Seeing Through the Haze: Simulation-Driven Atmospheric Analysis and Onboard Dehazing for Robust LEO/VLEO Earth Observation

Bharadwaj Chintalapati[1,2*], Arthur Precht[3], Jeffrey Hendrikse[4], Marcus Liwicki[1], Rene Laufer[1], Jens Eickhoff[2,5]

[1] Luleå University of Technology, Luleå, Sweden
[2] Airbus Defence and Space GmbH, Friedrichshafen, Immenstaad am Bodensee, Germany
[3] Ferschau, 80687 Munich, Germany
[4] Atmos Space Cargo GmbH, Lichtenau, Germany
[5] University of Stuttgart, Stuttgart, Germany
*Corresponding Author: lakshmi.venkata.bharadwaj.chintalapati@associated.ltu.se

Abstract
Atmospheric attenuation and cloud cover remain a primary bottleneck for optical Earth Observation (EO) mission operations, often rendering downlinked data unusable and straining limited bandwidth budgets. This work presents a novel framework to improve actionable intelligence by integrating real-time atmospheric assessment and image restoration directly into the onboard image processing chain. To overcome the scarcity of representative space-based training data, a custom simulation pipeline was developed to synthesize optical imagery that accounts for sensor noise, geometric distortions, and physically based atmospheric effects. The generated dataset was evaluated using the Structural Similarity Index Measure (SSIM) and employed to train lightweight convolutional neural networks for near real-time scene classification, along with an attention-based network for onboard image dehazing. Experimental benchmarking was performed on flight-representative computing architectures, including the Intel E3900 Atom processor and NVIDIA Jetson Orin Nano. The classification network attained an accuracy of 98.7%, while the dehazing algorithm exhibited high reconstruction fidelity with a patch-level SSIM of 0.91 and a Peak Signal-to-Noise Ratio of 19.69 dB - representing a 5.48 dB enhancement over raw atmospheric inputs. With a computational complexity of only 5.47 GMacs, the proposed model remains suitable for power-constrained edge processing environments. The framework improves data utility and bandwidth efficiency by filtering degraded images and improving scene interpretability prior to downlink, directly aiding autonomous decision-making in upcoming EO missions. The architecture is planned for on-orbit verification in the upcoming Atmos Space Cargo Phoenix 2.2 technology demonstration mission.

Keywords: Atmospheric Correction, VLEO, Scene Classification, Image Dehazing, Deep Learning, Edge Deployment.

1. Introduction

Recent Earth Observation (EO) architectures are moving away from monolithic single-satellite designs towards distributed constellations of Low Earth Orbit (LEO) SmallSats and Very Low Earth Orbit (VLEO) CubeSats [1]. Prior research has demonstrated the integration of edge high-computing hardware such as NVIDIA Jetson Orin Nano directly into the payload subsystem to advance autonomous multi-class object detection [2]. By deploying models (e.g. YOLOv11) onboard, it was seen that transmitting contextual metadata rather than raw imagery could reduce downlink requirements by over 90%, preserving critical bandwidth while maintaining high detection precision [2]. The hardware topology of this paradigm is illustrated in Fig. 1a and Fig. 1b.

However, the efficacy of onboard object detection systems is heavily predicated on the assumption of clear, high-contrast imagery as seen in Fig. 1(e). The space environment introduces severe radiometric and atmospheric variables that critically degrade satellite imagery before it ever reaches the model pipeline [3]. Although LEO/VLEO constellations provide unprecedented spatial resolution, they tend to generate Terabytes of raw optical data daily. Transmitting this volume of data to ground stations is severely constrained by brief satellite-to-ground visibility windows and the limited bandwidth of standard X-band communication links [4].

This downlink bottleneck is intensified by the *Atmospheric Barrier* as illustrated in Fig. 1(c) and Fig. 1(d). Climatological studies indicate that, at any given time, over 67% of the Earth's surface is obscured by cloud cover, aerosols, or severe haze [5, 6]. When a satellite blindly downlinks optical satellite raw data, a significant portion of the telemetry budget is wasted on transmitting heavily attenuated, unusable images. If such hazy and cloudy imagery is fed as input to an onboard object detection model, the confidence threshold drops, resulting in critical false negatives as seen in Fig. 1(f). This failure forces the system to downlink the raw images for human-in-the-loop verification on ground, entirely negating the benefits of edge computing and increasing the latency for *Actionable Intelligence* in time-sensitive defense and disaster monitoring applications. Current state-of-the-art EO pipelines rely predominantly on ground processing for atmospheric correction [7].

Addressing this inefficiency requires shifting data triage and partial atmospheric restoration from the ground segment directly to the space segment. However, transitioning atmospheric correction to the

edge exposes three critical gaps in current Deep Learning (DL) methodologies:

- Raw Data Scarcity: Existing DL models for image restoration and classification are predominantly trained on Level-1C or Level-2A data products [8]. These datasets have already undergone image correction and demosaicing. Training onboard models on ground-processed data induces domain shift, rendering the algorithms brittle when deployed on raw Level-0 data in-orbit.
- Hardware Constraints and Algorithmic Complexity: Image restoration (for example dehazing, restoration) is an ill-posed inverse problems traditionally solved using highly parameterized Convolutional Neural Networks (CNNs) [9] or heavy Vision Transformers (ViTs) [10]. Such architectures exceed the Size, Weight, and Power (SWaP) constraints on simple processors or power-capped edge GPUs.
- Lack of Integrated Onboard Pipeline: There is a distinct lack of hybrid frameworks that couple near real-time scene classification (to filter opaque clouds) with localized, computationally efficient image restoration prior to the downstream object detection stage [11].

To bridge the gap between ideal simulated deployments and the harsh realities of raw in-orbit operations, this research proposes a physics-based simulation of raw sensor data to address the raw data scarcity problem. Furthermore, this research proposes a framework combining lightweight classification and attention-based dehazing directly within the onboard processing chain. This study addresses the following primary question: *How does the onboard integration of a low complexity scene classification and attention-based dehazing algorithm impact the actionable yield of EO imagery under strict SWaP constraints?* To further evaluate the framework, the following hypotheses are established:

- *Primary Hypothesis*: If a hierarchical onboard pipeline (scene classification acting as a computational gatekeeper, followed by attention-based dehazing) is applied directly to existing or simulated raw optical satellite imagery on computationally constrained edge devices with less than 10 GMacs (Giga Multiply-Accumulate operations), then the structural fidelity and the downstream object detection yield will demonstrate an improvement over uncorrected, unfiltered raw pipelines.
- *Baseline Assumption:* Monolithic processing of all captured imagery (without triage) on edge hardware will exceed nominal Size, Weight, and Power (SWaP) constraints, yielding no operational advantage over traditional ground-based processing.

To test these hypotheses, the main contributions of this paper are summarized as follows:

(1) *Physics-Based Simulation Pipeline*. Generation of raw optical imagery that accounts for sensor noise, Modulation Transfer Function (MTF) degradation, geometric distortions, and atmospheric scattering, tailored to specific Commercial-Off-The-Shelf (COTS) sensor specifications.
(2) *Hierarchical Dual-Network Architecture*. The development of integrated onboard processing chain featuring a lightweight Convolutional Neural Network (CNN) for near real-time scene classification (opaque cloud filtering), coupled with an attention-based dehazing network only on recoverable hazy imagery to improve downstream multi-class object detection.
(3) *Edge-Hardware Deployment*. The empirical validation and deployment of the proposed models on flight-representative edge computing architectures.

2. Related Work

To contextualize the proposed onboard classification and restoration framework, this section assesses the EO architecture, computational trade-offs of atmospheric correction, and critical gap in training data.

2.1. Earth Observation Data Pipelines

Traditional EO missions employ *store-and-forward* architecture (Fig. 3), where geometrically and radiometrically degraded Level-0 sensor data is captured [3]. Computationally intensive corrections are then performed on-ground using high-performance clusters to generate Level-1C or Level-2A data products [13]. Conversely, Orbital Edge Computing (OEC) aims to shift these workloads to the satellite payload. For instance, ESA's Phi-sat-1 demonstrated onboard cloud screening [11], and commercial operators like Planet Labs are transitioning from ground-based filters [16], to on-board GPU data processing [17].

2.2. Raw Data Scarcity and Cloud Paradox

Most optical image processing and deep learning frameworks rely on highly processed public datasets.

