IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

ROMEO: Raw Onboard Multispectral Earth Observation — An Edge-AI Framework for Latency- and Bandwidth-Efficient Satellite Sensing

Charan M H[1], Abhishek G[2], Haridarshan K P[3], Mallikarjuna[4], K B Janaki*[5]

Department of Computer Science and Engineering, East West College of Engineering

Visvesvaraya Technological University, Bengaluru, Karnataka, India[1-5]

*Corresponding Author

Abstract: Wildfire and volcanic-eruption alerts from orbit are delayed less by detection accuracy than by the ground processing chain. This paper reports a working onboard pipeline that analyses raw Sentinel-2 Level-0 granules without any ground-segment product: radiometric calibration, coarse spatial coregistration and coarse georeferencing are applied on board, a spectral gate discards cloudy scenes, thirteen bands are reduced to the three carrying thermal signal, and a 23,714-parameter network quantised to 29.9 KB scans the granule. On an 18-granule catalogue the system reaches 100 % granule accuracy with no false detections, and classifies 240 of 240 unseen patches correctly under leave-granules-out evaluation. Removing coregistration costs 16.67 percentage points, identifying it as the enabling step. Quantisation reduces inference latency eightfold, and transmitting alerts instead of imagery reduces 30.67 MB to 105.74 KB. Because the granules are synthesised, the results characterise pipeline correctness rather than field performance.

Keywords: Coarse spatial coregistration, convolutional neural network, edge computing, integer quantisation, onboard artificial intelligence, raw Level-0 data, Sentinel-2, thermal hotspot detection, wildfire detection.

I. INTRODUCTION

Earth observation satellites acquire far more data than they can transmit. A Sentinel-2 granule of the extent used in this work occupies 1.70 MB across the thirteen instrument bands, and a mission operating under the conventional store-and-forward model must downlink every such granule, calibrate and orthorectify it on the ground, and only then search it for events of interest [1]. For cartography or land-cover statistics this is acceptable. For a wildfire or a volcanic eruption it is not: the quantity that matters is the delay between the sensor observing a thermal anomaly and a responder learning of it, and in a bent-pipe architecture that delay is set by the next ground station pass and the ground processing queue rather than by anything the detector does.

Moving the detector on board removes the ground loop. The satellite analyses its own imagery and transmits a short alert instead of a scene, so the link carries kilobytes of conclusions rather than megabytes of pixels. The Phi-sat-1 mission demonstrated that a convolutional network can be executed in orbit for cloud screening [2], and subsequent work has applied learned compression [3], [4] and eruption classification [5] to the same setting. These demonstrations share an assumption that is rarely examined: the network is applied to a high-level product whose production depends on the very ground chain that onboard processing is meant to avoid.

Operating directly on Level-0 data instead removes that contradiction, at the cost of confronting the raw signal as the sensor delivers it. Bands of a push-broom instrument are not mutually registered, because the detector arrays for different wavelengths occupy different positions in the focal plane; the short-wave infrared bands are read out inverted on alternate detector modules; the digital numbers carry a dark-current offset and a per-column gain variation; and single-event upsets induced by cosmic radiation corrupt isolated pixels [1]. Meoni et al. [6] addressed this by releasing THRawS, the first dataset of raw Sentinel-2 granules annotated for thermal hotspots, together with PyRawS and a lightweight coarse spatial coregistration (CSC) scheme whose cost is a fixed memory remap rather than a feature-matching search.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

In earlier work [18] the present authors surveyed five studies in this area, catalogued the limitations observed in each, and proposed a six-stage onboard pipeline intended to address them. That work was explicitly at the concept stage: the pipeline had not been built, and no accuracy value, latency measurement, or comparison was reported. The present paper closes that gap. Its contributions are as follows. First, the proposed pipeline is realised as a working system in which every stage from Level-0 ingestion to alert generation executes without any ground-segment product. Second, the contribution of coarse spatial coregistration is isolated by ablation: the same network is evaluated with and without the alignment stage, which converts a design assumption into a measured quantity. Third, the cost of deployment-oriented integer quantisation is measured in all three of the dimensions that matter on board - model size, inference latency, and agreement with the full-precision model. Fourth, detection accuracy is reported under a leave-granules-out protocol that excludes whole scenes, and one scene type entirely, from training, so that the figure quoted is not inflated by patches drawn from scenes the model has already seen. Fifth, the pipeline is exercised under injected single-event-upset noise, which quantifies what the radiometric stage contributes to robustness against the failure mode identified as dominant in [1]. Sixth, the objectives stated at the concept stage are revisited and the status of each is reported without embellishment.

The remainder of the paper is organised as follows. Section II reviews the literature and states the position of this study within it. Section III describes the implemented architecture, the reference data, the algorithms and the evaluation protocol. Section IV reports and discusses the results, including the status of the objectives stated at the concept stage and the threats to validity. Section V concludes.

II. LITERATURE REVIEW

Research on artificial intelligence aboard Earth observation satellites divides into three concerns: whether a network can be executed at all within the power and radiation envelope of a spacecraft, what it should be asked to compute, and what form of data it should receive.

Furano et al. [1] survey the first concern and conclude that the binding constraints are the uplink and downlink budget and the space radiation environment, in which single-event effects and single-event upsets corrupt both memory and arithmetic. Their analysis motivates two design decisions adopted here: the model must be small enough to be held and executed within a modest memory budget, and the pipeline must tolerate corrupted input pixels rather than assume clean data.

Giuffrida et al. [2] report the Phi-sat-1 mission, in which the CloudScout network was executed in orbit to discard cloud-covered scenes before downlink. The mission established that onboard inference is practicable, and the paper identifies the scarcity of mission-specific training data as the principal obstacle encountered. The cloud gate of Section III-C is the functional analogue of CloudScout in the present system, although it is implemented as a spectral test rather than a network because its only task here is to protect the detection stage from unusable scenes.

