enable longer duration and higher-orbit missions compared to the Intel Myriad 2 VPU.

In [69], the authors present an approach for onboard cloud coverage classification using a quantized CNN implemented on an FPGA. The study focuses on optimizing the performance of cloud detection in satellite imagery, addressing the challenges posed by quantization and resource utilization on FPGA platforms. They provided a specific CNN design, CloudSatNet-1, to maintain a high accuracy despite the quantization process.

To achieve this, they tried different bit widths for the CNN compression and evaluated accuracy, False Positive Rate (FPR), and other parameters. The Zynq-7020 board was used to test the proposed solution. The results demonstrate that the proposed CloudSatNet-1 achieves high accuracy in cloud coverage classification while being efficient enough for real-time processing on satellites. In particular, the proposed solution achieved a higher accuracy compared to the CloudScout CNN at the expense of a higher FPR. It represents a promising solution for on-board cloud detection in satellite systems, balancing accuracy with computational efficiency. The FPGA implementation shows significant advantages in terms of speed and resource efficiency compared to traditional CPU-based approaches. Figure 3 shows the architecture of CloudSaNet-1.

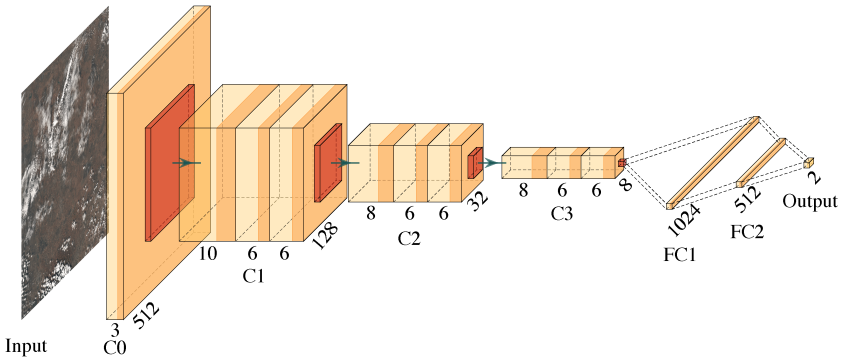

Figure 3. The architecture of CloudSaNet-1. Image taken from [69].

In [70], the authors present the ICU4SAT, a versatile instrument control unit. This design aims to enhance the capabilities of satellite instruments by providing a re-configurable control unit that can adapt to various mission requirements. ICU4SAT is designed to improve the autonomous operation of satellites, in particular facilitating advanced imaging and data processing. The ICU4SAT key features are as follows:

- Re-configurability: it can be adapted for different instruments and tasks, allowing it to support a wide range of satellite missions;
- Open Source: utilizing open-source components promotes collaboration and innovation, enabling users to modify and improve the system as needed;
- Integration with AI: the unit is designed to incorporate AI, enhancing its ability to process data and make decisions autonomously.

The system is built using open-source components, which facilitates accessibility and customization for different applications.

In [71], the authors explore the integration of FPGAs in enhancing the onboard processing capabilities for spacecraft pose estimation using CNNs. The authors present three approaches to landmark localization:

- Direct Regression on Full Image: a straightforward method but less accurate;
- Detection-Based Methods: utilizing heatmaps to improve accuracy;
- Combination of Detection and Cropping: involves detecting the spacecraft first and then applying landmark localization algorithms, which significantly enhances accuracy.

The paper details the implementation of CNN models on FPGAs, specifically using the Xilinx DPU IP Core with Xilinx UltraScale+ MPSoC hardware. The study highlights the network quantization techniques to optimize model performance on FPGAs. The authors show that the inference on the FPGA using 8-bit quantization has a negligible RMS drop (around 0.55) when compared to the 32-bit float inference on a PC. Finally, the authors implemented both the YOLOv3 and ResNet34-U-Net CNNs, finding that encoder–decoder models achieve better performance in landmark localization, with an inference time in the order of tens of milliseconds. Figure 4 shows the HW and SW implementation flow on the MPSoC (Xilinx, San Jose, CA, USA).

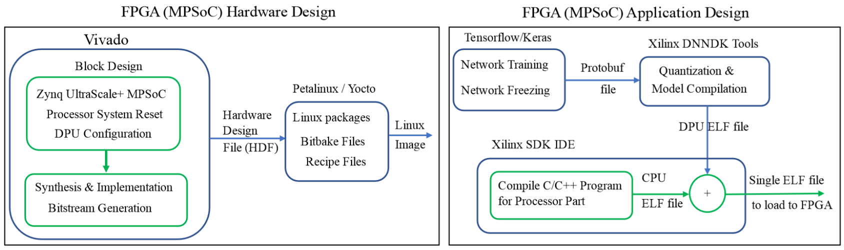

Figure 4. HW and SW inference flow on the MPSoC. Image taken from [71].

2.3. Neuromorphic Hardware and Memristors

Neuromorphic hardware represents a significant class of accelerators that offer a balance between high computational efficiency and low energy consumption. These systems are designed to mimic the neural architecture of the human brain, enabling efficient processing of NN models, particularly Spiking Neural Networks (SNNs).

Neuromorphic hardware, such as Intel's Loihi 2 [72,73], is specifically designed for high-efficiency, event-driven computations. Loihi 2 supports the development of SNNs, which can be more energy-efficient than traditional Artificial Neural Networks (ANNs). SNNs process information using discrete spikes, similar to biological neurons, which allows for lower power consumption and faster processing speeds in certain applications.

Memristors are another type of analog-signal-based hardware that demonstrate high efficiency and low power consumption. Memristors can store and process information simultaneously, making them highly suitable for implementing NNs. They offer non-volatile memory, which means they retain information without power, further reducing energy consumption.

The combination of neuromorphic hardware and memristors presents a promising approach for space applications, where power efficiency and computational capability are critical. These technologies can be used to develop advanced AI systems that operate efficiently in the harsh conditions of space.

Table 1 provides a comparison of neuromorphic hardware and memristors with traditional hardware accelerators.

Table 1. Comparison of neuromorphic hardware and memristors with traditional hardware accelerators.

Hardware	Computational Efficiency (TOPS/W)	Power Consumption (W)	Suitability for Space
Intel Loihi 2	10	<1	High
Memristors	Varies	<1	High
Google Edge TPU	4	2	Tested
Nvidia Jetson Orin Nano	6	7–15	Tested
Intel Movidius Myriad X	1.5	1–2	Tested

As shown in Table 1, neuromorphic hardware like Intel's Loihi 2 and memristors offer superior computational efficiency and lower power consumption compared to traditional hardware accelerators. Their suitability for space applications is high due to their energy efficiency and robustness in harsh environments.

2.4. Commercial Solutions

Nowadays, inside the product portfolio of some companies, we can find solutions specifically developed for AI acceleration onboard satellites. These solutions provide radiation-hardened devices to deploy spaceborne applications. In this subsection, we introduce some of the most interesting solutions provided by companies.

Ingeniars [74] provides the GPU@SAT [75,76] system, which is designed to enable AI and machine learning (ML) applications directly onboard satellites. This solution features a general-purpose GPU-like IP core integrated into a radiation-hardened FPGA to ensure the system can operate reliably in the harsh environment of space. By integrating GPU-like capabilities into satellite systems, GPU@SAT significantly enhances the computational efficiency and performance of AI applications, making them more feasible for space missions. This solution features an IP core that can be configured using OpenCL properties, allowing for the execution of kernels (small programs) on the hardware. This includes setting up necessary parameters, loading executable binary code, and allocating memory buffers. The system can schedule and execute multiple kernels, managing their sequence and dependencies efficiently. GPU@SAT is tailored for AI and machine learning (ML) applications, including computer vision tasks such as image and video processing. It leverages the parallel processing capabilities of Graphics Processing Units (GPUs) to accelerate computational tasks, which is crucial for real-time data analysis onboard satellites. Moreover, this system supports the development of edge computing applications in the context of Space-IoT, enabling smart processing directly onboard satellites. This enhances the capability for real-time or near-real-time data processing without constant communication with ground-based systems.

Mitsubishi Heavy Industries (MHI), Ltd. has developed an onboard AI-based object detector called Artificial Intelligence Retraining In Space (AIRIS) [77]. AIRIS aims to perform object detection from satellite-acquired images and its operations will be controlled by the space-grade Microprocessor Unit (MPU) SOISOC4, a product of development by the Japan Aerospace Exploration Agency (JAXA) and MHI. This MPU provides a high radiation resistance, allowing it to operate in the harsh radiation environment of deep space. AIRIS consists of an AI-equipped data processor and an Earth observation camera developed by the Tokyo University of Science. It will take images using the camera and it will use its AI to transmit to the ground only the subsection of images containing the target objects. Moreover, it will allow its onboard AI model to update by receiving a new version of it from the ground station. AIRIS will execute a demonstration as part of the "Innovative Satellite Technology Demonstration-4" mission aboard the small demonstration satellite "RAISE-4", which is scheduled for launch in 2025. In particular, AIRIS will perform vessel detection leveraging onboard AI inference and it will transmit to the ground only the portion of images containing the identified objects.

Blue Marble Communications (BMC) [78] provides a radiation-hardened Space Edge Processor (SEP) that integrates CPU, GPU, and FPGA capability into a high-performance, secure edge processor for spaceborne applications. This device features the following [79]:

- Industry-leading radiation performance and power efficiency;
- The SEP features an integral Ethernet/MPLS switch;
- AMD Ryzen V2748 CPU+GPU;
- AMD Versal coprocessor FPGA.

Both BruhnBruhn Innovation AB (BBI) and AIKO S.r.l. provide software to run AI-based algorithms on the SEP device [80,81].

Intel has developed the Loihi 2 neuromorphic chip, which is designed for high-efficiency, event-driven computations. Loihi 2 supports the development of SNNs, which can be more energy-efficient than traditional ANNs. This chip is particularly suitable for space applications due to its low power consumption and high computational efficiency.

MemComputing, Inc. offers memristor-based processing units that provide high efficiency and low power consumption for NN implementations. These units can store and process information simultaneously, making them highly suitable for space applications where power efficiency is critical. Memristor technology offers non-volatile memory, which means it retains information without power, further reducing energy consumption.

Both neuromorphic hardware and memristor-based solutions present promising approaches for enhancing the computational capabilities of satellites while maintaining low power consumption and high efficiency. These technologies are particularly valuable for developing advanced AI systems that can operate efficiently in the harsh conditions of space.