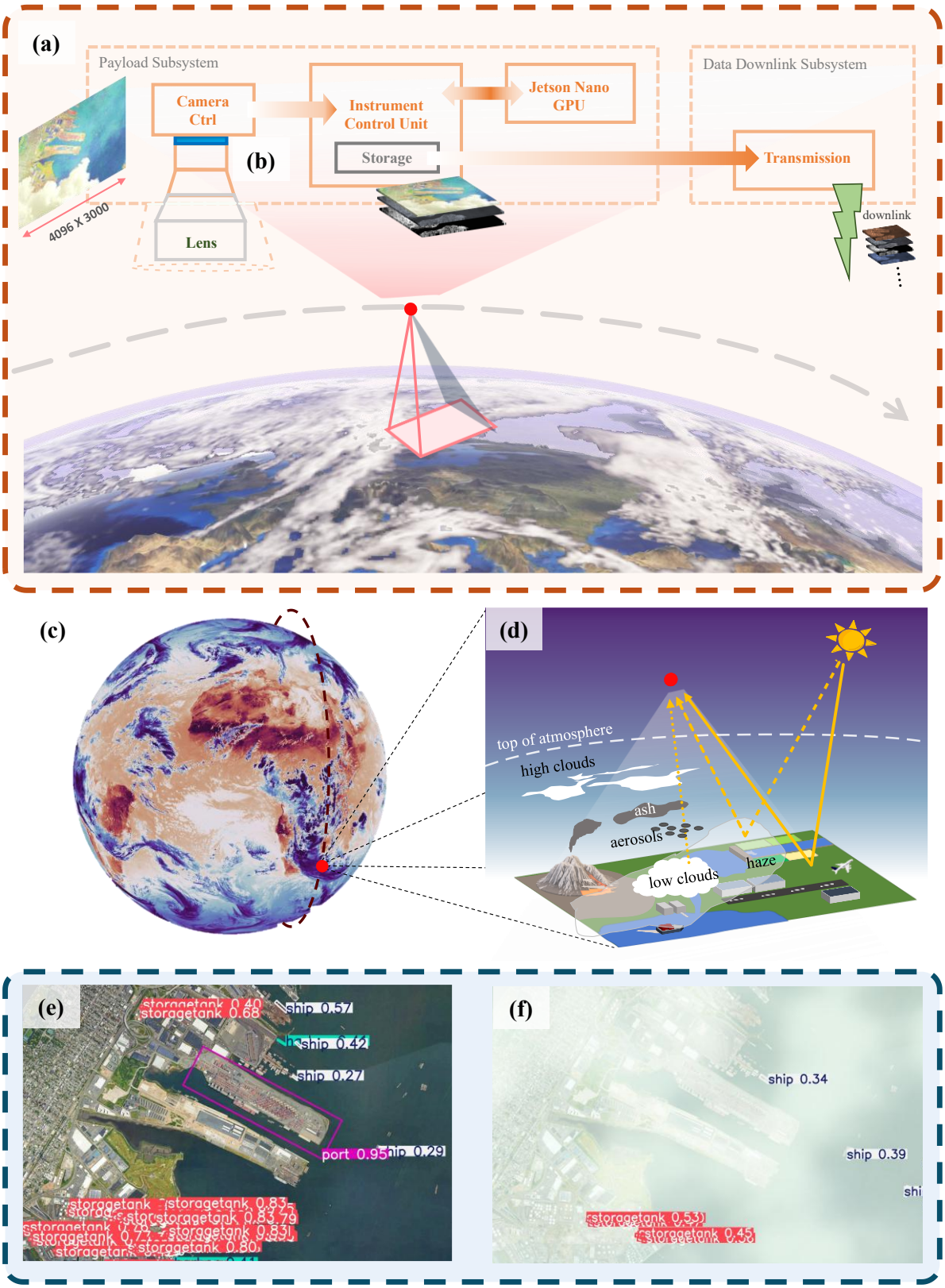

Fig. 1 Overview of the proposed edge-computing Earth Observation architecture and the impact of the Atmospheric Barrier. (**a**) Conceptual visualization of IceBrain-1 mission [2]. (**b**) Onboard Payload Subsystem featuring edge AI integration directly interfacing with the optical sensor. (**c**) Global cloud cover illustrating the prevalence of atmospheric obscuration [12]. (**d**) Physics of atmospheric interference, detailing scattering and attenuation from clouds, aerosols, and haze. (**e**) Object detection inference on clear optical imagery[2]. (**f**) Degraded object detection performance and false negatives under hazy, uncorrected conditions

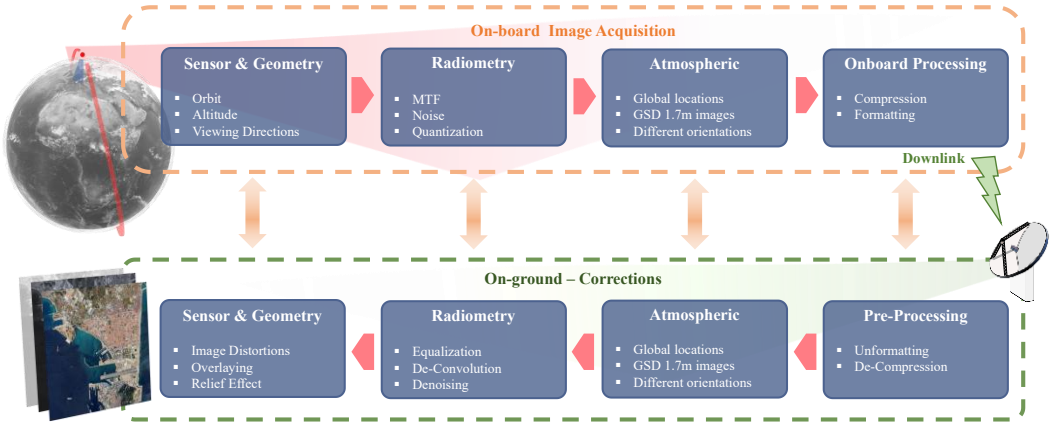

Fig. 2 Traditional architecture of Earth Observation data pipeline (adapted from [13])

Some of the examples are EuroSAT, SpaceNet, DOTA [8, 18]. Therefore, creating a distinct void in publicly available raw Level-0 imagery. This scarcity is exacerbated by the *Cloud Paradox* (Table 1). Satellite operates aggressively discard unusable cloudy or hazy images on the ground. Curated datasets of realistic cloudy-to-clear paired satellite images are virtually non-existent. Furthermore, releasing raw data exposes proprietary intellectual property regarding sensor calibration and defects.

2.3. Scene Classification for Data Triage

The data/telemetry wastage highlighted in Table 1 can be mitigated via real-time data triage. While traditional ground-based detection relies on multi-spectral thresholding [19], strict SWaP constraints have popularized lightweight Convolutional Neural Networks (CNNs) for binary classification directly on the edge.

Table 1 Exemplary public raw Earth Observation (EO) data sources with average scene category estimates and cloud paradox [5, 7, 14, 15]. *SAR sensor is cloud penetrating)

Sensor Type	Resolution	Avg. Raw Scenes (Global)	Avg. Clear Scenes (<20% CC) /Year	Paradox Ratio (Raw vs Clear)
Optical (Sentinel-2)	10-60m	~3.5 million	~1.05 million	~3:1
Multispectral (Landsat8/9)	15-100m	~547,500	~191,000	~3:1
SAR (Sentinel-1)	5-20m	~1.7 million	N/A*	1:1

However, existing onboard models generally act as independent, rigid gatekeepers [20], classifying images as either usable or cloudy and discarding the rest. This fails to account for the nuance of the atmospheric barrier, leaving a critical gap for hierarchical pipelines that can actively route recoverable, hazy imagery to an intermediate restoration module.

2.4. Constraints for Image Restoration

Traditional atmospheric correction relies on physical Radiative Transfer Models (RTMs) [21] or statistical priors like the Dark Current Prior (DCP) [22]. Both are suboptimal for isolated, real-time edge processing due to unavailable auxiliary atmospheric data, color distortion, and computational expense [22]. In recent years, deep learning methods like AOD-Net [9] and GridDehazeNet [23] have achieved superior reconstruction. However, their massive computational footprints (e.g., >20 GMacs) drastically exceed the thermal and power envelopes of targeted edge hardware, necessitating an optimized approach.

3. Raw Optical Imagery Simulation

To address the *Raw Data Scarcity* paradox, a physics-based simulation pipeline was developed to synthesize Level-0 data [24]. It systematically degrades pristine optical imagery via cascading layers (orbital geometrics, atmospheric scattering, opto-electronic noise). The simulation is mainly tailored to a high-resolution COTS payload: Sony IMX565 sensor (12.3 MP, 2.74μm pixel pitch) with an 800mm (f/8) lens [25].