Guerrisi et al. examine onboard compression, first with a convolutional autoencoder [3] and subsequently as a mission-level encoder-decoder pair for Phi-sat-2 [4]. Compression and detection are complementary responses to the same bandwidth problem: compression reduces the cost of transmitting an image, whereas detection removes the need to transmit most images at all. Both studies note that onboard schemes must be simplified to remain within the available power and memory, and [4] observes a migration cost whereby accuracy measured in a well-resourced ground environment degrades once the model is moved to flight hardware.

Di Stasio et al. [5] assess the feasibility of detecting volcanic eruptions on board, reporting classification accuracies of 80 % and 70 % for the two metrics considered. The study is restricted to sulphur dioxide measurements from Sentinel-5P and does not extend to an autonomous onboard system, which the authors state as remaining work.

The third concern, the form of the data, is addressed directly by Meoni et al. [6]. They observe that models demonstrated on Level-1C products inherit a ground chain that cannot be flown, propose a methodology for annotating raw Level-0 granules by projecting Level-1C detections back onto them, and release both the THRawS dataset and the PyRawS toolkit [8], [9]. Their coarse spatial coregistration compensates only the systematic component of the band-to-band shift, using one-time precomputed values, and therefore requires neither ground control points nor keypoint matching; they report 16.63 ms per granule on a central processing unit for this operation. The reference detector used to annotate the dataset is the threshold scheme of Massimetti et al. [7], which activates a pixel from fixed ratios of top-of-atmosphere reflectance in the bands B8A, B11 and B12.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

TABLE I. Summary of reviewed studies and observed limitations

Study	Methods used	Limitation observed	Ref.
Furano et al. (2020)	Survey of edge AI in space systems	Radiation causes single-event effects and upsets in AI hardware; bandwidth is the binding constraint	[1]
Giuffrida et al. (2022)	CloudScout CNN flown on Phi-sat-1	Insufficient mission-specific training data before launch	[2]
Guerrisi et al. (2022)	Convolutional autoencoder for onboard compression	Schemes must be simplified to fit onboard power and memory	[3]
Guerrisi et al. (2023)	Encoder-decoder compression for Phi-sat-2	Accuracy degrades when models migrate from ground to flight hardware	[4]
Di Stasio et al. (2022)	ML classification of eruption precursors	Single modality (SO2 from Sentinel-5P); not an autonomous onboard system	[5]
Meoni et al. (2024)	THRawS dataset, PyRawS, coarse spatial coregistration	Dataset and tooling only; no end-to-end detection system reported	[6]

Table I condenses the studies reviewed at the concept stage [18] together with the limitation observed in each. Taken together they establish that onboard inference is feasible and that raw data is the appropriate input, but none reports an end-to-end system that ingests Level-0 imagery, performs its own alignment, detects an event, georeferences it, and emits an alert, with the contribution of each stage measured. The present study differs from [1]-[5] not in the algorithms it employs, which are deliberately standard, but in reporting that end-to-end implementation together with the measurements needed to judge it, and in stating plainly which of the intended properties were achieved.

III. METHODS AND MATERIALS

A. System Architecture Overview

The implemented system preserves the six-stage organisation proposed in [18]. Each stage is a separate module with a defined interface, so that a stage can be replaced without disturbing the others, and every stage from ingestion to alert generation operates without reference to any ground-segment product. Fig. 1 shows the arrangement and the direction of data flow.

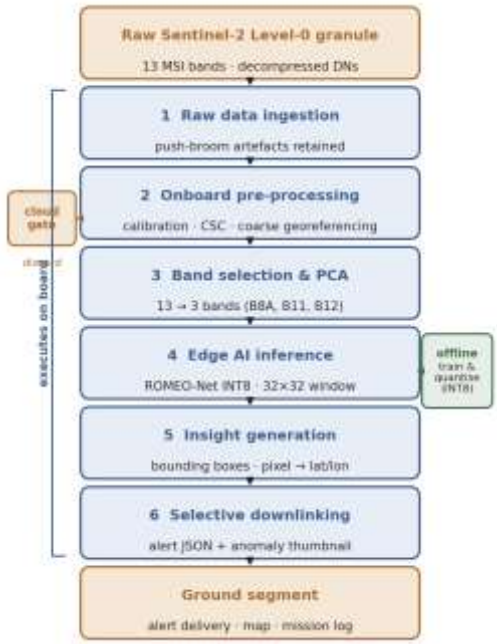


Fig. 1. Implemented six-stage pipeline. Data flow is downward from Level-0 ingestion to the ground segment. Stages 1 to 6 execute on board; the network is trained and quantised offline, and the cloud gate discards unusable granules before inference.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

B. Ingestion Stage

The input is a decompressed Level-0 granule: an array of 12-bit digital numbers per band, carrying the artefacts of the acquisition rather than a corrected image. Three bands are retained, B8A at 865 nm, B11 at 1610 nm and B12 at 2190 nm, all at the 20 m ground sampling distance of the instrument, which is the band collection used by the reference hotspot detector [7] and therefore by THRawS [6]. A granule in this study is a 256 x 256 pixel extract covering 5.12 km on a side. No radiometric or geometric correction has been applied at this point, and the bands are not mutually registered.

C. Pre-processing Stage

Three operations are applied in sequence, followed by a gate. Radiometric calibration removes the artefacts that the sensor imposes on the digital numbers. Impulsive outliers attributable to single-event upsets are identified by comparison with a 3 x 3 median and replaced only where the deviation exceeds a fixed threshold, so that the image is not otherwise smoothed; the per-column gain variation responsible for striping is flat-fielded against the column median; and the dark-current offset is subtracted. Each operation is separable and of linear cost in the number of pixels, which is the class of correction a small-satellite accelerator can afford [1].