3. Software Tools, Methods, and Solutions for AI Integration

3.1. Main Challenges

The integration of AI models on embedded platforms aboard satellites presents a series of complex challenges, both technological and methodological, as it does for low-power embedded devices [82]. These challenges extend beyond traditional hardware and resource limitations [28,83], encompassing issues related to reliability, security, data management, and operational resilience in extreme environments. In this subsection, we provide an examination of some of the main challenges and further issues that may arise.

Integrating AI models into satellite systems introduces significant energy constraints, particularly in low Earth orbit (LEO) satellite missions, which are reliant on limited energy resources, primarily solar panels. Power must be distributed among various systems, including communication, thermal control, instrumentation, and computation. AI models that demand intensive processing can quickly increase energy consumption, necessitating dynamic and adaptive power management.

The harsh conditions of space pose another significant challenge. Extreme temperature variations and high-energy radiation can cause hardware malfunctions, impacting the reliability of embedded components. Radiation can induce soft errors, leading to temporary data corruption, or hard errors, resulting in permanent hardware failure. AI models and hardware platforms must be resilient to such failures, incorporating fault-tolerant computing techniques, such as hardware redundancy, fault recovery through checkpoints, and algorithms capable of recognizing and avoiding radiation-induced errors. Error detection and correction mechanisms, such as error-correcting codes (ECCs), should be employed to prevent performance degradation and data corruption.

In addition, embedded platforms on satellites are constrained in terms of computational capacity and memory compared to terrestrial systems. Memory resources are often scarce and must be carefully managed to execute complex AI models without overloading the system. Techniques like pruning, sparsity, and model distillation are essential for reducing the size and complexity of AI models, making them executable in low-memory environments.

Another critical issue is the need for software upgradability and maintenance. Once launched, satellites have limited capacity for physical upgrades, making remote software updates crucial. Updating AI models and algorithms onboard is complicated by communication constraints, latency, and the high costs of data transmissions from space. A reliable and secure remote update system is necessary to ensure the robustness of updates against potential interruptions or data corruption during transmission. Additionally, techniques for edge learning, which enable satellites to adapt AI models to environmental or data changes without requiring full updates from the ground, should be considered.

Data management and communication bandwidth also present major challenges. Satellites may generate vast amounts of data, but the available bandwidth for transmitting data to Earth is usually limited. AI models can reduce the transmission load by processing data locally, but this introduces challenges in selecting and managing relevant information for transmission. Intelligent onboard data processing and filtering must be implemented to transmit only significant information or critical alerts, such as anomalies or key data points in Earth observation missions. Moreover, advanced data compression algorithms should be utilized to preserve the quality of critical information without compromising subsequent analyses.

Cybersecurity is another growing concern in space. Vulnerable satellites may compromise critical missions if subjected to intrusions or tampering. The integration of AI onboard requires that platforms be protected against unauthorized access and that models and data be encrypted to prevent manipulation [84,85]. Encryption and authentication mechanisms are essential to protect AI models and inferences, ensuring that communications between satellites and ground stations are authenticated and secure. AI could also be used to monitor system security in real time, detecting anomalous behaviours that could signal cyberattacks or system malfunctions.

Moreover, validating and verifying AI models in the space environment is a significant methodological challenge. Models must undergo extensive testing before launch and be monitored throughout the mission to ensure that they operate as expected, despite varying environmental conditions and limited resources. Testing in simulated environments, replicating space conditions such as radiation and temperature fluctuations, is crucial. Additionally, fallback models or simpler versions must be available in case primary models fail or behave unexpectedly due to unforeseen conditions.

3.2. SW Approaches for AI Integration

The integration of AI models into embedded platforms onboard satellites involves addressing a range of complex technological and methodological challenges. Among these, quantization [86] plays a key role by reducing the precision of numerical data, such as converting from 32-bit floating point to 8-bit integer, which significantly decreases memory usage and enhances inference speed. However, quantization introduces its own set of challenges. It is essential to balance computational efficiency with the accuracy of the model, as this process can result in rounding errors and a loss of precision. Testing and optimizing models becomes crucial to ensure that the quality of inference remains acceptable. Additionally, hardware compatibility poses a challenge since not all embedded devices support every type of quantization. Some devices may natively support 8-bit quantization, while others might require additional optimizations.

In this context, the use of lightweight frameworks and optimized runtimes, such as TensorFlow Lite, ONNX Runtime, and PyCoral API, is important for efficiently executing AI models on hardware with limited resources. However, ensuring compatibility between the selected framework and the satellite's hardware is critical, as not all frameworks support the required hardware architectures or model operations. Furthermore, the performance of these frameworks can vary depending on the device's specifications and the complexity of the model, necessitating thorough testing on actual hardware to achieve optimal results.

Another key approach is model compression [87,88], which includes techniques such as pruning, sparsity, and model distillation. These methods help reduce model complexity while maintaining their effectiveness. Nonetheless, pruning can negatively impact performance if not carried out correctly, so it is essential to test and validate the model after applying this technique to ensure satisfactory performance. Distillation, on the other hand, requires a training and fine-tuning process, which can be computationally intensive. In an onboard environment, it is vital to strike a balance between the effectiveness of distillation and the computational demands it places on the system.

Additionally, improving efficiency through pipelining and parallelization, where tasks are divided and executed simultaneously, introduces further complexity. Managing concurrency and data synchronization becomes a critical issue. It is essential to design the system carefully to prevent conflicts and ensure that data are processed accurately. Balancing the workload across resources is also crucial to avoid overloading some processing units while others remain idle, which demands thorough analysis and optimization.

In this subsection, we introduce some of the most recent works aimed at addressing the problems related to the integration of AI-based algorithms onboard satellites.

In [89], the authors show how to reduce the size of a CNN to reduce the upload time needed to transmit the CNN's weights from the ground to the satellite. In particular, starting from the CNN introduced in [90] (Baseline), they reduced the size of this CNN, creating two smaller versions (Reduced and Logistic) by reducing both the number of layers and channels per layer. This allowed a reduction of around 98.9% of the number of CNN parameters. Moreover, to reduce as much as possible the file size containing the CNN's weights, the authors converted the weights into 16- and 8-bit values. The conversion to 8 bit was carried as shown in Equation (1).

$$f\{x\} = \begin{cases} 127 & \text{, if } x \cdot 100 > 127 \\ -127 & \text{, if } x \cdot 100 < -127 \\ \lfloor x \cdot 100 \rceil & \text{, otherwise} \end{cases} \quad (1)$$

The authors considered the CNN inference performed as FP32, so no dedicated onboard hardware is needed during inference time. To evaluate the performance of the reduced CNNs, they considered the task of classifying plasma regions in Earth's magnetosphere. Concerning the accuracy, they found that the Baseline and Reduced CNNs showed quite similar accuracy, while the Logistic CNN showed a slightly higher accuracy. Moving to the file size, it is clear that reducing the bit widths of the weights from 32 to 8 greatly reduces the amount of data to be transmitted. Considering both the smaller version of the proposed CNN (the Logistic one) and the reduced bit width, the authors estimate a reduction time of 240×, proving the value of developing smaller CNNs for dedicated tasks.

In [91], the authors propose a novel approach for detecting changes in satellite imagery using Auto-Associative Neural Network (AANN). Their goal is to provide efficient change detection methods designed to operate onboard satellites, allowing for real-time processing and analysis. Solutions like this are useful to efficiently handle large amounts of data that are generated by the satellite during sensing applications that produce high-resolution images; storing and transmitting to the ground only the images that are valuable to the desired application could be a promising solution in these cases. The authors utilize Sentinel-2 imagery, focusing on urban changes while excluding natural variations. The proposed AANN compresses input images into lower-dimensional representations, which are then analyzed to determine changes by calculating the Euclidean distance between feature vectors derived from different time points. The proposed AANN provides a higher True Positive Rate (TPR), F1 score, and recall compared to Discrete Wavelet Transform (DWT).

In [92], the authors introduce a benchmarking pipeline for evaluating the performance of deep learning algorithms on edge devices for space applications. The proposed approach is model-agnostic and can be used to assess various deep learning algorithms, with a focus on computer vision tasks for space missions. The evaluation involves three use cases: detection, classification, and hyperspectral image segmentation. The key contributions of the proposed work are as follows:

- A benchmarking pipeline that utilizes standard Xilinx tools and a deep learning deployment chain to run deep learning techniques on edge devices;
- Performance evaluation of three state-of-the-art deep learning algorithms for computer vision tasks on the Leopard DPU Evalboard, which will be deployed onboard the Intuition-1 satellite;
- The analysis focuses on the latency, throughput, and performance of the models. Quantification of deep learning algorithm performance at every step of the deployment chain allows practitioners to understand the impact of crucial deployment steps like quantization or compilation on the model's operational abilities.

The tests were run on the Leopard DPU Evalboard.

In [93], the authors propose an approach for cloud detection using CNNs specifically designed for remote sensing applications. This study emphasizes the importance of efficient cloud screening in satellite imagery to enhance the quality of data for various applications, including climate monitoring and environmental assessments. They propose an encoder–decoder CNN structure and they built training and test samples using the SPARCS (Spatial Procedures for Automated Removal of Cloud and Shadow) [94] dataset. The main metrics taken into consideration by the authors are as follows: F1 score, mean Intersection over Union (mIoU), and overall accuracy. Different parameters were evaluated to understand their impact on the CNN performance:

- Input Bands: the network was tested with varying numbers of spectral bands, revealing that using four specific bands (red, green, blue, and infrared) yielded the best results;

- Input Size: different input sizes were assessed, with a size of 128 × 128 pixels providing optimal accuracy while reducing memory usage;
- Precision: experiments with half-precision computations demonstrated significant memory savings but at the cost of performance degradation in segmentation tasks;
- Convolutional Filters: adjusting the number of filters affected both memory consumption and accuracy, suggesting a balance is necessary to maintain valuable performance while optimizing resources;
- Encoder Depth: utilizing deeper residual networks improved performance metrics significantly but increased memory usage and inference time.

The authors showed that the proposed CNN outperformed existing state-of-the-art methods like Deeplab V3+ [95] in terms of accuracy and memory efficiency. The findings indicate that careful tuning of network parameters can lead to substantial improvements in cloud detection capabilities without excessive resource demands. Thus, the proposed CNN-based approach is effective for onboard cloud screening of satellite imagery. It highlights the potential for real-time processing in spaceborne systems, paving the way for enhanced data quality in remote sensing applications. This research contributes valuable insights into optimizing neural network architectures for specific tasks in remote sensing, particularly concerning resource constraints typical in onboard systems. Figure 5 shows the neural network's architecture for the cloud screening.