3.1. The Simulation Framework

The pipeline reverse-engineers the physical journey of light (Fig. 2). Stage 1 generates a *Clear Raw Optical Imagery* dataset (Fig. 3b) by injecting sensor-specific characteristics to improve onboard model robustness [2]. Stage 2 introduces dynamic, overlapping atmospheric occlusions (Fig. 3a-3c, 3e). The entire process is governed by a *Forward Observation Model* as seen in the Eq. (1),.

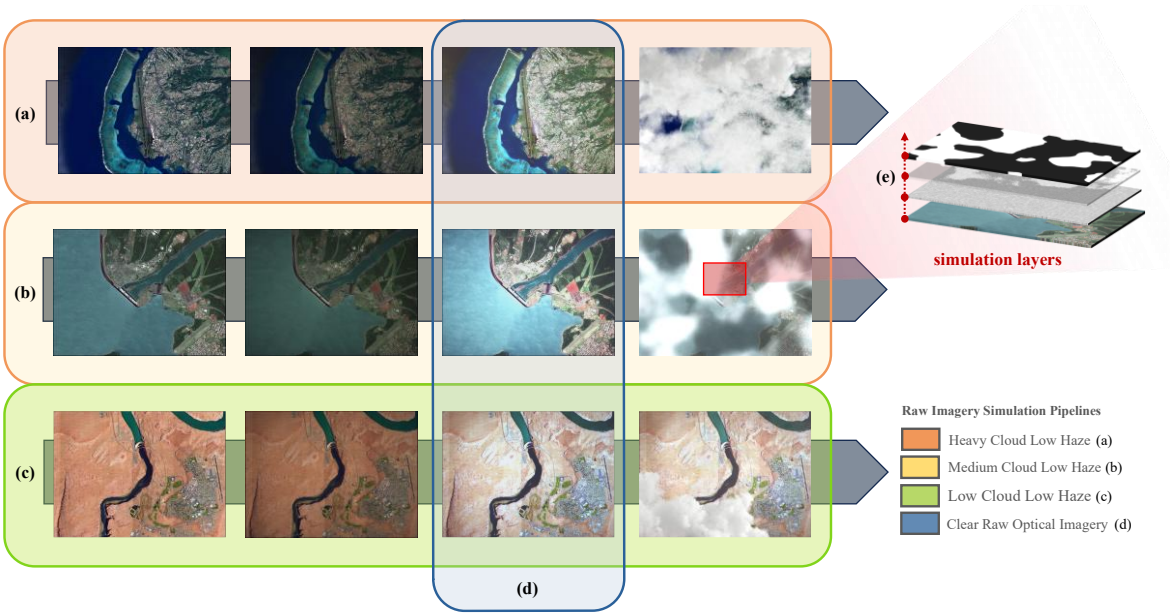

Fig. 3 Raw Imagery Simulation Pipeline. A framework degrading clear optical imagery to heavy cloud **(a)**, medium cloud **(b)**, and low cloud **(c)** atmospheric occlusions. **(d)** clear raw optical imagery is generated in first stage. **(e)** atmospheric occlusions as structured layers

$$I_{DN}(x,y) = Q\left\{PRNU \cdot \left[\begin{matrix}\left(L_{TOA}\left(T_{geo}(x,y)\right) * PSF_{sys}\right) \\ \cdot V(x,y) \cdot g \cdot t_{exp} \\ + N_{dark}\end{matrix}\right]\right\} \quad (1)$$

This equation converts Top-of-Atmosphere (TOA) radiance (L_{TOA}) is converted to discrete digital number (I_{DN}) [26]. Here, Q is quantization, Pixel Response Non-Uniformity ($PRNU$), T_{geo} is geometric transformation, PSF_{sys} is the Point Spread Function of the system, V is radial vignetting, g is sensor gain, and N_{dark} is thermal noise. Simulating exposure times (t_{exp}) and SNR dynamically reflect real-world LEO illumination windows derived from orbital ephemeris data.

3.2. Atmospheric Modeling

The atmospheric attenuation of the surface signal (Bottom-of-Atmosphere radiance, L_{BOA}) is modeled using a parameterized radiative transfer approach as detailed in Eq. (2) .

$$L_{TOA}(\lambda) = L_{path}(\lambda) + L_{BOA}(\lambda) \cdot \exp\left(-\frac{\tau_{550}\left(\frac{\lambda}{550}\right)^{-\alpha}}{cos\theta_z}\right) \quad (2)$$

As visualized in Fig. 4, τ_{550} is the Aerosol Optical Depth (AOD), α is the Ångström exponent, and θ_z is the solar zenith angle [21]. This accurately models both Rayleigh (blue band washout) and Mie scattering (haze), generating spectrally authentic occlusions rather than simplistic white overlays.

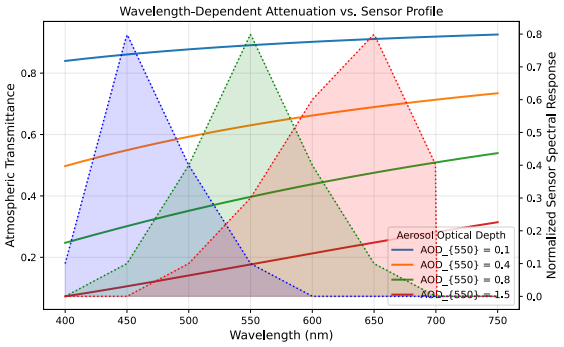

Fig. 4 Atmospheric Degradation Profiles

3.3. Geometric Distortions

The $T_{geo}(x,y)$ function applies to orbital spatial deformations. Considering the baseline IceBrain-1 mission's 500km nadir LEO orbit as reference. The Ground Sample Distance (GSD) is 1.7m. The center of the image is normally at 0°, and the extreme left or right edges of the image are ~0.4° The off-nadir perspective distortion due to the Earth curvature effects [27] is simulated using the Eq. (3).

$$GSD = GSD_{nadir} \cdot \left(\frac{Slant\_Range}{Altitude}\right) \quad (3)$$

Furthermore, COTS radial barrel distortion and exposure-dependent motion smear (~7.5 km/s velocity) are applied.

3.4. Sensor Modeling and MTF Degradation

The model simulates optical blurring, Point Spread Function (PSF_{blur}) and motion smear, captured by the system's Modulation Transfer Function (MTF) as plotted in Fig. 5. Radiometric shifts are then introduced, including vignetting ($V(x,y)$), stray light reflections, thermal dark noise (N_{dark}), and a 0.01% static defective pixel mask to mimic radiation degradation.

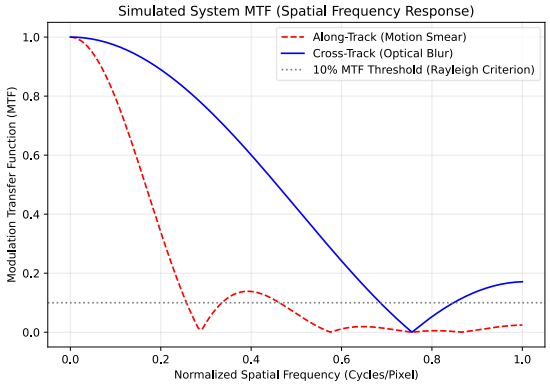

Fig. 5 Simulated System Modulation Transfer Function (MTF) and the grey dotted line indicating algorithm cut-off.

3.5. Cloud & Haze Simulation

A hybrid dataset strategy was used to simulate diverse atmospheric densities. Clouds were procedurally generated using multi-octave Perlin noise mapped to RGBA masks, allowing varied morphologies. A 4-class topographical mask generator [28] was integrated to synthesize crucial cloud shadows. Real-world Sentinel-2 imagery (30-40% occlusion) was also added to enhance meteorological robustness.

For haze, Koschmieder's Law [22] was applied via a gradient descent depth map for vertical fading. Synthetic haze densities AOD (τ_{550}), with values of 0.65 and 0.85 were combined with the Perlin clouds and empirical overlays to ensure the dataset encompasses a manifold of atmospheric perturbations.

3.6. Scene classification & Haze Dataset

The pipeline generated an *11-class, 38,970*-image classification dataset with 224x224 and 512x512 image resolutions with a class distribution of ~2500-3000 images per category. The dataset is augmented with the USTC_SmokeRS_dataset [29]. The whole dataset is categorized into three operational tiers to train the CNN gatekeeper:

Clear_Classes:
(*clear_terrain, clear_snow, clear_water, seaside, sea_ice*): These classes represent raw clear satellite imagery as seen in Fig. 3(d) routed directly to object detection to conserve computational power (GMacs).

***Obscured/Recoverable_Classes*:**
(*medium_haze, heavy_haze, dust, smoke*):Retains the structural info, routed to dehazing.

***Opaque/Unrecoverable_Classes*:** (*heavy_cloud, medium_cloud*). These scenes represent a loss of the Bottom-of-Atmosphere signal. The classifier is to be trained to flag these images.