Coarse spatial coregistration then aligns the bands. The short-wave infrared bands are de-rotated on even-numbered detector modules, and each band is translated by the fixed along-track and across-track displacement held in a per-satellite look-up table, as in [6]. The operation is a memory remap: no keypoints are extracted and no search is performed, which is what makes it affordable in flight. Border pixels invalidated by the translation are cropped.

Coarse georeferencing assigns coordinates to each band using the same displacement values applied to the granule corner coordinates supplied as ancillary metadata, following the scheme of [6]. This yields the mapping from pixel to latitude and longitude that Section III-F requires, without a digital elevation model or ground control points.

Finally a cloud gate estimates the fraction of the granule obscured by cloud from a spectral test that separates cloud from both snow and thermal anomalies, and granules exceeding 65 % cloud cover are discarded on board. The gate occupies the position held by CloudScout in [2]; its purpose here is to prevent the detection stage from expending effort, and the downlink from expending a slot, on a scene in which the ground is not visible.

D. Band Selection and Optimisation Stage

Two reductions are applied. The first is fixed at mission level: of the thirteen multispectral instrument bands, only the three that carry the thermal signature are acquired for this task, a 77 % reduction in data volume taken before any computation. The second is performed on board: principal component analysis of the retained cube reports the variance explained by each component and the loading of each band on the first component, which quantifies how much spectral redundancy remains and confirms that the short-wave infrared bands dominate the leading component over a thermal scene.

E. Detection Stage

A convolutional network, ROMEO-Net, is applied to overlapping 32 x 32 windows at a stride of 16 pixels, and the window probabilities are averaged into a per-pixel probability map. The network is held in 8-bit integer form and executed through a general-purpose inference runtime [14], since integer arithmetic is what the edge accelerators considered for this role provide natively. The classical threshold detector of [7] is computed alongside the network on the same coregistered granule, not as a component of the pipeline but as the reference against which the learned detector is judged.

F. Insight Generation Stage

The probability map is thresholded, connected components are extracted and small clusters rejected, and each surviving component is reduced to a bounding box. The box centre is converted to latitude and longitude by bilinear interpolation of the granule corner coordinates established in Section III-C, and the component area is converted to square kilometres from the ground sampling distance. The stage therefore converts a probability field into a list of located, sized and scored detections.

G. Selective Downlinking Stage

Only the alert is queued for transmission: a JSON object carrying the event class, the spacecraft and product identifiers, and the coordinates, confidence and area of each detection, together with a cropped thumbnail of the anomaly for visual confirmation. The granule itself is never transmitted. A granule in which no detection survives, or which the cloud gate rejected, produces no downlink at all.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

H. Implementation

The pipeline is implemented in Python. Array processing uses NumPy and SciPy, image rendering uses Pillow, and inference uses ONNX Runtime [14] with the quantised model; these four packages are the entire runtime dependency set, so the system runs on a single workstation without specialised hardware. Training is performed offline in JAX [15] and is not part of the flight path. A Flask application exposes the pipeline through a dashboard that renders every stage of a selected granule, a world map of granule footprints, an evaluation view that reproduces the experiments of Section IV on demand, a real-time comparison of downlink behaviour with and without the system, and a mission log to which every generated alert is appended; alerts may additionally be forwarded to an external channel. Correctness is protected by an automated suite of 30 tests covering the pipeline invariants, including the recovery of a known band displacement by the coregistration stage, the recovery of a known injected corruption by the calibration stage, the behaviour of the reference detector on constructed reflectances, and agreement between the integer and floating-point models.

I. Reference Granule Construction

Real THRawS granules [9] are distributed as an archive of approximately 100 GB, which exceeded the storage and bandwidth available to this study. A catalogue of reference granules was therefore synthesised from a physical sensor model so that the pipeline could be verified against a known ground truth. This is a deliberate methodological trade-off and its consequences are examined in Section IV-M; the import path for real granules is implemented and is exercised by the same code, so the substitution is one of data rather than of method.

Eighteen granules were generated, each carrying the identity, class, centre coordinates, acquisition date and spacecraft of a real event drawn from the THRawS event database, which comprises 88 events of which 58 are volcanic eruptions, 20 are wildfires and 10 are event-free. Surface reflectance is generated as a multi-octave value-noise field, and scene-specific behaviour is imposed for the event-free classes so that the catalogue contains the confusers a detector must reject: overcast ocean, bright desert, a snowfield that is bright in the near infrared but dark in the short-wave infrared, high-frequency urban texture, a river delta with dark water, and rainforest.

Thermal anomalies are generated as irregular fronts whose spectral behaviour follows the Planck relation over the three retained bands, so that emission grows with wavelength. Writing ρ_b for the reflectance in band b and m for the normalised hotspot mask,

$$\rho_b = \rho_b^{(terrain)} + a_b \cdot m + c_b \cdot m^2, \quad b \in \{B8A, B11, B12\} \qquad (1)$$

with $a_{12} > a_{11} > a_{8A}$, so that a hotspot is strongly B12-dominant and saturates at the 12-bit ceiling, which is the property the reference detector of [7] exploits.

The acquisition is then modelled. Following [6], the band-to-band displacement is decomposed as

$$S_{(Bn,Bm)} = \bar{S}_{(Bn,Bm)} + \Delta S_{(Bn,Bm)} \qquad (2)$$

where the first term is the systematic shift due to the push-broom geometry and the second is an aleatory component arising from mechanical disturbance. Only the systematic term is injected, by translating each band by its tabulated displacement, and the short-wave infrared bands are inverted on even detector modules. A per-column gain variation, additive read noise, a dark-current offset and salt-and-pepper corruption representing single-event upsets are applied, and the result is clipped to the 12-bit range. The catalogue is generated deterministically from a fixed seed and is therefore exactly reproducible.