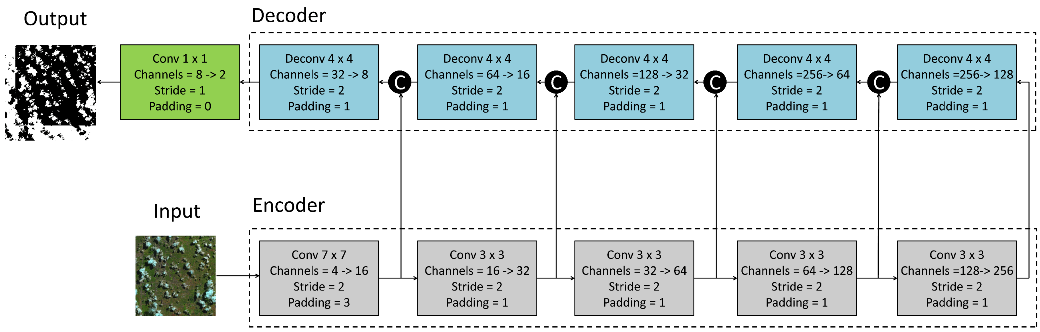

Figure 5. Cloud screening neural network architecture. Image taken from [93].

In [96], the authors investigate the use of SNNs on neuromorphic processors to improve energy efficiency in AI-based tasks for satellite communication. Their work aims to compare the performance and power consumption of various satellite communication applications using traditional AI accelerators versus neuromorphic hardware. They took into consideration three use cases: payload resource optimization; onboard interference detection and classification; and dynamic receive beamforming. The authors compare the performance of conventional CNNs implemented on Xilinx's VCK5000 Versal development card with SNNs running on Intel's Loihi 2 chip. Loihi 2 [97,98] is Intel's second-generation neuromorphic chip, introduced in 2021. It is designed for high-efficiency, event-driven computations, making it suitable for applications such as machine learning and real-time data processing, and it is supported by Lava [99], an open-source framework for developing neuro-inspired applications. The findings suggest that neuromorphic hardware could significantly enhance the efficiency of onboard AI systems in power-constrained environments like satellites, providing a promising avenue for future advancements in satellite communication technology.

In [100], the authors present a novel approach for onboard ship detection utilizing a custom hardware-oriented CNN, referred to as HO-ShipNet. This system is specifically designed for detecting ships in optical satellite images, which is crucial for maritime monitoring and safety. The HO-ShipNet architecture (shown in Figure 6) is tailored for embedded systems, enabling efficient processing and deployment on platforms with limited computational resources, such as satellites or drones. This model achieved high accuracy in distinguishing between ship and non-ship scenarios, which is critical for effective maritime surveillance. The authors highlight the increasing importance of integrating advanced AI technologies in maritime applications, showcasing how embedded neural networks can enhance real-time decision-making processes in complex environments. The main tasks taken into consideration are as follows:

- Maritime Security: enhancing the capability to monitor illegal fishing, smuggling, and other illicit activities at sea;
- Environmental Monitoring: assisting in tracking oil spills and other environmental hazards;
- Traffic Management: improving the management of shipping routes and port operations.

Notably, the authors emphasize explainability, allowing users to understand the decision-making process of the neural network. This is particularly important in applications where trust and transparency are essential. Figure 7 shows the architecture, including both PL and PS developed for the HO-ShipNet network.

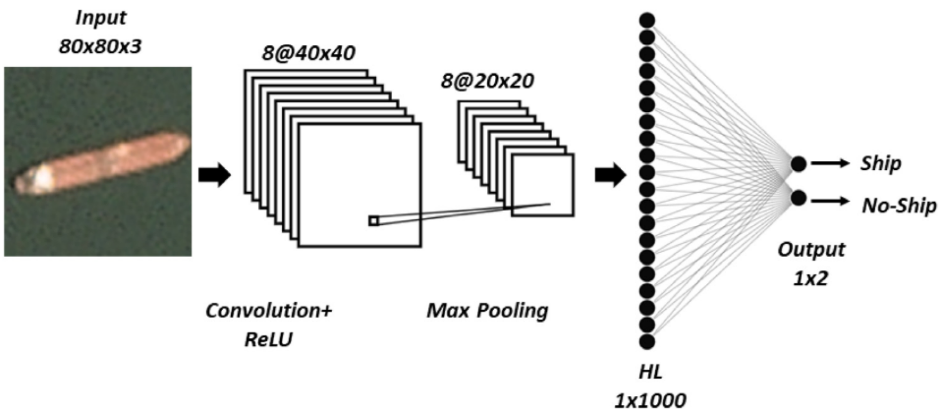

Figure 6. HO-ShipNet architecture. Image taken from [100].

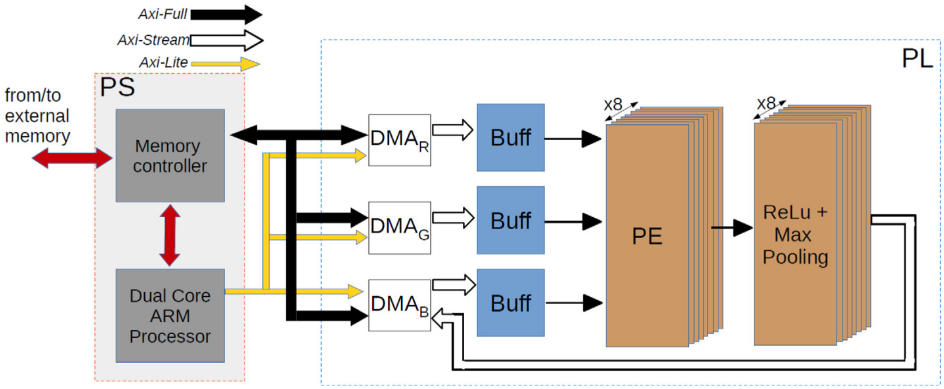

Figure 7. PL and PS architecture developed for the HO-ShipNet. Image taken from [100].

4. AI Applications in Satellites Tasks and Missions

AI is emerging as a key technology in the space sector, particularly aboard satellites. Its applications range from enhancing the operational efficiency of satellites to gathering and analyzing scientific data, as well as boosting the capabilities of exploratory missions. In this section, we provide an overview of the main AI applications in the space context that can benefit from AI-based onboard processing capability and provide works specifically designed to improve each of these applications.

4.1. Earth Observation and Image Analysis

One of the primary applications of AI aboard satellites is in Earth observation, where AI is used to analyze vast numbers of images and amounts of Geospatial data. Satellites equipped with AI can detect environmental changes, such as deforestation, glacier melting, pollution, and natural disasters like hurricanes or wildfires. AI helps filter and select the most relevant images for transmission to Earth, reducing the amount of data sent and speeding up decision-making processes. Other benefits of adopting AI-powered solutions onboard satellites for Earth observation applications include enabling direct observation and recognition of natural disasters from orbit, thus enabling faster disaster response and mitigation. Nonetheless, there are many challenges in integrating AI-based solutions onboard satellites, especially small satellites. These challenges mainly concern hardware resource consumption such as power consumption, and the availability of radiation-tolerant hardware accelerators while maintaining good performance of the AI models.

In [58], the authors propose the CloudScout CNN for onboard cloud coverage detection. The aim is to identify hyperspectral images that are covered by the cloud and thus that do not have to be downloaded to the ground station, saving time and energy resources. The dataset was built using hyperspectral images from the Sentinel-2 mission. The proposed CNN has been tailored to run on the Intel Myriad 2 VPU to achieve a valuable trade-off between low inference time and low-power consumption. CluodScout achieves 92% accuracy, a 1% False Positive (FP) rate, an inference time of 325 ms, and a power consumption of 1.8 W. The model footprint, around 2.1 MB, makes it feasible to upload a new version of the CNN during the mission life cycle.

The Φ-Sat-1 mission, launched on 3 September 2020, is a pioneering project by the European Space Agency (ESA) that demonstrates the use of AI for Earth observation [101]. This mission is notable for being the first to implement an onboard deep neural network (DNN) for processing satellite imagery directly in space. Key features of the Φ-Sat-1 mission include the following:

- AI Integration: The satellite employs a Convolutional Neural Network to perform cloud detection, filtering out unusable images before they are transmitted back to Earth. This enhances data efficiency by significantly reducing the volume of images sent down, which is crucial given that cloud cover can obscure over 30% of satellite images.
- Technological Demonstrator: as part of the Federated Satellite Systems (FSSCat) initiative, Φ-Sat-1 serves as a technological demonstrator, showcasing how AI can optimize satellite operations and data collection.
- Payload: the satellite is equipped with a hyperspectral camera and an AI processor (Intel Movidius Myriad 2), which enables it to analyze and process data onboard.
- Future Developments: following the success of Φ-Sat-1, plans are already in motion for the Φ-Sat-2 mission, which aims to expand on these capabilities by integrating more advanced AI applications for various Earth observation tasks.

Overall, Φ-Sat-1 marks a significant step forward in utilizing AI technologies in space, setting the stage for more sophisticated applications in future missions.

The authors of [102] discuss advancements in image processing techniques for high-resolution Earth observation satellites, specifically focusing on the French space agency (CNES) experience with satellites like Pléiades. The paper focuses on the following:

- Image Acquisition: high-resolution satellites capture panchromatic images with fine spatial resolution and multispectral images with coarser sampling due to downlink constraints.
- Processing Chain: the paper outlines a next-generation processing chain that includes onboard compression, correction of compression artifacts, denoising, deconvolution, and pan-sharpening techniques.
- Compression Techniques: a fixed-quality compression method is detailed, which minimizes the impact of compression on image quality while optimizing bitrate based on scene complexity.
- Denoising Performance: the study shows that non-local denoising algorithms significantly improve image quality, outperforming previous methods by 15% in terms of root mean squared error.

- Adaptation for CMOS Sensors: the authors also discuss adapting these processing techniques for low-cost CMOS Bayer colour matrices, demonstrating the versatility of the proposed image processing chain.

This research contributes to enhancing the quality and efficiency of satellite imagery processing, which is crucial for various applications in remote sensing and Earth observation.