Explicitly separating *clear_snow* and *heavy_cloud* addresses the common algorithmic pitfall of *cloud-snow* confusion, forming a foundational bedrock required to validate the proposed Orbital Edge Computing architecture.

4. Onboard Neural Network Architecture

To process the simulated raw optical imagery under the strict Size, Weight, and Power (SWaP) constraints of a 16U CubeSat, a hardware-optimized, hierarchical neural network architecture was engineered. This section details the autonomous triage logic, the multi-class gatekeeper network, the attention-based dehazing algorithm, and the downstream object detection algorithm [2].

4.1. Autonomous Triage Logic

The fundamental principle of Orbital Edge Computing is computational economy: executing deep neural networks only when strictly necessary. To achieve this, an overarching Autonomous Triage Logic was implemented, mapping distinct processing tiers to specific hardware components, as illustrated in Fig. 6.

This logic sequence acts as the payload's central decision matrix. Following image acquisition, the data enters the *Gatekeeper* phase. The *Scene Classification* network is the main model that processes the images in this phase. To maximize hardware efficiency, a low-power CPU node was considered as baseline in the IceBrain-1 payload subsystem. The theoretical assumption was to host the scene classification network onboard this CPU node and the power-intensive GPU to remain largely dormant until selectively engaged. Based on the classification result, the image is either immediately discarded (preserving computation and telemetry if *cloudy*) or routed to the GPU for Stage-0 processing. If the classification result is *clear*, the image is directly routed to GPU for object detection.

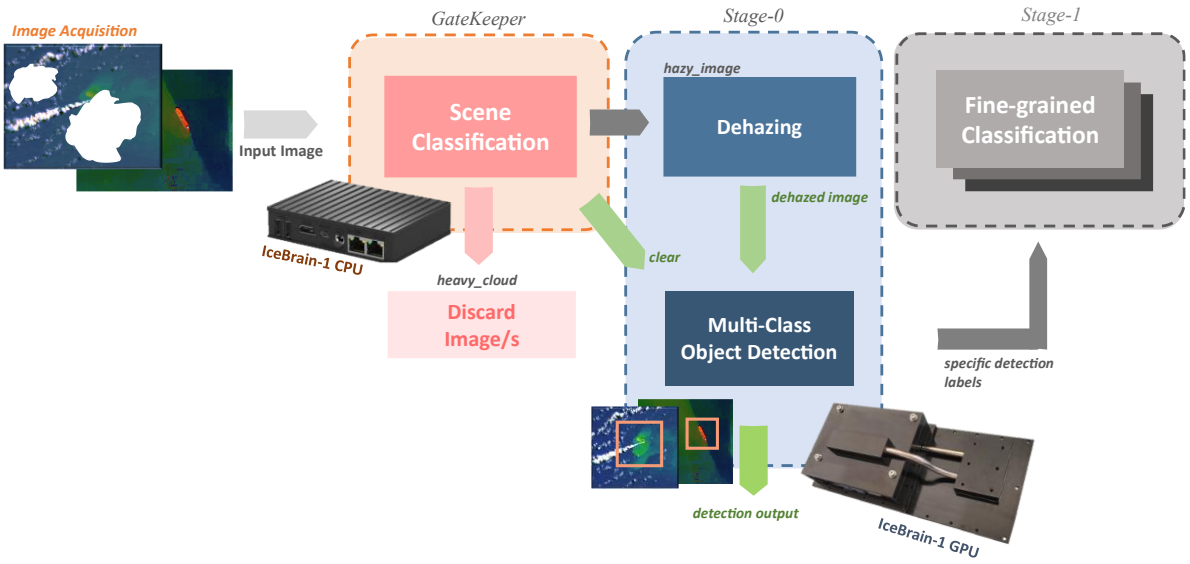

Fig. 6 Onboard Autonomous Triage and Routing

If the classification result is *hazy,* the image is also routed to GPU for image restoration. Once restored, it is passed downstream for object detection. Based on the detection results, the images are further routed for Stage-1 fine-grained image processing (e.g. disaster monitoring). At the time of drafting this paper, a separate study on potential fine-grained use-case is being planned. The intention for this physical separation of triage and its deployment in the IceBrain-1payload was to preserve the telemetry bandwidth and the satellite's power and thermal budget. The contemporary on-ground deep learning trends often favor Multi-Task Learning (MTL) architecture, where a single monolithic model performs classification, restoration, and detection simultaneously [30]. Such an approach is fundamentally incompatible with the strict SWaP constraints of a satellite payload.

4.2. Custom Scene Classification Network

To serve as the triage gatekeeper, a specialized, low-latency Convolutional Neural Network (CNN) was designed. Standard lightweight models struggle with the severe domain overlap inherent to EO data. To resolve this, a Dual-Stream Gated Fusion Network was engineered (*IceBrainSceneNet)*, as detailed in the Fig. 7. The native 4096×3000 raw imagery is first split into manageable 1 x 3 x 224 x 224 input tensors. The architecture then bifurcates into two parallel extraction streams:

- *RGB Features Stream*: Extracts standard spatial and colorimetric features via traditional convolutions C_l.
- *Texture Stream:* A dedicated Texture Extractor processes high-frequency structural data C_p, isolating the jagged topography of terrestrial from the smooth gradients of atmospheric masks.

These streams are dynamically merged via a *Gated Fusion Network* (Fig. 7, blue dashed callout) [31]. The fusion block computes a squeezed spatial attention map, passing it through a convolution and Sigmoid activation ($Conv + \sigma$). These weights scale RGB features before combining with the texture features, defined as: $(c_l \otimes (Conv_\sigma(c_p)) \oplus c_p)$.

This yields a 128-channel fused feature map, passing it through successive Attention Blocks to 256 channels [32], a pooling grid, and a flatten layer. The network outputs are 11 distinct classes as mentioned in the Section 3.6, grouping them into safe-to-process classes (green dashed box), hazy candidates requiring restoration (yellow dashed box), and opaque discards (red dashed box).

4.3. Dual Attention Dehazing Network

When the Gatekeeper flags a scene for restoration, the data is routed to the GPU-hosted *Dual Attention Dehazing* Network (*DA-Net*). This model surpasses the restoration fidelity of legacy convolutions while avoiding the quadratic memory bottleneck of heavy Vision Transformers [33]. As illustrated in the Fig. 8, the 4096×3000 hazy image is split into $1 \times 3 \times 512 \times 512$ spatial crops to preserve fine-grained structural details. This fine-tuned DA-Net architecture utilizes a highly optimized U-Net topology [34]:

- *Encoder*: Alternating 3 x 3 Convolutions and Dual Attention Blocks [35], down-sample the spatial resolution context of the haze.

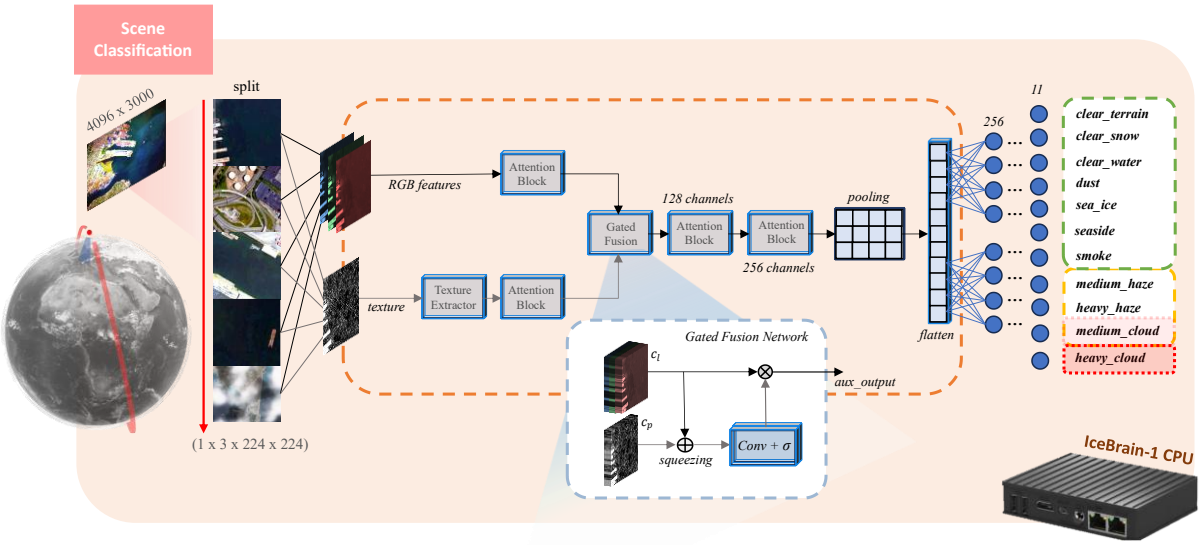

Fig. 7 Scene Classification (*IceBrainSceneNet)* Architecture

- *Decoder*: Utilizes Up-Sample mechanisms to dynamically recalibrate receptive fields, fusing multi-scale skip connections from the encoder.
- *Reconstruction & Blend*: The decoder outputs a restored tensor that undergoes a final *merge & blend* operation to seamlessly stitch the 512 x 512 patches back into a contiguous clear image.