J. Coarse Spatial Coregistration

The pre-processing stage inverts the systematic component of the displacement. For a band B_k with tabulated along-track and across-track displacement s_k,

$$\hat{B}_k(i, j) = B_k(i - s_k^{at}, j - s_k^{ct}) \qquad (3)$$

the inverse translation being applied with zero fill and the invalidated border subsequently cropped. Because only the systematic term is compensated, the residual misregistration is the aleatory term of (2) together with the quantisation of the tabulated values, and it is this residual that Section IV-D measures. Per-band corner coordinates follow from the same displacement applied to the granule corners,

$$PC_{(Bk)} = PC_{(B8A)} + |s_k| \cdot G_a / G_L \qquad (4)$$

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

where G_a is the granule extent in degrees and G_L its length in pixels, which is the coarse georeferencing scheme of [6].

K. Reference Detector

The classical detector is the simplified Massimetti scheme as implemented in [6]. A pixel is activated when any of four logical conditions holds on the top-of-atmosphere reflectances ρ_{8a}, ρ_{11} and ρ_{12}. The first two conditions are

$$\alpha = (\rho_{12}/\rho_{11} \geq 1.4) \wedge (\rho_{12}/\rho_{8a} \geq 1.2) \wedge (\rho_{12} \geq 0.15) \quad (5)$$

$$\beta = (\rho_{11}/\rho_{8a} \geq 2.0) \wedge (\rho_{11} \geq 0.5) \wedge (\rho_{12} \geq 0.5) \quad (6)$$

and the activation is the disjunction

$$p = \alpha \vee \beta \vee S \vee \gamma \quad (7)$$

where S is a saturation condition and gamma a condition evaluated on the surrounding of alpha or beta. Connected components smaller than nine pixels are discarded, as specified in [7], to suppress isolated false activations.

L. Network and Training Protocol

ROMEO-Net is deliberately small. Three 3 x 3 convolutional blocks of 16, 32 and 64 channels, each followed by a rectified linear unit and 2 x 2 max pooling, are succeeded by global average pooling and a two-way linear classifier, giving 23,714 parameters in total. The configuration is listed in Table III. The architecture follows the lightweight design principle of [19] rather than introducing a novel topology, because the contribution claimed here concerns the pipeline rather than the network.

Training patches are cut after the full pre-processing chain, so that the model is fitted to exactly the data the deployed system presents to it. Following the patch decomposition of [6], 32 x 32 patches are taken at 25 % overlap and a patch is labelled as containing an event when its annotation exceeds five pixels, yielding 1,080 patches of which 168 (15.6 %) are positive. The resulting class imbalance is handled by oversampling the event class, after which the data are partitioned into 1,124 training and 124 validation patches. Optimisation uses Adam at a learning rate of 2 x 10^-3 for 30 epochs with a fixed seed, which requires 18.6 s on a general-purpose processor.

The fitted network is exported to ONNX and quantised to 8-bit integers by post-training static quantisation in the quantise-dequantise format [13], calibrated on 256 patches drawn from the same pre-processed distribution. Quantisation is reported as a measured trade-off in Section IV-F rather than assumed to be free.

M. Evaluation Protocol

Performance is reported at two levels. At patch level a patch is scored against its annotation. At granule level a granule is declared to contain an event when at least one detection survives Section III-F, and is scored against the class of the event it represents; this is the decision the spacecraft actually takes, since it determines whether a downlink occurs. Writing TP, FP, FN and TN for the four outcomes,

$$P = TP/(TP+FP), \quad R = TP/(TP+FN), \quad F1 = 2PR/(P+R) \quad (8)$$

and accuracy as the proportion of granules correctly classified.

Two protocols are used. The first partitions patches at random, which is the conventional choice and is reported for comparability. The second, leave-granules-out, excludes four whole granules from training and evaluates only on those scenes; one of the excluded granules is a snowfield, a scene type that appears nowhere in the training set. The second protocol is the one on which the generalisation claim of this paper rests, because patches drawn from a scene the model has already seen are not independent of it.

TABLE II. Reference granule catalogue and patch summary

Property	Value
Granules in catalogue	18
Event granules (eruption / fire)	10 (6 / 4)
Event-free granules	8
Granule size	256 x 256 px (5.12

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

	km x 5.12 km)
Ground sampling distance	20 m
Bands retained	B8A (865 nm), B11 (1610 nm), B12 (2190 nm)
Quantisation of digital numbers	12 bit (0-4095)
Patches extracted (overlap 25 %)	1,080
Positive patches	168 (15.6 %)
Training / validation patches	1,124 / 124
Leave-granules-out partition	840 training / 240 unseen patches
Source event database	THRawS: 88 events (58 eruption, 20 fire, 10 event-free)

TABLE III. ROMEO-Net configuration and training settings

Component	**Configuration**
Input	32 x 32 x 3 patch, raw digital numbers scaled to [0, 1]
Convolutional block 1	3 x 3, 16 channels, ReLU, 2 x 2 max pool
Convolutional block 2	3 x 3, 32 channels, ReLU, 2 x 2 max pool
Convolutional block 3	3 x 3, 64 channels, ReLU, 2 x 2 max pool
Head	Global average pooling, linear 64 -> 2, softmax
Parameters	23,714
Optimiser	Adam, learning rate 2 x 10^-3, 30 epochs, batch 64
Deployed form	ONNX, static INT8 (QDQ), 256 calibration patches
Inference window	32 x 32, stride 16, detection threshold 0.60

IV. RESULTS AND DISCUSSION

A. Granule-Level Detection Performance

Table IV reports the outcome on all 18 granules of the catalogue. Every event granule produced at least one surviving detection and every event-free granule produced none, giving 10 true positives, 8 true negatives and no false detections, hence precision 1.000, recall 1.000 and granule accuracy 100 %. The single granule rejected by the cloud gate was an overcast ocean scene, which is the intended behaviour: it is event-free, and rejecting it before inference is a correct outcome reached without spending inference effort.