As discussed in [103], the integration of AI-based image recognition in small-satellite Earth observation missions offers a mix of substantial opportunities and notable challenges. One of the key advantages is the increased efficiency in handling data. AI can process images directly onboard the satellite, filtering out unusable data—such as images obscured by clouds—before transmission. This capability significantly reduces the volume of data sent back to Earth, making better use of bandwidth and lowering transmission costs. Additionally, satellites equipped with AI gain enhanced autonomy, enabling them to analyze data in real time and make decisions about which information to transmit. This results in faster response times, particularly valuable for applications like disaster monitoring and environmental assessments. Moreover, AI supports more advanced functionalities, including cloud detection, object recognition, and anomaly identification directly in space. This leads to more actionable insights, increasing the utility of the satellite's imagery for diverse stakeholders. The use of onboard AI also contributes to cost reduction by minimizing the data processing requirements on the ground, thus cutting operational expenses. However, the integration of AI into space missions presents several challenges. The hardware used in space is subject to strict limitations in terms of power consumption and processing capacity. AI models must be both efficient and lightweight, which can restrict their complexity and effectiveness. In addition, the space environment exposes hardware to radiation, which can affect the reliability of AI systems, requiring more robust designs to ensure uninterrupted operation. Developing AI models for space also involves extensive training with relevant datasets, which is often complicated by the unique conditions of space missions. These models must be adaptable to handle a variety of operational scenarios. Furthermore, the integration of AI into existing satellite systems is a technically demanding process, involving compatibility issues with software, data formats, and communication protocols. This adds complexity to mission planning and execution. In conclusion, while AI-based image recognition holds transformative potential for enhancing small-satellite Earth observation missions, careful consideration of the associated technical challenges is essential for successful implementation.

In [104], the authors propose a distillation method to reduce the size of deep neural networks (DNNs) for satellite onboard image segmentation, specifically for ship detection tasks. The goal is to simplify large, high-performance deep neural networks (DNNs) to fit within the limited computational resources of CubeSat platforms while minimizing accuracy loss. The motivation of this work is that Cubesat satellites can benefit from onboard data reduction and data reduction can be effectively performed using modern deep neural networks (DNNs). At the same time, modern deep neural networks (DNNs) are often too big and require too many computation capabilities. With this aim, the authors propose to use a teacher–student approach to reduce the size of state-of-the-art Deep Neural Networks (DNNs). In particular, they use a distillation process to train a smaller student DNN to mimic the outputs of the teacher DNN. The student DNN is optimized to have around 1 million parameters to fit within Cubesat processing payloads. They use a weighted MSE loss function to train the student DNN. The authors propose a distillation-based method to significantly reduce the size of DNNs for onboard satellite image segmentation while maintaining high accuracy, enabling their use within the constraints of CubeSat platforms. Moreover, they highlight that combining distillation with other compression methods like pruning and quantization can achieve even higher reduction rates.

The integration of onboard anomaly detection systems in marine environmental protection leverages artificial intelligence to identify threats to marine ecosystems, such as oil spills, harmful algal blooms, and sediment discharges. The European Space Agency's

Φ-sat-2 mission features a marine anomaly detection application that assigns an anomaly score to maritime images [105]. Figure 8 shows the architecture of this solution. This score increases as the image content deviates from normal water conditions. When the score exceeds a certain threshold, alerts are triggered to facilitate rapid responses from authorities. The main key features of the proposed solution are as follows:

- Image Prioritization: the system prioritizes downloading images with the highest anomaly scores, optimizing data transmission and operator efficiency.
- Alert Mechanism: immediate alerts are sent for significant incidents detected onboard, enhancing response times.
- AI Efficiency: the application uses minimal annotated data for training, making it suitable for satellites with limited computational power and adaptable to various environments beyond marine settings.

This innovative approach demonstrates the potential of AI in enhancing maritime monitoring and environmental protection efforts, establishing a framework that can be utilized across different ecosystems and satellite missions.

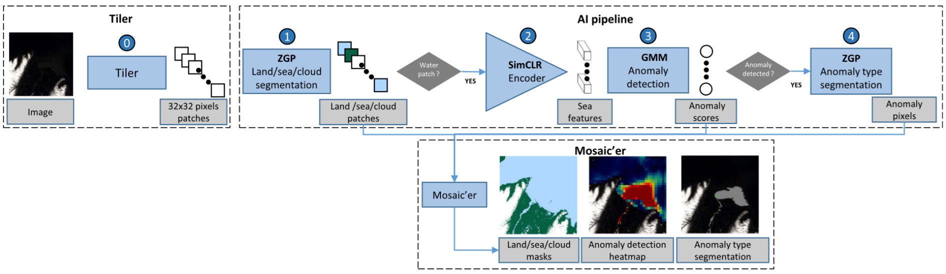

Figure 8. Architecture of the anomaly detection pipeline proposed in [105]. Image taken from [105].

4.2. *Anomaly Detection and Predictive Maintenance*

AI is widely used for monitoring and predictive maintenance of satellites. By analyzing operational data, machine learning algorithms can detect anomalies in onboard systems and predict failures before they occur. This improves satellite management, reduces unplanned downtime, extends the operational life of equipment, and enhances mission reliability [106,107].

In [108], the authors provide an in-depth exploration of various machine learning (ML) methods aimed at improving the detection of anomalies in satellite telemetry data. They highlight different types of anomalies that can occur in telemetry, emphasizing the importance of efficient detection strategies to maintain satellite health and performance. The paper explores several machine learning approaches, such as Gaussian Mixture Models (GMMs), which model the distribution of telemetry data to identify outliers, and the Local Outlier Factor (LOF), a technique that detects anomalies by assessing the local density of data points. It also discusses the use of autoencoders—neural networks that learn to represent data efficiently—making them particularly useful for identifying deviations from normal behaviour. The research applies these techniques to actual satellite telemetry data, evaluating their effectiveness in detecting anomalies. However, the dataset used in the evaluation presents certain challenges, such as limited time coverage and a scarcity of critical anomalies, requiring careful performance assessment of each algorithm. Despite these challenges, the paper concludes that machine learning techniques offer significant potential for improving anomaly detection in satellite telemetry. These methods can reduce reliance on human operators and increase the overall operational efficiency of satellite systems. Nonetheless, the study acknowledges that further research and refinement of these approaches are necessary to fully adapt them for practical use in space missions.

The authors in [109] outline a framework designed to monitor and predict the health of Control Moment Gyroscopes (CMGs), which play a crucial role in the attitude control systems of satellites. The study introduces a Prognostic Health Management (PHM) framework, tailored specifically for CMGs, to emphasize the need for early detection of potential failures and to estimate the Remaining Useful Life (RUL) of these critical components. This proactive approach aims to prevent unexpected failures and enhance satellite reliability. The framework leverages historical telemetry data from CMGs, using them to train models that can detect failure patterns and predict the future operational lifespan of the gyroscopes. Various machine learning techniques are explored in this context, including both regression models and classification algorithms, which analyze the telemetry data to forecast possible failures. The models are evaluated based on their performance metrics, specifically their accuracy in predicting failures and estimating the RUL. The ultimate goal is to ensure that these models can operate reliably within the challenging conditions of space. In terms of practical applications, the implementation of this PHM framework can significantly enhance operational efficiency. By enabling proactive maintenance, the framework can extend the lifespan of CMGs and help reduce the likelihood of unforeseen failures. Moreover, the approach has broader applications beyond CMGs; it can be adapted for other satellite components and systems, contributing to a comprehensive strategy for satellite health management. This research underlines the importance of predictive analytics in the space sector, offering a pathway to improved reliability and reduced operational costs for satellite systems.

The authors in [110] explore a new cloud detection system using Convolutional Neural Networks (CNNs) on commercial off-the-shelf (COTS) microcontrollers. This system is designed for small satellites to autonomously analyze and prioritize cloud-free images for transmission, thereby optimizing data collection and improving the quality of Earth observation data. By deploying the CNN on COTS hardware, the research demonstrates that commercial components can effectively perform machine learning tasks in space, making the technology both accessible and cost-effective. The study also examines how quantization affects CNN performance by comparing results from the embedded system with those from a more powerful PC setup. Despite the constraints, the embedded system achieves performance comparable to more advanced platforms. Overall, the findings suggest that integrating AI with commerical off-the-shelf (COTS) components can significantly enhance satellite operations, paving the way for more efficient real-time data processing in future space missions.

In [111], the authors explore the integration of machine learning (ML) technologies in satellite systems, specifically focusing on Failure Detection, Isolation, and Recovery (FDIR) functions. As satellite technology becomes increasingly complex, traditional FDIR methods are deemed inadequate for handling multiple simultaneous failures and the prognosis of future issues, leading to potentially catastrophic outcomes. The authors highlight the potential of machine learning (ML) algorithms for improving error detection and prognosis through advanced onboard systems that can operate under strict power, mass, and radiation constraints. Machine learning (ML) algorithms can enhance the autonomy and reliability of satellite operations by learning from in-flight data to detect and respond to failures more effectively. This work also provides a comparison of various commercial off-the-shelf (COTS) edge AI boards, focusing on their power draw, processing capabilities (TOPS), and suitability for space applications. The boards taken into consideration in this paper are Nvidia Jetson Xavier NX, Huawei Atlas 200, Google Coral, and Intel Movidius Myriad 2. The results highlight how it is important to select an appropriate hardware device based on mission requirements.

4.3. Autonomous Navigation and Control

AI plays a crucial role in implementing autonomous navigation systems in satellites. These systems allow satellites to make real-time decisions regarding routes and manoeuvres, reducing their dependence on ground-based control. In space exploration missions, AI enables probes and satellites to autonomously adjust their trajectories to avoid obstacles or modify their orbits in response to environmental changes. This capability is particularly valuable in complex environments, such as missions to asteroids or distant planets [112,113].

In [114], the authors present a novel Guidance, Navigation, and Control (GNC) system designed to enhance the autonomy and efficiency of spacecraft during on-orbit manipulation tasks. This research is driven by the increasing demand for advanced robotic manipulation in space, particularly for servicing missions involving non-cooperative targets like malfunctioning satellites. The proposed solution includes two main systems:

- AI Modules. The proposed GNC system incorporates two state-of-the-art AI components:
  - Deep Learning (DL)-based Pose Estimation: this algorithm estimates the pose of a target from 2D images using a pre-trained neural network, eliminating the need for prior knowledge about the target's dynamics;
  - Trajectory Modeling and Control: this technique utilizes probabilistic modeling to manage the trajectories of robotic manipulators, allowing the system to adapt to new situations without complex on-board trajectory optimizations. This minimizes disturbances to the spacecraft's attitude caused by manipulator movements.
- Centralized Camera Network. The system employs a centralized camera network as its primary sensor, integrating a 7 Degrees of Freedom (DoF) robotic arm into the GNC architecture.

The intelligent GNC system was tested through simulations of a conceptual mission named AISAT, which involves a micro-satellite performing manipulations around a non-cooperative CubeSat. The simulations were conducted in Matlab/Simulink, utilizing a physics rendering engine to visualize the operations realistically.

Intelligent GNC architectures could serve as a foundation for developing fully autonomous orbital robotic systems.