Attention as a Sparsity Filter

The internal routing of the Dual Attention Block is modified in this model (Fig. 8, bottom callout). Following Layer Normalization and a Multi-Layer Perceptron (MLP), the signal passes through a 1 x 1 Conv and a 5 x 5 Conv. It is then explicitly split into three parallel streams. Local Channel Attention, Global Channel Attention, and Spatial Attention, before the final concatenation. By learning to generate spatial masks that assign low weights to clear pixels, the network ignores easy regions of the image.

4.4. Multi-Class Object Detection Model

The terminal Stage-1 phase of the onboard processing chain is the extraction of actionable intelligence via object detection, utilizing modified YOLO11m architecture as seen in [2]. The foundation for this model was already established in the prior work: *"The LTU Novel EO dataset is founded on the fundamental question: What are certain events that can be detected from satellite orbit?… The dataset will be further enhanced by a greater number of labeled images, different object category images, and simulation of atmospheric conditions"[2].*

To meet this progression, the legacy model was fundamentally retrained using the Clear Raw Optical Imagery (Level-0) as shown in the Fig. 3(d). This model relies mainly on a CSPDarknet-based *Backbone* utilizing sequential Cross-Stage Partial with kernel size 2 (C3k2) modules and a Spatial Pyramid Pooling Fast (SPPF) layer. It also integrates Cross-Stage Partial with Spatial Attention (C2PSA) blocks within both the backbone and the Neck to enhance feature extraction across multi-scale targets. Finally, the decoupled *Detection Head* calculates bounding box loss and Classification loss independently.

5. Results and Discussion

To ensure the physical accuracy and statistical validity of the raw optical imagery simulation detailed in Section 3, a validation approach was applied prior to deep learning model training. The simulated sensor degradation was validated against the hardware specifications of the target optical CMOS sensor model and were set to constant. A fractional factorial design approach was chosen with Structural Similarity Index Measure (SSIM) as the response variable (Y) and five main experimental factors (X). These factors include altitude, off-nadir angle, AOD, exposure and Sun elevation. After 16 runs, the off-nadir angle and exposure factors came out to be the most important configuration settings that could entirely degrade the quality of the image.

Based on this experiment, the *clear raw optical satellite imagery* dataset included images that were generated at all possible exposure and off-nadir configuration settings for the intended mission profile. Ultimately, the most robust verification of a synthetic simulation pipeline is its zero-shot empirical performance on in-orbit real-world data.

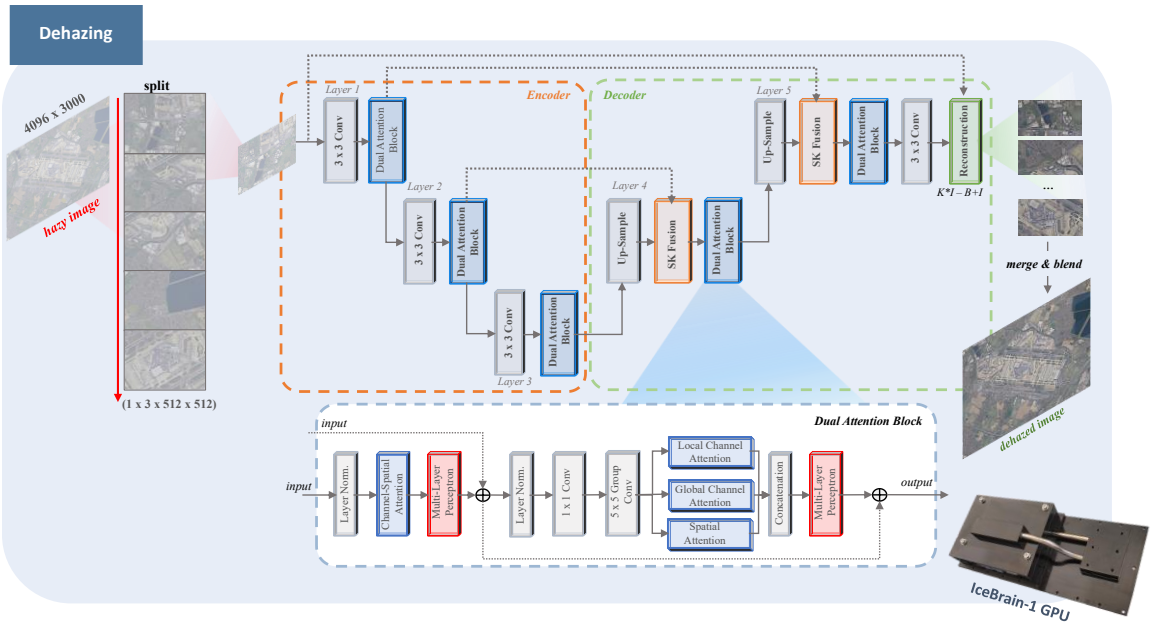

Fig. 8 Edge Optimized Dual Attention Dehazing Network (DA-Net) [33]

The primary objective of this research is to investigate whether the integration of a hierarchical, processing pipeline improves telemetry bandwidth efficiency and actionable data yield under strict edge hardware constraints. Based on the empirical data extracted from the flight-representative deployment, this section presents a training, evaluation, and analysis of the scene classification gatekeeper and dehazing module under operational conditions.

5.1. Training and Evaluation

Scene Classification Network

The Gated Fusion Network demonstrated convergence during training for 150 epochs. As illustrated in Fig. 9 (top), the Validation F1 (Macro) score stabilized efficiently, achieving an overall accuracy of 98.7% on the isolated test set. The validation of architecture is observed in the Confusion Matrix (Fig. 9, right). Apart from the proposed custom classification model, several standard classification architectures were also trained in parallel. Models like ConvNext family [36], Swin Transformer [37], and Vision Transformer (ViT) [38] outperformed the custom model achieving better macro-F-1 score on test data. However, they suffered from severe false positive rates when differentiating between meteorological clouds and highly reflective terrestrial topology like snow during inference tests. In the proposed custom model, the integration of the Texture Feature Stream successfully decoupled these classes. Out of 486 *clear_snow* samples, only one was misclassified as *heavy_cloud*, representing a near-perfect spatial frequency separation.

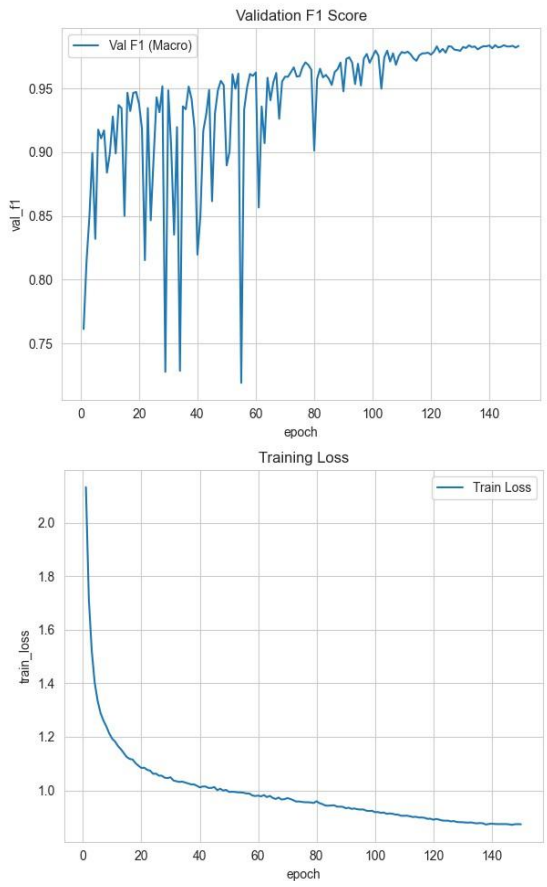

Fig. 9 Training and Evaluation of the Gatekeeper Scene Classification. (top) validation F1 Score. (bottom) training loss

Minor misclassifications were constrained to physically similar aerosol phenomena, such as dust being occasionally confused with smoke (5 overall instances) or seaside (7 instances).