A result of this form requires qualification rather than celebration, and two qualifications apply. First, the catalogue contains 18 granules, so a single misclassification would move the accuracy by 5.6 percentage points; the figure is therefore a demonstration that the pipeline is correct end to end, not a benchmark comparable with values reported on large archives. Second, the reference granules are synthetic, and Section IV-M sets out what that excludes. The threshold sweep reinforces the point: granule accuracy, precision and recall remain at 1.000 across every detection threshold from 0.30 to 0.90 in steps of 0.05, which shows that the deployed operating point of 0.60 is not delicately tuned, but equally that this catalogue does not discriminate between operating points and so cannot be used to select one.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

TABLE IV. Detection performance on the 18-granule catalogue

Quantity	Value
Granules evaluated	18
True positives / true negatives	10 / 8
False positives / false negatives	0 / 0
Precision	1.000
Recall	1.000
F1 score	1.000
Granule-level accuracy	100.00 %
Granules rejected by the cloud gate	1 (overcast ocean, event-free)
Accuracy without coregistration (ablation)	83.33 %
Unseen patches correct (leave-granules-out)	240 / 240
Unseen granules correct (leave-granules-out)	4 / 4
Threshold sweep, 0.30-0.90	precision, recall and F1 constant at 1.000

B. Effect of the Pre-processing Stage

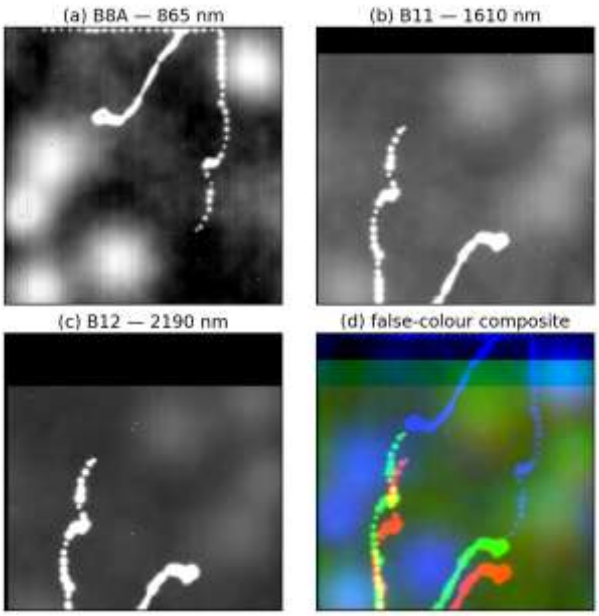


Fig. 2. Raw Level-0 granule of an eruption event: (a)-(c) the three retained bands, (d) false-colour composite with R = B12, G = B11, B = B8A. The hotspot appears as three separated colour trails because the bands are not mutually registered.

Fig. 2 shows the three retained bands of an eruption granule together with their false-colour composite. The hotspot is present in all three bands but displaced between them, and in the composite it therefore appears not as one feature but as three separated colour trails.

Fig. 3 shows the same granule after each pre-processing operation, all three panels cropped to the common valid region so that the comparison is exact. Radiometric calibration removes the striping and the impulsive corruption, visible as the disappearance of isolated bright and dark pixels between panels (a) and (b), but it does not move anything: the colour trails remain separated, because calibration is radiometric and the defect is geometric. After coarse spatial

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

coregistration in panel (c) the three trails coincide and the hotspot appears as a single saturated feature. This is the visual counterpart of the ablation reported in Section IV-E.

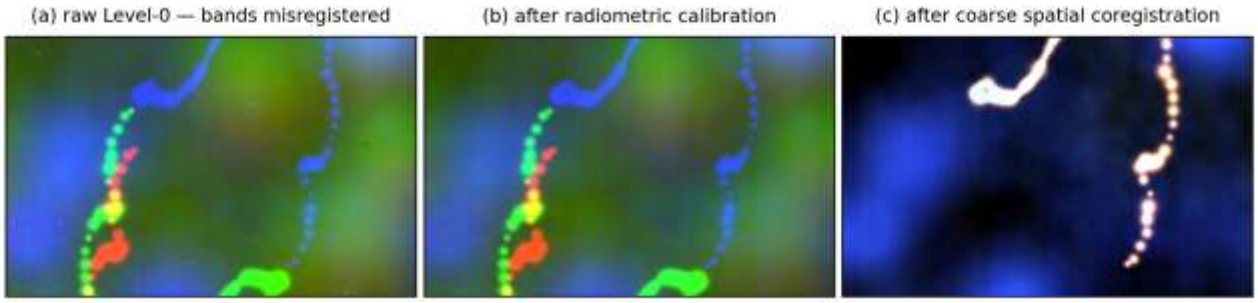


Fig. 3. Effect of the pre-processing stage on the granule of Fig. 2, all panels cropped to the common valid region: (a) raw Level-0, (b) after radiometric calibration, (c) after coarse spatial coregistration. Calibration removes striping and impulsive corruption but not the displacement; only after coregistration do the three bands describe the same ground location.