In [115], the authors present the DeepNav initiative, which is a research project funded by the Italian Space Agency (ASI). DeepNav aims to enhance autonomous navigation techniques for small satellites operating in deep space, particularly around asteroids. This is crucial for missions that require precise maneuvering and data collection from these celestial bodies. This project is set to last for 18 months, during which various methodologies will be explored and tested. DeepNav's key Features are as follows:

- Autonomous Navigation. DeepNav focuses on creating systems that allow satellites to navigate without constant human intervention, which is vital for missions that operate far from Earth;
- Deep Learning Techniques. The project leverages deep learning algorithms to process and analyze data collected from asteroid surfaces, enabling better decision-making and navigation strategies.

The project represents a significant step forward in the field of aerospace engineering, particularly in the context of small satellite missions.

In [116], the authors present the FUTURE mission. This mission aims at developing an innovative approach to enhancing spacecraft autonomy, specifically focusing on orbit determination using AI. The FUTURE mission aims to reduce reliance on ground operators by improving the onboard autonomy of a $6U$ CubeSat. This will be achieved through advanced optical sensors and AI-based algorithms that process data for positional knowledge. Its key components are as follows:

- Optical sensors. The CubeSat will be equipped with optical sensors to gather data about Earth features such as lakes and coastlines;
- AI processing. The data collected will be processed onboard to generate positional inputs for navigation filters, enhancing the accuracy of orbit determination.

The authors introduce the navigation filter architecture, a preliminary design of the navigation filter focusing on how it will process data from the optical sensors, and then they assesses the potential accuracy of orbit determination achievable through onboard processing. Finally, they discuss opportunistic observations of celestial objects, such as the

Moon, to validate autonomous navigation methods during specific flight conditions. The future phases of the mission will include further enhancements to the navigation filter. This could lead to improved autonomous operations not just in low Earth orbit (LEO) but also in missions targeting other celestial bodies, thereby expanding the operational capabilities of CubeSats in deep space exploration.

4.4. Data Management, Data Compression, and Communication Optimization

AI facilitates the intelligent management of data collected onboard satellites by optimizing data compression and selecting the most relevant information for transmission. This is critical since satellite communication resources are limited, and transmitting large volumes of data can be costly and slow. Using AI, satellites can process data onboard, identify the most valuable scientific or operational information, and transmit only this to Earth.

Moreover, AI can be used to optimize communication networks, improving resource management and load distribution among various satellites. AI algorithms can dynamically adjust the network to better handle data traffic, minimize interference, and ensure optimal coverage [117].

In [118], the authors addresses the significant challenge of data transmission from satellites to Earth, particularly in Earth observation applications. The authors highlight that satellite downlink bandwidth is often a limiting factor for transmitting high-resolution images, which can hinder timely data delivery and analysis.To address this issue, they introduce a novel approach utilizing federated learning for onboard image compression. This method allows satellites to collaboratively learn compression algorithms without needing to send raw data back to Earth, thus conserving bandwidth. The proposed framework enables multiple satellites to train a shared model while keeping their data localized. This decentralized approach not only enhances privacy but also optimizes the learning process by leveraging diverse datasets from various satellites.

The authors conduct experiments demonstrating that their federated learning strategy significantly improves compression efficiency compared to traditional methods, thereby alleviating the downlink bottleneck. Improving image transmission capabilities could impact various applications in Earth observation, including environmental monitoring and disaster response.

In [119], the authors explore the development of an advanced algorithm for compressing multispectral images directly on satellites. This approach is crucial for efficient data transmission and storage in Earth observation missions. The primary goal of this paper is to enhance the compression of multispectral images using AI techniques, which can significantly reduce the amount of data sent back to Earth. The proposed AI-based algorithm leverages machine learning (ML) to improve compression rates while maintaining image quality. This involves training models on existing datasets to optimize the compression process.

The implementation of the algorithm demonstrates improved performance compared to traditional methods, achieving higher compression ratios without substantial loss of critical image information. This technology is particularly beneficial for low Earth orbit (LEO) satellites, where bandwidth is limited and efficient data handling is essential for timely analysis and response.

In [120], the authors focus on enhancing the efficiency of data processing in small satellites, particularly in the context of hyperspectral imaging. The authors leverage deep learning inference techniques to optimize onboard data processing to address the significant challenge posed by the large volumes of data generated by hyperspectral sensors, which complicates the downlinking process to ground stations. The study proposes an onboard processing framework that utilizes deep learning (DL) algorithms to analyze and compress hyperspectral data before transmission, thus reducing the amount of data that needs to be downlinked and minimizing bandwidth usage. The findings indicate that employing deep learning (DL) models can improve data processing times and significantly enhance the accuracy and speed of reflectance retrievals from satellite data.

The increasing volume of SAR data necessitates efficient processing methods to overcome limitations in downlink capacity and reduce latency in data products. In [11], the authors discuss how the employment of deep learning (DL) and machine learning (ML) algorithms can help in reducing the bandwidth requirement and enhance the real-time capabilities, which is crucial for time-sensitive applications like disaster response and environmental monitoring. The authors also highlight the challenges related to onboard processing, including both the onboard computing power and the power consumption. To overcome this challenges it is mandatory to develop efficient algorithms tailored to the onboard processing capabilities.

In [121], the authors explore the transformative role of artificial intelligence (AI) in enhancing satellite communication systems. It emphasizes how AI technologies can optimize various aspects of satellite operations, addressing the increasing demand for connectivity in a rapidly evolving digital landscape. They introduce a traffic demand prediction method utilizing deep learning (DL)-based algorithms. This approach aims to effectively manage the dynamic traffic demands faced by satellite networks, ensuring efficient resource allocation and service delivery under varying conditions. Moreover, the authors highlighted AI as a crucial tool for improving operational efficiency in satellite communications. The integration of AI can automate and streamline processes such as data processing, signal analysis, and anomaly detection, thereby reducing operational costs and enhancing service reliability.

5. Conclusions and Future Works

The integration of AI models, particularly NNs, into embedded platforms onboard satellites represents a significant advancement in satellite technology. This review has highlighted the various hardware solutions, including ASICs, FPGAs, VPUs, and neuromorphic hardware, that enable efficient NN inference in space. Each hardware type offers unique advantages in terms of computational efficiency, power consumption, and suitability for the harsh conditions of space.

ASICs like the Google Edge TPU and Nvidia Jetson boards have demonstrated high computational efficiency and low power consumption, making them suitable for short-duration satellite missions. FPGAs, while offering flexibility and re-programmability, provide higher radiation tolerance, which is essential for long-duration missions. Neuromorphic hardware and memristors present promising approaches for developing advanced AI systems that operate efficiently in space, offering superior computational efficiency and lower power consumption.

The commercial solutions discussed, such as Ingeniars' GPU@SAT, MHI's AIRIS, and Blue Marble Communications' SEP, showcase the industry's efforts to develop radiation-hardened devices for AI acceleration onboard satellites. These solutions enhance the computational capabilities of satellites, enabling real-time data processing and reducing the reliance on ground-based systems.

Despite these advancements, several challenges remain. The limited computational and memory resources onboard satellites, the need for fault-tolerant systems, and the harsh space environment pose significant obstacles to the deployment of complex AI models. Additionally, the integration of AI models into satellite systems introduces cybersecurity risks that must be addressed to ensure the integrity and reliability of satellite operations.

Future research should focus on the following areas to further enhance the integration of AI onboard satellites:

- Optimization of AI Models: Developing lightweight and energy-efficient AI models that can operate within the constraints of satellite hardware. Techniques such as model compression, pruning, and quantization should be explored to reduce the computational requirements of AI models.

- Radiation Tolerance: Enhancing the radiation tolerance of AI hardware through the development of radiation-hardened components and fault-tolerant systems. This includes testing and validating AI hardware in simulated space environments to ensure reliability.
- Neuromorphic and Memristor-Based Systems: Further research into neuromorphic hardware and memristor-based systems for space applications. These technologies offer promising solutions for developing energy-efficient AI systems that can operate in the harsh conditions of space.
- Cybersecurity: Implementing robust cybersecurity measures to protect AI models and data from unauthorized access and manipulation. This includes encryption, authentication, and continuous monitoring for potential threats.
- Real-Time Data Processing: Developing advanced AI algorithms for real-time data processing onboard satellites. This includes applications such as Earth observation, anomaly detection, and autonomous navigation, which require low-latency responses.
- Collaboration and Standardization: Promoting collaboration between industry, academia, and government agencies to standardize AI hardware and software solutions for space applications. This will facilitate the development of interoperable systems and accelerate the adoption of AI technologies in the space sector.

In conclusion, the integration of AI onboard satellites holds great promise for enhancing the capabilities of satellite missions. By addressing the current challenges and focusing on future research directions, we can unlock the full potential of AI in space, enabling more efficient, autonomous, and resilient satellite operations.

Author Contributions: All authors contributed equally to this research work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, C.; Jin, J.; Kuang, L.; Yan, J. LEO constellation design methodology for observing multi-targets. *Astrodynamics* **2018**, *2*, 121–131. [CrossRef]
2. Jaffer, G.; Malik, R.A.; Aboutanios, E.; Rubab, N.; Nader, R.; Eichelberger, H.U.; Vandenbosch, G.A. Air traffic monitoring using optimized ADS-B CubeSat constellation. *Astrodynamics* **2024**, *8*, 189–208. [CrossRef]
3. Bai, S.; Zhang, Y.; Jiang, Y.; Sun, W.; Shao, W. Modified Two-Dimensional Coverage Analysis Method Considering Various Perturbations. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *60*, 2763–2777. [CrossRef]
4. ESA. Artificial Intelligence in Space. Available online: https://www.esa.int/Enabling_Support/Preparing_for_the_Future/Discovery_and_Preparation/Artificial_intelligence_in_space (accessed on 20 October 2024).
5. Thangavel, K.; Sabatini, R.; Gardi, A.; Ranasinghe, K.; Hilton, S.; Servidia, P.; Spiller, D. Artificial intelligence for trusted autonomous satellite operations. *Prog. Aerosp. Sci.* **2024**, *144*, 100960. [CrossRef]
6. Thangavel, K.; Spiller, D.; Sabatini, R.; Amici, S.; Longepe, N.; Servidia, P.; Marzocca, P.; Fayek, H.; Ansalone, L. Trusted autonomous operations of distributed satellite systems using optical sensors. *Sensors* **2023**, *23*, 3344. [CrossRef]
7. Al Homssi, B.; Dakic, K.; Wang, K.; Alpcan, T.; Allen, B.; Boyce, R.; Kandeepan, S.; Al-Hourani, A.; Saad, W. Artificial Intelligence Techniques for Next-Generation Massive Satellite Networks. *IEEE Commun. Mag.* **2024**, *62*, 66–72. [CrossRef]
8. Nanjangud, A.; Blacker, P.C.; Bandyopadhyay, S.; Gao, Y. Robotics and AI-Enabled On-Orbit Operations With Future Generation of Small Satellites. *Proc. IEEE* **2018**, *106*, 429–439. [CrossRef]
9. Alves de Oliveira, V.; Chabert, M.; Oberlin, T.; Poulliat, C.; Bruno, M.; Latry, C.; Carlavan, M.; Henrot, S.; Falzon, F.; Camarero, R. Satellite Image Compression and Denoising With Neural Networks. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
10. Guerrisi, G.; Schiavon, G.; Del Frate, F. On-Board Image Compression using Convolutional Autoencoder: Performance Analysis and Application Scenarios. In Proceedings of the IGARSS 2023—2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023; pp. 1783–1786. [CrossRef]
11. Garcia, L.P.; Furano, G.; Ghiglione, M.; Zancan, V.; Imbembo, E.; Ilioudis, C.; Clemente, C.; Trucco, P. Advancements in On-Board Processing of Synthetic Aperture Radar (SAR) Data: Enhancing Efficiency and Real-Time Capabilities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 16625–16645. [CrossRef]
12. Guerrisi, G.; Frate, F.D.; Schiavon, G. Artificial Intelligence Based On-Board Image Compression for the Φ-Sat-2 Mission. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 8063–8075. [CrossRef]
13. Russo, A.; Lax, G. Using artificial intelligence for space challenges: A survey. *Appl. Sci.* **2022**, *12*, 5106. [CrossRef]