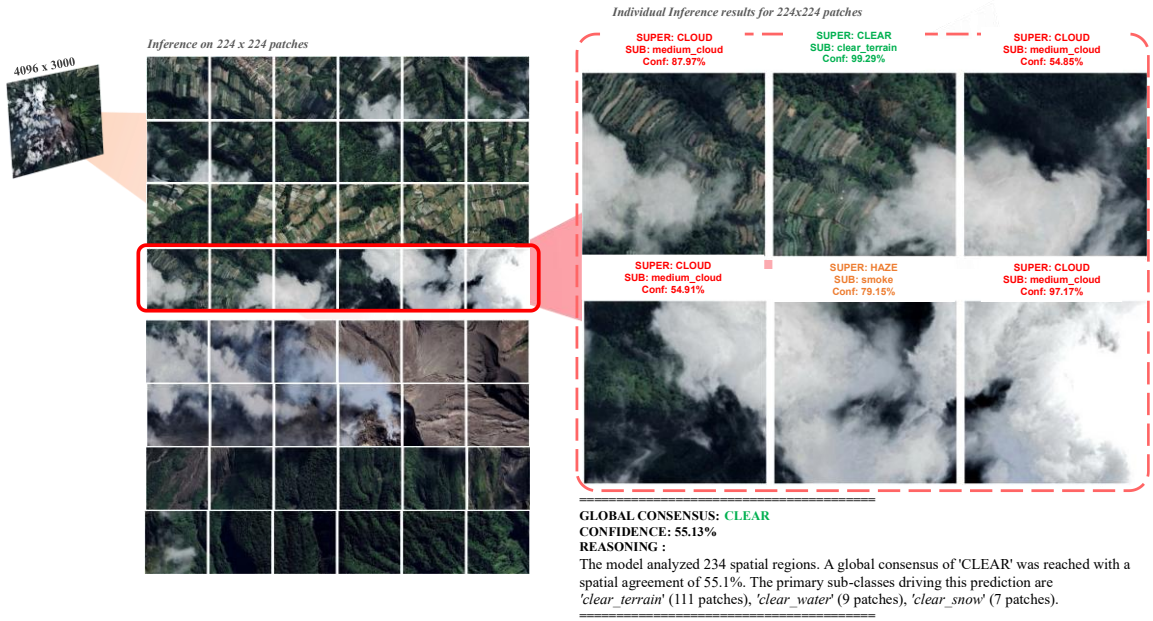

Fig. 10 Operational Scene Classification Inference

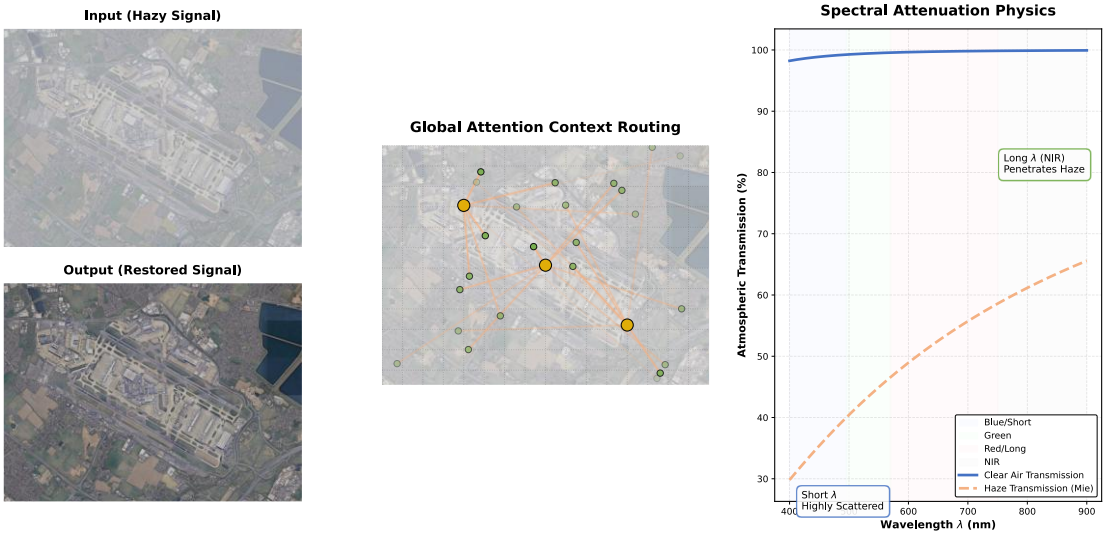

Fig. 11 Qualitative Dehazing Inference

Dual Attention Dehazing Network

The DA-Net architecture was trained with 4446 hazy images. These images include the simulated hazy images combined with existing datasets like RSID dataset [36]. The architecture was initially trained on 224 x 224 image patches achieving 0.92 Structural Similarity Index (SSIM). After evaluating the network on ground-truth data, the network was re-trained with 512 x 512 image patches. The training loss decayed smoothly across 120 epochs. Concurrently, the Validation Peak Signal-to-Noise Ratio climbed rapidly, stabilizing near 19.6dB. On the other hand, the Validation SSIM converged tightly around 0.67 during the intermediate training phase. To deploy this model within the <10 GMacs ceiling, the network was trained with composite L1, Perceptual loss [40] and subjected to Post-Training Quantization (PTQ) [41].

5.2. Scene Classification Inference

To evaluate the operational viability of the *Autonomous Triage Logic*, the classification model was subjected to a full-scale raw 4096 × 3000 acquisition containing heterogeneous atmospheric conditions (a volcanic event). Fig. 10, shows the model inference results on individual 224×224 spatial patches.

Rather than discarding the entire image due to partial cloud cover, the gatekeeper routed the data on a per-patch basis. It successfully identified opaque *medium_cloud* patches and *clear_terrain* patches.

Crucially, the model identified highly localized smoke patches (79.15% confidence) near the volcanic vent. Ultimately, the inference resulted in a Global Consensus of *CLEAR* (55.13% spatial agreement across 234 regions), proving that the pipeline can intelligently salvage actionable intelligence from partially obscured orbital passes. The result of the model is only the reasoning text as seen in the Fig. 10 (bottom-right).

5.3. Dehazing Network Inference

To evaluate the raw reconstruction capability of the DA-Net architecture, isolated offline evaluations were performed on heavily obscured 512x512 spatial patches against their ground-truth counterparts. During terminal inference on these targeted patches, the network achieved a quantitative reconstruction fidelity of 19.69 dB PSNR and SSIM of 0.91. Compared to this baseline, uncorrected atmospheric inputs (which measured 14.21 PSNR and 0.7189 SSIM), the dehazing algorithm delivered a significant enhancement of +5.48 dB PSNR and +0.195 SSIM. As visualized in the qualitative comparisons (Fig. 11), the Attention-based Sparsity Filter successfully preserved the native colorimetry of the underlying terrain while stripping the non-homogeneous atmospheric mask. Full-frame 4096 x 3000 reconstruction metrics, including merge and blend operations, are detailed separately in Table 3. Visualization of the Global Attention Context Routing (as seen in Fig. 11) mechanism, demonstrating how local features (small green nodes) communicate with global routing hubs (large orange nodes) to efficiently capture long-range contextual dependencies across the image grid.

5.4. Deployment Results

The models were deployed on the target flight-representative hardware, evaluating classification efficiency on the low-power CPU node and dehazing performance on the GPU node.

Scene Classification

For the classification, three models were selected after the training and evaluation based on the size & weight of the models. As shown in the Table 2, the custom IceBrainSceneNet was benchmarked against the State-of-the-Art (SOTA) ConvNeXt architecture (Base and Small variants) in both FP32 and quantized FP16 formats. The results demonstrate the extreme architectural parsimony of the proposed model. While the ConvNeXtBase model required 171MB of memory and 88.7 million parameters, yielding 0.39 Frames Per Second (FPS). The IceBrainSceneNet required less than 1MB of memory (0.46 million parameters). This 99.4% reduction in parameter space allowed the FP16-quantized model to achieve an inference speed of 1.72 FPS, while consuming only 2.13W of power.

Dehazing Network

When the gatekeeper identifies a hazy scene, the image is passed to the GPU for restoration. Table 3 details the performance of the Attention-based dehazing network on native, high-resolution (4096 x 3000) IceBrain-1 simulated imagery. The power and temperature were obtained directly onboard using software command tools like NVIDIA *tegrastats* and *jtop*. The network achieved robust structural restoration, yielding Peak Signal-to-Noise Ratios (PSNR) up to 24.26 dB and Structural Similarity Index Measures (SSIM) peaking at 0.869. However, the telemetry highlights the stringent operational constraints of edge computing: processing 12-megapixel raw imagery induced transient GPU utilization spikes of 99.7%, elevating the System on Chip (SoC) temperature to ~60°C and causing power draws up to 19.5 W leading to thermal throttling. To validate the hypothesis established in Section 1, that edge models require training on raw sensor characteristics rather than ground-processed datasets, the dehazing network underwent an out-of-distribution inference test using unseen Sentinel-2 imagery (Table 3). The results starkly illustrate the "Domain Gap". While the model achieved an SSIM of ~0.85 on the native IceBrain-1 sensor data, performance on Sentinel-2 imagery degraded to an average SSIM of ~0.50 (minimum 0.380). Despite faster inference times (due to the varying spatial dimensions and sparsity of the S2 inputs), the network struggled to reconstruct structural fidelity.