C. Detection Stage

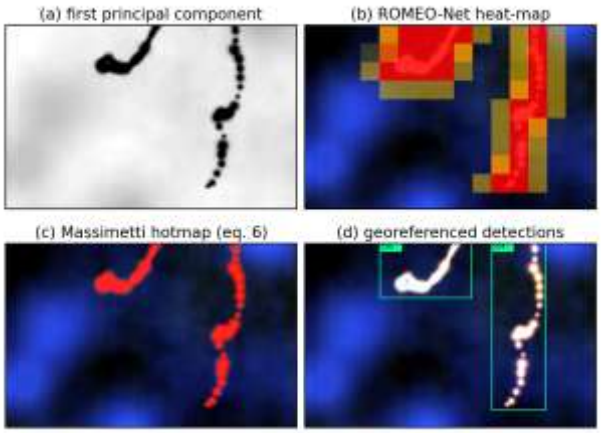


Fig. 4. Detection stage: (a) first principal component of the coregistered cube, (b) ROMEO-Net probability heat-map, (c) Massimetti reference hotmap of (7), (d) surviving components as georeferenced bounding boxes with confidence.

Fig. 4 shows the outputs of the detection stage. The first principal component in panel (a) is dominated by the short-wave infrared bands, as expected over a thermal scene. The network probability map in panel (b) and the Massimetti hotmap in panel (c) agree on the location of the anomaly, which is the expected outcome given that the annotations on which the network was trained derive from the same physical signature the threshold scheme exploits. The two differ in character rather than in conclusion: the threshold scheme activates individual pixels and is therefore sharp but fragmented, whereas the network assigns a graded probability over a neighbourhood and is therefore smoother and more contiguous. Panel (d) shows the surviving components reduced to georeferenced bounding boxes. For this granule the pipeline reported two detections at 97.8 % and 94.3 % confidence covering 1.434 km^2 each.

D. Coregistration Accuracy

Residual misregistration after the coregistration stage was measured by sub-pixel phase correlation of each band against the master band B8A. Over the catalogue, 30 estimates met the reliability criterion, the remainder being rejected because a low-texture scene does not present the gradient structure the estimator requires - a limitation shared with the keypoint matchers the coarse scheme replaces. Across the reliable estimates the mean absolute residual is 0.150 px along-track and 0.243 px across-track, with standard deviations of 0.293 px and 0.472 px respectively. Both standard deviations are below one pixel, which is the criterion stated in [6] and adopted as a requirement at the concept stage, so the alignment is sub-pixel in the sense required and the residual is smaller than the smallest anomaly the detector is asked to find.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

E. Ablation of the Coregistration Stage

The contribution of coarse spatial coregistration was isolated by evaluating the same quantised network twice over the catalogue: once on coregistered granules and once on calibrated but unregistered granules, with no retraining and no other change. Granule accuracy falls from 100 % to 83.33 %, shown in Fig. 5(c). Three granules are misclassified without alignment.

The mechanism is the one Fig. 2 makes visible. The network decides from a 32 x 32 window, and the discriminative evidence for a hotspot is the relation between the three bands at the same ground location: B12 strongly exceeding B11, which in turn exceeds B8A. When the bands are displaced by tens of pixels, a window no longer contains the three measurements of one place but three measurements of three different places, and the spectral relation the detector depends upon is destroyed. Coregistration is therefore not an enhancement applied to improve an already working detector; it is the step that makes the evidence available at all, which is the central claim of [6] rendered as a measurement on this system.

F. Cost of Integer Quantisation

Quantisation was assessed in the three dimensions that matter on board, with results in Fig. 5(a) and 5(b). Model size falls from 93.6 KB in single precision to 29.9 KB in 8-bit integer form, a factor of 3.1. Mean inference time per granule falls from 70.25 ms to 8.83 ms, a factor of 8.0; the comparison is between the integer runtime and the reference floating-point implementation over identical windows, so it reflects both the arithmetic and the runtime. Agreement between the two models, measured as the proportion of pixels receiving the same binary decision, is 96.19 %. No granule-level decision changed between the two models over the catalogue, so the 4 % of pixels that differ lie at the margins of detected components and do not alter any reported detection. Quantisation is accordingly accepted here as a favourable trade: a threefold reduction in footprint and an eightfold reduction in latency for a pixel-level disagreement that does not propagate to any decision.

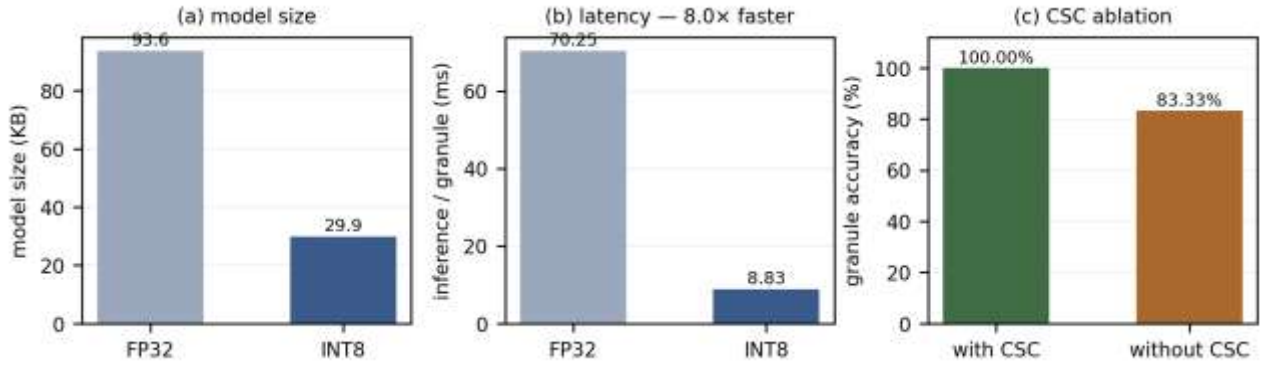


Fig. 5. Measured trade-offs: (a) model size and (b) mean inference time per granule for the single-precision and 8-bit integer models, and (c) granule accuracy with and without the coarse spatial coregistration stage, the model being identical in both cases.