14. Ortiz, F.; Monzon Baeza, V.; Garces-Socarras, L.M.; Vasquez-Peralvo, J.A.; Gonzalez, J.L.; Fontanesi, G.; Lagunas, E.; Querol, J.; Chatzinotas, S. Onboard processing in satellite communications using ai accelerators. *Aerospace* **2023**, *10*, 101. [CrossRef]
15. Leyva-Mayorga, I.; Martinez-Gost, M.; Moretti, M.; Perez-Neira, A.; Vazquez, M.A.; Popovski, P.; Soret, B. Satellite Edge Computing for Real-Time and Very-High Resolution Earth Observation. *IEEE Trans. Commun.* **2023**, *71*, 6180–6194. [CrossRef]
16. Li, C.; Zhang, Y.; Xie, R.; Hao, X.; Huang, T. Integrating Edge Computing into Low Earth Orbit Satellite Networks: Architecture and Prototype. *IEEE Access* **2021**, *9*, 39126–39137. [CrossRef]
17. Zhang, Z.; Qu, Z.; Liu, S.; Li, D.; Cao, J.; Xie, G. Expandable on-board real-time edge computing architecture for Luojia3 intelligent remote sensing satellite. *Remote Sens.* **2022**, *14*, 3596. [CrossRef]
18. Pacini, T.; Rapuano, E.; Tuttobene, L.; Nannipieri, P.; Fanucci, L.; Moranti, S. Towards the Extension of FPG-AI Toolflow to RNN Deployment on FPGAs for On-board Satellite Applications. In Proceedings of the 2023 European Data Handling & Data Processing Conference (EDHPC), Juan Les Pins, France, 2–6 October; pp. 1–5. [CrossRef]
19. Razmi, N.; Matthiesen, B.; Dekorsy, A.; Popovski, P. On-Board Federated Learning for Satellite Clusters With Inter-Satellite Links. *IEEE Trans. Commun.* **2024**, *72*, 3408–3424. [CrossRef]
20. Meoni, G.; Prete, R.D.; Serva, F.; De Beusscher, A.; Colin, O.; Longépé, N. Unlocking the Use of Raw Multispectral Earth Observation Imagery for Onboard Artificial Intelligence. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 12521–12537. [CrossRef]
21. Pelton, J.N.; Finkleman, D. Overview of Small Satellite Technology and Systems Design. In *Handbook of Small Satellites: Technology, Design, Manufacture, Applications, Economics and Regulation*; Springer: Cham, Switzerland, 2020; pp. 125–144.
22. Manoj, S.; Kasturi, S.; Raju, C.G.; Suma, H.; Murthy, J.K. Overview of On-Board Computing Subsystem. In Proceedings of the Smart Small Satellites: Design, Modelling and Development: Proceedings of the International Conference on Small Satellites, ICSS 2022, Punjab, India, 29–30 April 2022; Springer Nature: Singapore, 2023; Volume 963, p. 23.
23. Cratere, A.; Gagliardi, L.; Sanca, G.A.; Golmar, F.; Dell'Olio, F. On-Board Computer for CubeSats: State-of-the-Art and Future Trends. *IEEE Access* **2024**, *12*, 99537–99569. [CrossRef]
24. Schäfer, K.; Horch, C.; Busch, S.; Schäfer, F. A Heterogenous, reliable onboard processing system for small satellites. In Proceedings of the 2021 IEEE International Symposium on Systems Engineering (ISSE), Vienna, Austria, 13 September–13 October 2021; pp. 1–3. [CrossRef]
25. Ray, A. Radiation effects and hardening of electronic components and systems: An overview. *Indian J. Phys.* **2023**, *97*, 3011–3031. [CrossRef]
26. Bozzoli, L.; Catanese, A.; Fazzoletto, E.; Scarpa, E.; Goehringer, D.; Pertuz, S.A.; Kalms, L.; Wulf, C.; Charaf, N.; Sterpone, L.; et al. EuFRATE: European FPGA Radiation-hardened Architecture for Telecommunications. In Proceedings of the 2023 Design, Automation & Test in Europe Conference & Exhibition (DATE), Antwerp, Belgium, 17–19 April 2023; pp. 1–6. [CrossRef]
27. Pavan Kumar, M.; Lorenzo, R. A review on radiation-hardened memory cells for space and terrestrial applications. *Int. J. Circuit Theory Appl.* **2023**, *51*, 475–499. [CrossRef]
28. Ghiglione, M.; Serra, V. Opportunities and challenges of AI on satellite processing units. In Proceedings of the 19th ACM International Conference on Computing Frontiers, New York, NY, USA, 17–22 May 2022; pp. 221–224. [CrossRef]
29. Tao, C.; Gao, J.; Wang, T. Testing and Quality Validation for AI Software–Perspectives, Issues, and Practices. *IEEE Access* **2019**, *7*, 120164–120175. [CrossRef]
30. Chen, G.; Guan, N.; Huang, K.; Yi, W. Fault-tolerant real-time tasks scheduling with dynamic fault handling. *J. Syst. Archit.* **2020**, *102*, 101688. [CrossRef]
31. Valente, F.; Eramo, V.; Lavacca, F.G. Optimal bandwidth and computing resource allocation in low earth orbit satellite constellation for earth observation applications. *Comput. Netw.* **2023**, *232*, 109849. [CrossRef]
32. Estébanez-Camarena, M.; Taormina, R.; van de Giesen, N.; ten Veldhuis, M.C. The potential of deep learning for satellite rainfall detection over data-scarce regions, the west African savanna. *Remote Sens.* **2023**, *15*, 1922. [CrossRef]
33. Wang, P.; Jin, N.; Davies, D.; Woo, W.L. Model-centric transfer learning framework for concept drift detection. *Knowl.-Based Syst.* **2023**, *275*, 110705. [CrossRef]
34. Khammassi, M.; Kammoun, A.; Alouini, M.S. Precoding for High-Throughput Satellite Communication Systems: A Survey. *IEEE Commun. Surv. Tutor.* **2024**, *26*, 80–118. [CrossRef]
35. Salim, S.; Moustafa, N.; Reisslein, M. Cybersecurity of Satellite Communications Systems: A Comprehensive Survey of the Space, Ground, and Links Segments. *IEEE Commun. Surv. Tutor.* 2024, *in press*. [CrossRef]
36. Elhanashi, A.; Gasmi, K.; Begni, A.; Dini, P.; Zheng, Q.; Saponara, S. Machine learning techniques for anomaly-based detection system on CSE-CIC-IDS2018 dataset. In Proceedings of the International Conference on Applications in Electronics Pervading Industry, Environment and Society, Genova, Italy, 26–27 September 2022; Springer: Cham, Switzerland, 2022; pp. 131–140.
37. Elhanashi, A.; Dini, P.; Saponara, S.; Zheng, Q. Integration of deep learning into the iot: A survey of techniques and challenges for real-world applications. *Electronics* **2023**, *12*, 4925. [CrossRef]
38. Chen, X.; Xu, Z.; Shang, L. Satellite Internet of Things: Challenges, solutions, and development trends. *Front. Inf. Technol. Electron. Eng.* **2023**, *24*, 935–944. [CrossRef]
39. Rech, P. Artificial Neural Networks for Space and Safety-Critical Applications: Reliability Issues and Potential Solutions. *IEEE Trans. Nucl. Sci.* **2024**, *71*, 377–404. [CrossRef]