Table 2 Inference in FPS, Power and Temperature readings of Scene Classification Models

Model	Size (KB)	Weight (Million)	Power (W)	Inference (FPS)
ConvNextBase-FP32	171051	88.7	2.59	0.39
ConvNextBase-FP16	171051	88.6	2.58	0.40
ConvNextSmall-FP32	2972	28.6	2.58	0.70
ConvNextSmall-FP16	2972	28.5	2.58	0.71
IceBrainSceneNet-FP32	915	0.4648	2.25	1.70
IceBrainSceneNet-FP16	915	0.4648	2.13	1.72

Table 3 Dehazing network deployment inference results on native IceBrain-1 imagery and Sentinel-2 imagery

Imagery	Time (s)	Temp (C) ± 1.2	Power (W) ± 0.4	PSNR	SSIM
IceBrain-1 (4096 x 3000)	13.081	56.9	14.9	19.786	0.8521
	12.304	54.6	18.7	24.261	0.8238
	17.251	58.8	12.2	20.001	0.6133
	12.577	58.8	15.3	22.549	0.8694
Sentinel -2 (varied resolution)	5.999	56.1	15.3	8.855	0.3805
	4.855	56.6	18.5	14.265	0.6083
	4.864	56.8	18.0	14.184	0.6848
	4.923	57.1	18.6	12.686	0.6056

This drop in PSNR and SSIM proves the domain-gap and that deep learning models optimized for processed Level-1C/Level-2A data are incompatible with raw edge deployments. confirming the necessity of the proposed physics-based simulation pipeline developed to train edge-agnostic networks.

5.5. Ablation Study

To quantify the system-level benefits of the proposed framework, an ablation study was conducted. This experiment evaluates the architectural efficiency of heterogeneous onboard computing.

Does Dehazing Improve Downstream Detection Accuracy?

The ultimate validation of Stage-0 image restoration is its impact on actionable intelligence (Stage-1). As visualized in the positive feedback results in Fig. 12 (left), feeding the raw, hazy images directly into the object detection network resulted in catastrophic gradient loss, failing to detect critical infrastructure. Following the dehazing process, structural contrast was recovered, enabling the model to confidently detect complex targets including *harbor* (0.67), *airport* (0.76), and *dam* (0.85).

However, deep learning models inherently possess a few limitations. As demonstrated in the negative feedback results in Fig. 12(right), the dehazing algorithm occasionally struggled with highly complex, non-homogeneous terrestrial phenomena. In scenes featuring active volcanic where smoke is mixed with dense meteorological clouds, the model exhibited spatial over-smoothing, slightly degrading the confidence of adjacent bounding box predictions. This empirical limitation underscores the need for localized attention refinement in future iterations, ensuring the network does not inadvertently suppress high-frequency structural anomalies during the dehazing process.

Operational Impact of Autonomous Triage

To quantify the system-level benefits of the proposed framework, an ablation study was conducted focusing on the operational impact of the Autonomous Triage Logic. Rather than evaluating pure algorithmic speed, this experiment evaluates the architectural efficiency of heterogeneous onboard computing. Mainly focusing on the computational, thermal, and telemetry savings achieved by preventing heavily occluded imagery from unnecessarily booting the high-power payload GPU. The ablation test was executed on the target payload hardware, which utilizes a dual-node heterogeneous architecture: an Intel E3900 Atom CPU and an NVIDIA Jetson Orin Nano GPU. This design is a deliberate systems engineering choice for power-constrained CubeSats. The low-power CPU acts as an "always-on" gatekeeper managing initial triage, while the high-power GPU is kept in a low-power idle state, waking only to execute the intensive attention-based dehazing and YOLOv11m object detection models. A representative batch of ten high-resolution (4096 x 3000) raw images was evaluated. To reflect the climate average of >50% atmospheric occlusion, the batch contained 6 heavily occluded scenes and 4 clear/recoverable scenes. In the *Baseline Pipeline (Single-Node, No Triage)*, forcing all ten images through the GPU for full feature-extraction (object detection) accumulated a significant energy debt of approximately 2,025 Joules. More critically, as observed in previous dehazing tests, this continuous GPU utilization induced transient peak power spikes up to 19.5W.

Fig. 12 Dehazing improving detection (left), Dehazing not improving detection (right)

By contrast, the *Proposed Pipeline (Dual-Node, With Triage)* exhibited architectural stability. The CPU-based Gatekeeper comprehensively analysed 352 spatial regions per image, operating at a stable peak power of just ~2.74W. It processed the 10-image batch in 645 seconds, consuming 1,743.4 Joules. Crucially, the Gatekeeper successfully identified all 6 occluded images as CLOUD. By immediately terminating these pipelines, the system prevented the GPU from waking up for 60% of the workload. The remaining 4 CLEAR images bypassed the dehazing network entirely, passing directly to the YOLOv11m detector on the GPU (executing at an efficient ~0.4s and ~2.5J per image, peaking at 7.01W).

By utilizing the CPU as a low-power sieve, the total active system energy expenditure was reduced by ~13.4%. More importantly, the absolute maximum peak power load placed on the satellite bus was reduced by 64% (from 19.5W down to 7.01W), drastically improving the thermal safety margins of the payload. Furthermore, by filtering the 6 opaque clouds prior to downlink, the required telemetry buffer dropped from 350 MB to just 140 MB, a 60% bandwidth reduction.

Mission Lifetime Extrapolation

To contextualize these savings, the results were extrapolated over the anticipated 3-year lifespan of a VLEO technology demonstration mission (e.g., the upcoming *Atmos Space Cargo Phoenix 2.2*). Assuming an active payload yield of 1.31 million images over the mission's life, the 60% bandwidth reduction translates to the prevention of 788,400 unusable images from entering the telemetry queue. At 35 MB per raw image, this equates to 27.59 Terabytes of saved telemetry. In X-band communication architectures, this bandwidth salvage equates to thousands of hours of pass time reallocated exclusively for the transmission of refined, actionable intelligence. Ultimately, this ablation study suggests that heterogeneous onboard triage is not only an algorithmic convenience, but an architectural framework for sustainable orbital operations. Therefore, the empirical results from the hardware deployment and ablation study validate the primary hypothesis and disprove the baseline assumption.

6. Conclusion

This research presented an end-to-end Orbital Edge Computing framework designed to mitigate the fundamental data bottlenecks of Earth Observation SmallSats. By developing a physics-based Raw Optical Imagery Simulator, the critical lack of Level-0 training data was addressed, mapping complex radiometric and atmospheric scattering onto commercial sensor profiles. To process this data within the strict constraints of a 16U CubeSat or similar, a heterogeneous dual-network architecture was validated. The deployment of a lightweight Dual-Stream CNN on an Intel Atom CPU successfully acted as a gatekeeper, selectively routing only recoverable imagery to an efficient (5.47 GMacs) Dual Attention Dehazing Network deployed on GPU.

The results demonstrated that by discarding 60% of heavily clouded imagery prior to GPU engagement or downlink (350 MB to 140MB), the architecture reduced active thermal energy debt by over 56% and

increased effective actionable downlink capacity. Restoring raw hazy data onboard directly enabled robust multi-class object detection, preventing domain-shift failures commonly observed when ground-trained AI models are subjected to the harsh realities of the space environment. Furthermore, while the current study utilized the CPU for initial triage to strictly adhere to peak power constraints, the extended inference time resulted in a higher net energy consumption per image. For future orbital missions, it is recommended to investigate deploying the scene classification module on the GPU. Exploiting the GPU's parallel processing capabilities would drastically reduce inference latency, which may ultimately lower the total energy expenditure per image despite the higher transient power draw.

While this study validated the framework on flight-representative terrestrial testbeds, the natural progression of this research is an In-Orbit Demonstration (IOD). Future work will focus on deploying this heterogeneous pipeline aboard the upcoming *Atmos Space Cargo Phoenix 2.2* VLEO mission. Planned updates include studies on Stage-2 (Fig. 6) applications featuring fine-grained use-cases. Furthermore, studies are planned on Continuous Edge Learning architectures, specifically utilizing Knowledge Distillation to analyze if any improvements in edge-based detection accuracy can be achieved. Overall, the focus of this research is to enable autonomous, cognitive Earth Observation systems.