G. Generalisation Under Leave-Granules-Out Evaluation

Under the random patch partition the network reaches 100 % validation accuracy on 124 patches. That figure is reported for comparability only, because patches drawn at random from the same granules are not independent: a validation patch may overlap, or be adjacent to, a training patch from the same scene.

The leave-granules-out protocol removes this objection. A separate model was trained on 840 patches with four whole granules excluded - an eruption, a wildfire, and two event-free scenes, one of which is a snowfield whose spectral behaviour, bright in the near infrared and dark in the short-wave infrared, appears nowhere in the training set. Evaluated only on the excluded scenes, the model classified 240 of 240 unseen patches correctly (19 true positives, 221 true negatives, no false detections) and reached the correct decision on 4 of 4 unseen granules. That the snowfield is correctly rejected is the more informative half of this result, since it shows the network discriminating on the spectral relation rather than on brightness alone.

H. Robustness to Single-Event Upsets

Radiation-induced corruption is the failure mode identified as dominant in [1], and the pipeline was exercised against it by injecting additional salt-and-pepper corruption at rates from 0.2 % to 6 % of pixels per band, each configuration

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

being evaluated twice: with the impulse-repair step of the calibration stage enabled and with it disabled. Fig. 6 reports the result.

With repair enabled, granule accuracy remains at 100 % at every injected rate up to 6 %. With repair disabled, accuracy falls to 66.67 % as soon as corruption reaches 0.2 % and does not recover. The gap between the two curves is the measured contribution of a single inexpensive operation within the radiometric stage, and it is large: a median filter applied only where a pixel deviates from its neighbourhood is sufficient to absorb two orders of magnitude of injected upset rate. The result also indicates where the fragility lies, which is not in the network but in the detection arithmetic downstream of it: isolated saturated pixels satisfy the B12-dominance relation locally and are promoted into spurious components unless they are removed first.

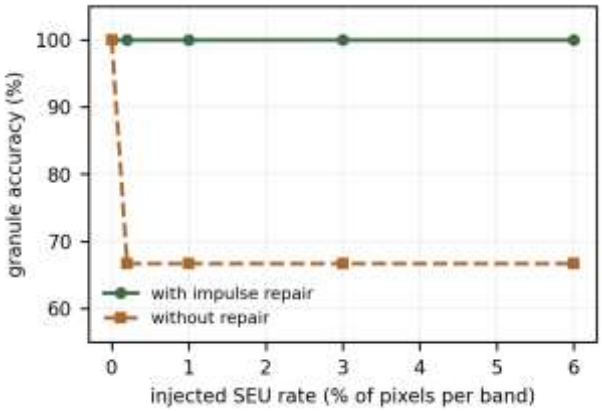


Fig. 6. Granule accuracy under injected single-event-upset corruption, with and without the impulse-repair operation of the radiometric stage. Accuracy is held at 100 % up to a 6 % corruption rate when repair is enabled.

I. Computational Cost

Stage timings were measured over 20 repetitions on one granule on a general-purpose processor without acceleration, and are reported in Table V and Fig. 7. Coarse spatial coregistration requires 0.128 ms, cloud screening 0.084 ms, the quantised network 5.018 ms and the reference threshold detector 1.063 ms. Radiometric calibration dominates at 23.944 ms, because the median comparison used for impulse repair is the only operation in the chain that inspects a neighbourhood per pixel.

The coregistration figure admits direct comparison with the literature. The granules used here are 256 x 256 pixels whereas [6] reports timings for a full 1152 x 1296 granule, a factor of 22.8 in area; scaling the measured 0.128 ms accordingly gives 2.92 ms, against the 16.63 ms reported in [6] for the same operation on a central processing unit. The comparison should be read as an order-of-magnitude agreement rather than an improvement, since the processors differ and the present implementation handles three bands rather than thirteen; what it establishes is that the alignment stage remains in the millisecond regime, which is the property that makes it flyable.

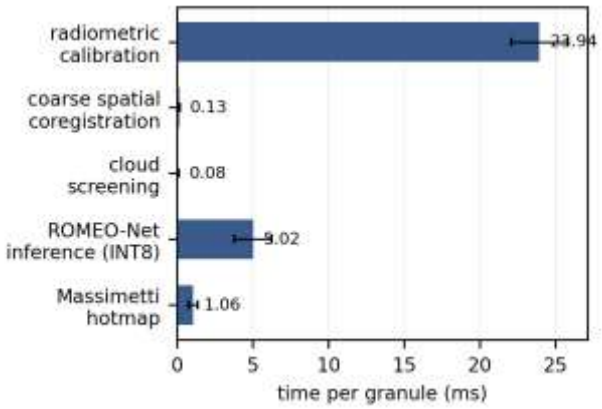


Fig. 7. Mean time per granule for each pipeline stage over 20 repetitions on a general-purpose processor; error bars are one standard deviation. Radiometric calibration dominates because impulse repair inspects a neighbourhood per pixel.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

TABLE V. Stage timings and deployment diagnostics

Diagnostic	Value
Radiometric calibration	23.944 +/- 1.876 ms
Coarse spatial coregistration	0.128 +/- 0.039 ms
Cloud screening	0.084 +/- 0.012 ms
ROMEO-Net inference (INT8)	5.018 +/- 1.221 ms
Massimetti reference detector	1.063 +/- 0.305 ms
CSC scaled to 1152 x 1296 granule	2.92 ms (16.63 ms reported in [6])
Model size, FP32 / INT8	93.6 KB / 29.9 KB
Mean inference per granule, FP32 / INT8	70.25 ms / 8.83 ms
Pixel-decision agreement, INT8 vs FP32	96.19 %
Residual misregistration, s.d. along / across	0.293 px / 0.472 px
Raw catalogue / alert volume	30.67 MB / 105.74 KB (290x)

J. Downlink Reduction

Over the whole catalogue the raw granules amount to 30.67 MB, whereas the alerts generated from them amount to 105.74 KB, a reduction by a factor of 290. Individual alerts range from 9.70 KB to 11.00 KB with a mean of 10.57 KB, the variation arising from the number of detections and the cropped thumbnail each alert carries; against the 1.70 MB of a single raw granule this is a per-granule reduction of 160. Event-free granules contribute nothing to the downlink at all, which is where the larger part of the saving originates: the system transmits for 10 of 18 granules rather than for all 18. At the S-band rate of 1 Mbit/s the raw catalogue requires approximately 245 s of contact time against 0.85 s for the alerts.