40. Bodmann, P.R.; Saveriano, M.; Kritikakou, A.; Rech, P. Neutrons Sensitivity of Deep Reinforcement Learning Policies on EdgeAI Accelerators. *IEEE Trans. Nucl. Sci.* **2024**, *71*, 1480–1486. [CrossRef]
41. Buckley, L.; Dunne, A.; Furano, G.; Tali, M. Radiation test and in orbit performance of mpsoc ai accelerator. In Proceedings of the 2022 IEEE Aerospace Conference (AERO), Big Sky, MT, USA, 5–12 March 2022; pp. 1–9.
42. Dunkel, E.R.; Swope, J.; Candela, A.; West, L.; Chien, S.A.; Towfic, Z.; Buckley, L.; Romero-Cañas, J.; Espinosa-Aranda, J.L.; Hervas-Martin, E.; et al. Benchmarking deep learning models on myriad and snapdragon processors for space applications. *J. Aerosp. Inf. Syst.* **2023**, *20*, 660–674. [CrossRef]
43. Ramaswami, D.P.; Hiemstra, D.M.; Yang, Z.W.; Shi, S.; Chen, L. Single Event Upset Characterization of the Intel Movidius Myriad X VPU and Google Edge TPU Accelerators Using Proton Irradiation. In Proceedings of the 2022 IEEE Radiation Effects Data Workshop (REDW) (in Conjunction with 2022 NSREC), Provo, UT, USA, 18–22 July 2022; pp. 1–3. [CrossRef]
44. Boutros, A.; Nurvitadhi, E.; Ma, R.; Gribok, S.; Zhao, Z.; Hoe, J.C.; Betz, V.; Langhammer, M. Beyond Peak Performance: Comparing the Real Performance of AI-Optimized FPGAs and GPUs. In Proceedings of the 2020 International Conference on Field-Programmable Technology (ICFPT), Maui, HI, USA, 9–11 December 2020; pp. 10–19. [CrossRef]
45. Google. Edge TPU. Available online: https://coral.ai/products/ (accessed on 20 October 2024).
46. Google. TensorFlow Models on the Edge TPU. Available online: https://coral.ai/docs/edgetpu/models-intro (accessed on 20 October 2024).
47. Google. Run Inference on the Edge TPU with Python. Available online: https://coral.ai/docs/edgetpu/tflite-python/ (accessed on 20 October 2024).
48. Google. Models for Edge TPU. Available online: https://coral.ai/models/ (accessed on 20 October 2024).
49. Lentaris, G.; Leon, V.; Sakos, C.; Soudris, D.; Tavoularis, A.; Costantino, A.; Polo, C.B. Performance and Radiation Testing of the Coral TPU Co-processor for AI Onboard Satellites. In Proceedings of the 2023 European Data Handling & Data Processing Conference (EDHPC), Juan-Les-Pins, France, 2–6 October 2023; pp. 1–4. [CrossRef]
50. Rech Junior, R.L.; Malde, S.; Cazzaniga, C.; Kastriotou, M.; Letiche, M.; Frost, C.; Rech, P. High Energy and Thermal Neutron Sensitivity of Google Tensor Processing Units. *IEEE Trans. Nucl. Sci.* **2022**, *69*, 567–575. [CrossRef]
51. Goodwill, J.; Crum, G.; MacKinnon, J.; Brewer, C.; Monaghan, M.; Wise, T.; Wilson, C. NASA spacecube edge TPU smallsat card for autonomous operations and onboard science-data analysis. In Proceedings of the Small Satellite Conference, Virtual, 7–12 August 2021; SSC21-VII-08.
52. Nvidia. Jetson Orin for Next-Gen Robotics. Available online: https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/ (accessed on 20 October 2024).
53. Nvidia. Jetson Orin Nano Developer Kit Getting Started. Available online: https://developer.nvidia.com/embedded/learn/get-started-jetson-orin-nano-devkit (accessed on 20 October 2024).
54. Nvidia. TensorRT SDK. Available online: https://developer.nvidia.com/tensorrt (accessed on 20 October 2024).
55. Slater, W.S.; Tiwari, N.P.; Lovelly, T.M.; Mee, J.K. Total ionizing dose radiation testing of NVIDIA Jetson nano GPUs. In Proceedings of the 2020 IEEE High Performance Extreme Computing Conference (HPEC), Waltham, MA, USA, 21–25 September 2020; pp. 1–3.
56. Rad, I.O.; Alarcia, R.M.G.; Dengler, S.; Golkar, A.; Manfletti, C. Preliminary Evaluation of Commercial Off-The-Shelf GPUs for Machine Learning Applications in Space. Master's Thesis, Technical University of Munich, Munich, Germany, 2023.
57. Del Castillo, M.O.; Morgan, J.; Mcrobbie, J.; Therakam, C.; Joukhadar, Z.; Mearns, R.; Barraclough, S.; Sinnott, R.; Woods, A.; Bayliss, C.; et al. Mitigating Challenges of the Space Environment for Onboard Artificial Intelligence: Design Overview of the Imaging Payload on SpIRIT. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Seattle, WA, USA, 17–21 June 2024; pp. 6789–6798.
58. Giuffrida, G.; Diana, L.; de Gioia, F.; Benelli, G.; Meoni, G.; Donati, M.; Fanucci, L. CloudScout: A deep neural network for on-board cloud detection on hyperspectral images. *Remote Sens.* **2020**, *12*, 2205. [CrossRef]
59. Dunkel, E.; Swope, J.; Towfic, Z.; Chien, S.; Russell, D.; Sauvageau, J.; Sheldon, D.; Romero-Cañas, J.; Espinosa-Aranda, J.L.; Buckley, L.; et al. Benchmarking deep learning inference of remote sensing imagery on the qualcomm snapdragon and intel movidius myriad x processors onboard the international space station. In Proceedings of the IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 5301–5304.
60. Furano, G.; Meoni, G.; Dunne, A.; Moloney, D.; Ferlet-Cavrois, V.; Tavoularis, A.; Byrne, J.; Buckley, L.; Psarakis, M.; Voss, K.O.; et al. Towards the Use of Artificial Intelligence on the Edge in Space Systems: Challenges and Opportunities. *IEEE Aerosp. Electron. Syst. Mag.* **2020**, *35*, 44–56. [CrossRef]
61. AMD. Virtex-5QV Family Data Sheet. Available online: https://docs.amd.com/v/u/en-US/ds192_V5QV_Device_Overview (accessed on 20 October 2024).
62. Microchip. ProASIC 3 FPGAs. Available online: https://www.microchip.com/en-us/products/fpgas-and-plds/fpgas/proasic-3-fpgas (accessed on 20 October 2024).
63. Bosser, A.; Kohler, P.; Salles, J.; Foucher, M.; Bezine, J.; Perrot, N.; Wang, P.X. Review of TID Effects Reported in ProASIC3 and ProASIC3L FPGAs for 3D PLUS Camera Heads. In Proceedings of the 2023 IEEE Radiation Effects Data Workshop (REDW) (in conjunction with 2023 NSREC), Kansas City, MI, USA, 24–28 July 2023; pp. 1–6.
64. Microchip. RTG4 Radiation-Tolerant FPGAs. Available online: https://www.microchip.com/en-us/products/fpgas-and-plds/radiation-tolerant-fpgas/rtg4-radiation-tolerant-fpgas (accessed on 20 October 2024).

65. Berg, M.D.; Kim, H.; Phan, A.; Seidleck, C.; Label, K.; Pellish, J.; Campola, M. *Microsemi RTG4 Rev C Field Programmable Gate Array Single Event Effects (SEE) Heavy-Ion Test Report*; Technical Report; 2019. Available online: https://ntrs.nasa.gov/citations/20190001593 (accessed on 20 October 2024).
66. Tambara, L.A.; Andersson, J.; Sturesson, F.; Jalle, J.; Sharp, R. Dynamic Heavy Ion SEE Testing of Microsemi RTG4 Flash-based FPGA Embedding a LEON4FT-based SoC. In Proceedings of the 2018 18th European Conference on Radiation and Its Effects on Components and Systems (RADECS), Gothenburg, Sweden, 16–21 September 2018; pp. 1–6.
67. Kim, H.; Park, J.; Lee, H.; Won, D.; Han, M. An FPGA-Accelerated CNN with Parallelized Sum Pooling for Onboard Realtime Routing in Dynamic Low-Orbit Satellite Networks. *Electronics* **2024**, *13*, 2280. [CrossRef]
68. Rapuano, E.; Meoni, G.; Pacini, T.; Dinelli, G.; Furano, G.; Giuffrida, G.; Fanucci, L. An fpga-based hardware accelerator for cnns inference on board satellites: Benchmarking with myriad 2-based solution for the cloudscout case study. *Remote Sens.* **2021**, *13*, 1518. [CrossRef]
69. Pitonak, R.; Mucha, J.; Dobis, L.; Javorka, M.; Marusin, M. Cloudsatnet-1: Fpga-based hardware-accelerated quantized cnn for satellite on-board cloud coverage classification. *Remote Sens.* **2022**, *14*, 3180. [CrossRef]
70. Nannipieri, P.; Giuffrida, G.; Diana, L.; Panicacci, S.; Zulberti, L.; Fanucci, L.; Hernandez, H.G.M.; Hubner, M. Icu4sat: A general-purpose reconfigurable instrument control unit based on open source components. In Proceedings of the 2022 IEEE Aerospace Conference (AERO), Big Sky, MT, USA, 5–12 March 2022.
71. Cosmas, K.; Kenichi, A. Utilization of FPGA for onboard inference of landmark localization in CNN-based spacecraft pose estimation. *Aerospace* **2020**, *7*, 159. [CrossRef]
72. Intel. Next-Level Neuromorphic Computing: Intel Lab's Loihi 2 Chip. Available online: https://www.intel.com/content/www/us/en/research/neuromorphic-computing-loihi-2-technology-brief.html (accessed on 20 October 2024).
73. Davies, M.; Srinivasa, N.; Lin, T.H.; Chinya, G.; Cao, Y.; Choday, S.H.; Dimou, G.; Joshi, P.; Imam, N.; Jain, S.; et al. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro* **2018**, *38*, 82–99. [CrossRef]
74. Ingeniars. GPU@SAT. Available online: https://www.ingeniars.com/in_product/gpusat/ (accessed on 20 October 2024).
75. Benelli, G.; Todaro, G.; Monopoli, M.; Giuffrida, G.; Donati, M.; Fanucci, L. GPU@ SAT DevKit: Empowering Edge Computing Development Onboard Satellites in the Space-IoT Era. *Electronics* **2024**, *13*, 3928. [CrossRef]
76. Benelli, G.; Giuffrida, G.; Ciardi, R.; Davalle, D.; Todaro, G.; Fanucci, L. GPU@ SAT, the AI enabling ecosystem for on-board satellite applications. In Proceedings of the 2023 European Data Handling & Data Processing Conference (EDHPC), Juan-Les-Pins, France, 2–6 October 2023; pp. 1–4.
77. MHI. AIRIS. Available online: https://www.mhi.com/news/240306.html (accessed on 20 October 2024).
78. Communications, B.M. BMC Products. Available online: https://www.bluemarblecomms.com/products/ (accessed on 20 October 2024).
79. Blue Marble Communications (BMC); BruhnBruhn Innovation (BBI). Space Edge Processor and Dacreo AI Ecosystem. Available online: https://bruhnbruhn.com/wp-content/uploads/2024/03/SAT2024-SEP-Apps-Demo-Flyer.pdf (accessed on 20 October 2024).
80. Blue Marble Communications (BMC); BruhnBruhn Innovation (BBI). dacreo: Space AI Cloud Computing. Available online: https://bruhnbruhn.com/dacreo-space-ai-cloud-computing/ (accessed on 20 October 2024).
81. AIKO. AIKO Onboard Data Processing Suite. Available online: https://aikospace.com/projects/aiko-onboard-data-processing-suite/ (accessed on 20 October 2024).
82. Dini, P.; Diana, L.; Elhanashi, A.; Saponara, S. Overview of AI-Models and Tools in Embedded IIoT Applications. *Electronics* **2024**, *13*, 2322. [CrossRef]
83. Furano, G.; Tavoularis, A.; Rovatti, M. AI in space: Applications examples and challenges. In Proceedings of the 2020 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT), Frascati, Italy, 19–21 October 2020; pp. 1–6. [CrossRef]
84. Dini, P.; Elhanashi, A.; Begni, A.; Saponara, S.; Zheng, Q.; Gasmi, K. Overview on intrusion detection systems design exploiting machine learning for networking cybersecurity. *Appl. Sci.* **2023**, *13*, 7507. [CrossRef]
85. Dini, P.; Saponara, S. Analysis, design, and comparison of machine-learning techniques for networking intrusion detection. *Designs* **2021**, *5*, 9. [CrossRef]
86. Wei, L.; Ma, Z.; Yang, C.; Yao, Q. Advances in the Neural Network Quantization: A Comprehensive Review. *Appl. Sci.* **2024**, *14*, 7445. [CrossRef]
87. Dantas, P.V.; Sabino da Silva, W., Jr.; Cordeiro, L.C.; Carvalho, C.B. A comprehensive review of model compression techniques in machine learning. *Appl. Intell.* **2024**, *54*, 11804–11844. [CrossRef]
88. Deng, L.; Li, G.; Han, S.; Shi, L.; Xie, Y. Model compression and hardware acceleration for neural networks: A comprehensive survey. *Proc. IEEE* **2020**, *108*, 485–532. [CrossRef]
89. Ekelund, J.; Vinuesa, R.; Khotyaintsev, Y.; Henri, P.; Delzanno, G.L.; Markidis, S. AI in Space for Scientific Missions: Strategies for Minimizing Neural-Network Model Upload. *arXiv* **2024**, arXiv:2406.14297.
90. Olshevsky, V.; Khotyaintsev, Y.V.; Lalti, A.; Divin, A.; Delzanno, G.L.; Anderzén, S.; Herman, P.; Chien, S.W.; Avanov, L.; Dimmock, A.P.; et al. Automated classification of plasma regions using 3D particle energy distributions. *J. Geophys. Res. Space Phys.* **2021**, *126*, e2021JA029620. [CrossRef]