Author contributions B.C. wrote the main manuscript text, did the conceptualization, methodology, investigation, visualization, and formal analysis. A.P. participated in validation. J.H. participated in validation. R.L. performed validation and project administration. M.L. performed validation and provided supervision. J.E. performed supervision, project administration and provided resources.

Funding Open access funding provided by Lulea University of Technology. The authors declare that no funds, grants, or other support was received during the preparation of this manuscript.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Marcuccio, S., Ullo, S., Carminati, M., Kanoun, O.: Smaller satellites, larger constellations: trends and design issues for Earth observation systems. IEEE Aerosp. Electron. Syst. Mag. 34, 50–59 (2019). https://doi.org/10.1109/MAES.2019.2928612
2. Chintalapati, B., Jafari, A., Laufer, R., Liwicki, M., Eickhoff, J.: Advancing multi-class object detection from LEO/VLEO: model evaluation and onboard deployment tailored for a 16U CubeSat. CEAS Space J. (2026). https://doi.org/10.1007/s12567-025-00697-6
3. Schott, J.R.: Remote sensing: the image chain approach. Oxford University Press, Oxford (2023)
4. Knapek, M., Horwath, J., Moll, F., Epple, B., Courville, N., Bischl, H., Giggenbach, D.: Optical high-capacity satellite downlinks via high-altitude platform relays. In: Proceedings of the SPIE 6709, Free-Space Laser Communications VII (2007)
5. King, M.D., Platnick, S., Menzel, W.P., Ackerman, S.A., Hubanks, P.A.: Spatial and temporal distribution of clouds observed by MODIS onboard the terra and aqua satellites. IEEE Trans. Geosci. Remote Sens. 51, 3826–3852 (2013). https://doi.org/10.1109/TGRS.2012.2227333
6. Stubenrauch, C.J., Rossow, W.B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B.C., Menzel, W.P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., Zhao, G.: Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel. Bull. Am. Meteorol. Soc. 94, 1031–1049 (2013). https://doi.org/10.1175/BAMS-D-12-00117.1
7. Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., Gascon, F.: Sen2Cor for Sentinel-2. In: SPIE Remote Sensing, p. 3 (2017)
8. Helber, P., Bischke, B., Dengel, A., Borth, D.: EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification. In: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 204–207 (2019). https://doi.org/10.1109/IGARSS.2019.8900609
9. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: AOD-Net: All-in-one dehazing network. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 770–778 (2017). https://doi.org/10.1109/ICCV.2017.90
10. Song, Y., He, Z., Qian, H., Du, X.: Vision transformers for single image dehazing. IEEE Trans. Image Process. 32, 1920–1935 (2023). https://doi.org/10.1109/TIP.2023.3255146
11. Giuffrida, G., Fanucci, L., Meoni, G., Batic, M., Buckley, L., Dunne, A., Van Dijk, C., Esposito, M., Hefele, J., Vercruyssen, N., Furano, G., Pastena, M., Aschbacher, J.: The Φ-Sat-1 mission: the first on-board deep neural network demonstrator for satellite Earth observation. IEEE Trans. Geosci. Remote Sens. 60, 1–14 (2022). https://doi.org/10.1109/TGRS.2021.3125567
12. EUMETSAT: First images from Meteosat Third Generation Sounder 1. EUMETSAT. https://www.eumetsat.int/image-week-first-images-meteosat-third-generation-sounder-1 (2026). Accessed 23 Apr 2026
13. Fourest, S.: Satellite imagery: from acquisition principles to processing of optical images for observing the Earth. Cépaduès Éditions, Toulouse (2012)
14. Frazier, A.E., Hemingway, B.L.: A technical review of Planet SmallSat data: practical considerations for processing and using PlanetScope imagery. Remote Sens. 13, 3930 (2021). https://doi.org/10.3390/rs13193930
15. Diana, L., Dini, P.: Review on hardware devices and software techniques enabling neural network inference onboard satellites. Remote Sens. 16, 2182 (2024). https://doi.org/10.3390/rs16122182

16. Xia, G.S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., Datcu, M., Pelillo, M., Zhang, L.: DOTA: A large-scale dataset for object detection in aerial images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3974–3983 (2018). https://doi.org/10.1109/CVPR.2018.00418
17. Loveland, T., Dwyer, J., Woodcock, C., Roy, D.: Landsat 8 status and Landsat/Sentinel-2 synergy. NASA (2014)
18. Sentinel Data Access Annual Report 2023. Copernicus (2024). COPE-SRCO-RP-2400521
19. Zhu, Z., Woodcock, C.E.: Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ. 118, 83–94 (2012). https://doi.org/10.1016/j.rse.2011.10.028
20. Mateo-García, G., Aybar, C., Acciarini, G., Růžička, V., Meoni, G., Longépé, N., Gómez-Chova, L.: Onboard cloud detection and atmospheric correction with deep learning emulators. In: IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, pp. 1875–1878 (2023). https://doi.org/10.1109/IGARSS52108.2023.10282462
21. Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M., Morcrette, J.J.: Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An overview. IEEE Trans. Geosci. Remote Sens. 35, 675–686 (1997). https://doi.org/10.1109/36.581987
22. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. 33, 2341–2353 (2011). https://doi.org/10.1109/TPAMI.2010.168
23. Liu, X., Ma, Y., Shi, Z., Chen, J.: GridDehazeNet: Attention-based multi-scale network for image dehazing. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 7314–7323 (2019). https://doi.org/10.1109/ICCV.2019.00741
24. Giggenbach, D., Shrestha, A.: Atmospheric absorption and scattering impact on optical satellite-ground links. Int. J. Satell. Commun. Network. 40, 157–176 (2022). https://doi.org/10.1002/sat.1426
25. Eickhoff, J.: Payload for Onboard AI-based Image Recognition. (2023). https://doi.org/10.13140/RG.2.2.14216.39686
26. Janesick, J.R.: Scientific charge-coupled devices. SPIE Press, Bellingham (2001)
27. Toutin, T.: Geometric processing of remote sensing images: models, algorithms and methods. Int. J. Remote Sens. 25, 1893–1924 (2004). https://doi.org/10.1080/0143116031000101684
28. Czerkawski, M., Atkinson, R., Michie, C., Tachtatzis, C.: SatelliteCloudGenerator: Controllable cloud and shadow synthesis for multi-spectral optical satellite images. Remote Sens. 15, 4138 (2023). https://doi.org/10.3390/rs15174138
29. Ba, R., Chen, C., Yuan, J., Song, W., Lo, S.: SmokeNet: Satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention. Remote Sens. 11, 1702 (2019). https://doi.org/10.3390/rs11141702
30. Crawshaw, M.: Multi-task learning with deep neural networks: a survey. arXiv preprint arXiv:2009.09796 (2020)
31. Bhuiyan, A., Liu, Y., Siva, P., Javan, M., Ben Ayed, I., Granger, E.: Pose guided gated fusion for person re-identification. In: 2020 IEEE International Conference on Image Processing (ICIP), pp. 101–105 (2020). https://doi.org/10.1109/ICIP40778.2020.9190757
32. Hou, Q., Zhou, D., Feng, J.: Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 13713–13722 (2021). https://doi.org/10.1109/CVPR46437.2021.01350
33. Kim, N., Choi, I.S., Han, S.S., Jeong, C.S.: DA-Net: Dual attention network for haze removal in remote sensing image. IEEE Access 12, 136297–136312 (2024). https://doi.org/10.1109/ACCESS.2024.3459588
34. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
35. Sun, L., Wang, Y., Zhu, Q.: Dual-attention-based block matching for dynamic point cloud compression. J. Imaging 11, 15 (2025). https://doi.org/10.3390/jimaging11100332
36. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A ConvNet for the 2020s. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11976–11986 (2022). https://doi.org/10.1109/CVPR52688.2022.01167
37. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin Transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 10012–10022 (2021). https://doi.org/10.1109/ICCV48922.2021.00986
38. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations (ICLR) (2021)
39. Zhu, Z.H., Lu, W., Chen, S.B., Ding, C.H.Q., Tang, J., Luo, B.: Real-world remote sensing image dehazing: benchmark and baseline. IEEE Trans. Geosci. Remote Sens. 63, 1–17 (2025). https://doi.org/10.1109/TGRS.2025.3584234
40. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: European Conference on Computer Vision (ECCV), pp. 694–711. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_43
41. Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., Kalenichenko, D.: Quantization and training of neural networks for efficient integer-arithmetic-only inference. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2704–2713 (2018). https://doi.org/10.1109/CVPR.2018.00286