91. Guerrisi, G.; Del Frate, F.; Schiavon, G. Satellite on-board change detection via auto-associative neural networks. *Remote Sens.* **2022**, *14*, 2735. [CrossRef]
92. Ziaja, M.; Bosowski, P.; Myller, M.; Gajoch, G.; Gumiela, M.; Protich, J.; Borda, K.; Jayaraman, D.; Dividino, R.; Nalepa, J. Benchmarking deep learning for on-board space applications. *Remote Sens.* **2021**, *13*, 3981. [CrossRef]
93. Ghassemi, S.; Magli, E. Convolutional neural networks for on-board cloud screening. *Remote Sens.* **2019**, *11*, 1417. [CrossRef]
94. Hughes, M.J.; Hayes, D.J. Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing. *Remote Sens.* **2014**, *6*, 4907–4926. [CrossRef]
95. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [CrossRef] [PubMed]
96. Lagunas, E.; Ortiz, F.; Eappen, G.; Daoud, S.; Martins, W.A.; Querol, J.; Chatzinotas, S.; Skatchkovsky, N.; Rajendran, B.; Simeone, O. Performance Evaluation of Neuromorphic Hardware for Onboard Satellite Communication Applications. *arXiv* **2024**, arXiv:2401.06911.
97. Orchard, G.; Frady, E.P.; Rubin, D.B.D.; Sanborn, S.; Shrestha, S.B.; Sommer, F.T.; Davies, M. Efficient neuromorphic signal processing with loihi 2. In Proceedings of the 2021 IEEE Workshop on Signal Processing Systems (SiPS), Coimbra, Portugal, 20–22 October 2021; pp. 254–259.
98. Intel. Intel Advances Neuromorphic with Loihi 2, New Lava Software Framework and New Partners. Available online: https://www.intel.com/content/www/us/en/newsroom/news/intel-unveils-neuromorphic-loihi-2-lava-software.html#gs.ezemn0 (accessed on 20 October 2024).
99. LAVA. Lava Software Framework. Available online: https://lava-nc.org/ (accessed on 20 October 2024).
100. Ieracitano, C.; Mammone, N.; Spagnolo, F.; Frustaci, F.; Perri, S.; Corsonello, P.; Morabito, F.C. An explainable embedded neural system for on-board ship detection from optical satellite imagery. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108517. [CrossRef]
101. Giuffrida, G.; Fanucci, L.; Meoni, G.; Batič, M.; Buckley, L.; Dunne, A.; van Dijk, C.; Esposito, M.; Hefele, J.; Vercruyssen, N.; et al. The Φ-Sat-1 Mission: The First On-Board Deep Neural Network Demonstrator for Satellite Earth Observation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [CrossRef]
102. Cucchetti, E.; Latry, C.; Blanchet, G.; Delvit, J.M.; Bruno, M. Onboard/on-ground image processing chain for high-resolution Earth observation satellites. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2021**, *43*, 755–762. [CrossRef]
103. Chintalapati, B.; Precht, A.; Hanra, S.; Laufer, R.; Liwicki, M.; Eickhoff, J. Opportunities and challenges of on-board AI-based image recognition for small satellite Earth observation missions. *Adv. Space Res.* 2024, *in press*. [CrossRef]
104. de VIEILLEVILLE, F.; Lagrange, A.; Ruiloba, R.; May, S. Towards distillation of deep neural networks for satellite on-board image segmentation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 1553–1559. [CrossRef]
105. Goudemant, T.; Francesconi, B.; Aubrun, M.; Kervennic, E.; Grenet, I.; Bobichon, Y.; Bellizzi, M. Onboard Anomaly Detection for Marine Environmental Protection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 7918–7931. [CrossRef]
106. Begni, A.; Dini, P.; Saponara, S. Design and test of an lstm-based algorithm for li-ion batteries remaining useful life estimation. In Proceedings of the International Conference on Applications in Electronics Pervading Industry, Environment and Society, Genoa, Italy, 26–27 September 2022; Springer: Cham, Switzerland, 2022; pp. 373–379.
107. Dini, P.; Ariaudo, G.; Botto, G.; Greca, F.L.; Saponara, S. Real-time electro-thermal modelling and predictive control design of resonant power converter in full electric vehicle applications. *IET Power Electron.* **2023**, *16*, 2045–2064. [CrossRef]
108. Murphy, J.; Ward, J.E.; Mac Namee, B. An Overview of Machine Learning Techniques for Onboard Anomaly Detection in Satellite Telemetry. In Proceedings of the 2023 European Data Handling & Data Processing Conference (EDHPC), Juan-Les-Pins, France, 2–6 October 2023; pp. 1–6. [CrossRef]
109. Muthusamy, V.; Kumar, K.D. Failure prognosis and remaining useful life prediction of control moment gyroscopes onboard satellites. *Adv. Space Res.* **2022**, *69*, 718–726. [CrossRef]
110. Salazar, C.; Gonzalez-Llorente, J.; Cardenas, L.; Mendez, J.; Rincon, S.; Rodriguez-Ferreira, J.; Acero, I.F. Cloud detection autonomous system based on machine learning and cots components on-board small satellites. *Remote Sens.* **2022**, *14*, 5597. [CrossRef]
111. Murphy, J.; Ward, J.E.; Namee, B.M. Low-power boards enabling ml-based approaches to fdir in space-based applications. In Proceedings of the 35th Annual Small Satellite Conference, Salt Lake City, UT, USA, 6–11 August 2021.
112. Pacini, F.; Dini, P.; Fanucci, L. Cooperative Driver Assistance for Electric Wheelchair. In Proceedings of the International Conference on Applications in Electronics Pervading Industry, Environment and Society, Genoa, Italy, 28–29 September 2023; Springer: Cham, Switzerland, 2023; pp. 109–116.
113. Pacini, F.; Dini, P.; Fanucci, L. Design of an Assisted Driving System for Obstacle Avoidance Based on Reinforcement Learning Applied to Electrified Wheelchairs. *Electronics* **2024**, *13*, 1507. [CrossRef]
114. Hao, Z.; Shyam, R.A.; Rathinam, A.; Gao, Y. Intelligent spacecraft visual GNC architecture with the state-of-the-art AI components for on-orbit manipulation. *Front. Robot. AI* **2021**, *8*, 639327. [CrossRef]
115. Buonagura, C.; Pugliatti, M.; Franzese, V.; Topputo, F.; Zeqaj, A.; Zannoni, M.; Varile, M.; Bloise, I.; Fontana, F.; Rossi, F.; et al. Deep Learning for Navigation of Small Satellites About Asteroids: An Introduction to the DeepNav Project. In Proceedings of the International Conference on Applied Intelligence and Informatics, Reggio Calabria, Italy, 1–3 September 2022; Springer: Cham, Switzerland, 2022; pp. 259–271.

116. Buonagura, C.; Borgia, S.; Pugliatti, M.; Morselli, A.; Topputo, F.; Corradino, F.; Visconti, P.; Deva, L.; Fedele, A.; Leccese, G.; et al. The CubeSat Mission FUTURE: A Preliminary Analysis to Validate the On-Board Autonomous Orbit Determination. In Proceedings of the 12th International Conference on Guidance, Navigation & Control Systems (GNC) and 9th International Conference on Astrodynamics Tools and Techniques (ICATT), Sopot, Poland, 12–16 June 2023; pp. 1–15.
117. Fourati, F.; Alouini, M.S. Artificial intelligence for satellite communication: A review. *Intell. Converg. Netw.* **2021**, *2*, 213–243. [CrossRef]
118. Gómez, P.; Meoni, G. Tackling the Satellite Downlink Bottleneck with Federated Onboard Learning of Image Compression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 17–21 June 2024; pp. 6809–6818.
119. Guerrisi, G.; Bencivenni, G.; Schiavon, G.; Del Frate, F. On-Board Multispectral Image Compression with an Artificial Intelligence Based Algorithm. In Proceedings of the IGARSS 2024—2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 7–12 July 2024; pp. 2555–2559. [CrossRef]
120. Garimella, S. Onboard deep learning for efficient small satellite reflectance retrievals and downlink. In Proceedings of the Image and Signal Processing for Remote Sensing XXIX, Amsterdam, The Netherlands, 4–5 September 2023; Volume 12733, pp. 20–23.
121. Navarro, T.; Dinis, D.D.C. Future Trends in AI for Satellite Communications. In Proceedings of the 2024 9th International Conference on Machine Learning Technologies, Oslo, Norway, 24–26 May 2024; pp. 64–73.