K. Coregistration as the Enabling Step

The most informative single result in this study is the ablation of Section IV-E, because it concerns a design decision rather than a performance figure. Coarse spatial coregistration is the cheapest operation in the chain at 0.128 ms, two orders of magnitude below the radiometric stage and a fortieth of the network, yet removing it costs 16.67 percentage points of granule accuracy. Cost and consequence are therefore inversely related, and the practical implication for a mission designer is that the alignment stage is the last thing that should be removed under a compute budget, not the first.

The reason is specific to raw data and would not arise on a Level-1C product, where registration has already been performed on the ground. This is what justifies treating raw-data processing as a distinct problem rather than as a minor variation of established onboard inference: the operation that dominates the result is one that does not exist in the conventional pipeline, and a study conducted on corrected products would have no occasion to discover it.

L. Status of the Objectives Stated at the Concept Stage

The concept-stage work [18] specified a six-stage pipeline together with functional and non-functional requirements. Since that specification is the standard against which the present implementation should be judged, the status of each requirement is stated in Table VI, and the assessment is deliberately conservative.

Six requirements are met. The six processing stages are implemented and execute without any ground-segment product. Coarse spatial coregistration is implemented by the look-up-table method of [6] and its residual is measured below one pixel, which satisfies the sub-pixel requirement. The latency requirement is met with margin: the alignment stage scales to 2.92 ms against the 16.63 ms reference. The lightweight-model requirement is met, the deployed network occupying 29.9 KB. Event procurement from a reference database is implemented against the 88-event THRawS catalogue, and raw granule preparation and export are implemented.

One requirement is partially met. The system is evaluated on synthesised granules carrying the identities of real events rather than on the real archive; the import path for genuine THRawS granules is implemented and tested, so what remains is the acquisition of the data, not the construction of a capability.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

Two requirements remain open and should not be represented otherwise. Execution on representative flight hardware has not been attempted: all timings in this paper are from a general-purpose processor, and the migration cost reported in [4] means that they cannot be assumed to transfer to a vision processing unit or an integer accelerator. Pixel-level segmentation of the anomaly, as opposed to detection and bounding, is not implemented; it would require the segmentation annotations of a dataset such as SegTHRawS rather than a change of method.

TABLE VI. Status of the requirements stated at the concept stage [18]

ID	Requirement from [18]	Status	Evidence or obstacle
R1	Six-stage onboard pipeline on Level-0 data	Met	All six stages implemented; no ground-segment product used (Section III)
R2	Coarse spatial coregistration of raw bands	Met	Look-up-table method of [6]; ablation in Section IV-E
R3	Sub-pixel registration accuracy (s.d. < 1 px)	Met	0.293 px along-track, 0.472 px across-track (Section IV-D)
R4	Lightweight model within onboard memory	Met	23,714 parameters, 29.9 KB after INT8 quantisation
R5	Coregistration latency comparable to [6]	Met	2.92 ms scaled to a full granule against 16.63 ms
R6	Event procurement and raw data preparation	Met	88-event THRawS catalogue; granule export implemented
R7	Evaluation on real THRawS granules	Partial	Import path implemented and tested; evaluation performed on synthesised granules
R8	Execution on representative flight hardware	Open	All timings from a general-purpose processor; migration cost of [4] not assessed
R9	Pixel-level segmentation of anomalies	Open	Detection and bounding only; requires segmentation annotations

M. Limitations and Threats to Validity

The principal limitation of this study is the nature of its reference data, and it constrains every accuracy figure reported above. The granules were synthesised from the sensor model of Section III-I, so the detection task was to recover a known signature from a known corruption process. A real granule would additionally contain atmospheric scattering and absorption that vary across the scene, thin cirrus that neither the cloud gate nor the detector was designed against, sun-glint over water, terrain shadow, sensor degradation accumulated in orbit, and thermal anomalies of anthropogenic origin - gas flares, industrial furnaces and burning landfill - which present the same spectral signature as the events of interest and constitute the dominant false-positive population in operational use. None of these is represented. The accuracy of 100 % therefore quantifies the capacity of the model and the correctness of the pipeline; it is not a prediction of field performance and should not be compared directly with values reported on measured archives. The authors regard restating this plainly as preferable to allowing a favourable number to stand without qualification.

Four further limitations follow. First, the catalogue comprises 18 granules of 5.12 km extent, whereas an operational granule is 1152 x 1296 pixels; results on a small catalogue of small extracts carry correspondingly wide uncertainty, and the flat threshold sweep of Section IV-A is direct evidence that the catalogue does not discriminate finely enough to select an operating point. Second, the comparison with the classical detector of [7] is not an independent validation, because the annotations used to train the network derive from the same spectral signature that detector exploits; agreement between them confirms consistency, not correctness. Third, all timings are from a general-purpose processor, and no claim is made about flight hardware. Fourth, the residual-misregistration estimate of Section IV-D could be formed for only 30 of the 36 band pairs, the remainder being low-texture scenes in which the estimator is unreliable, so the reported standard deviations characterise the textured subset rather than the whole catalogue.

V. CONCLUSION

The six-stage onboard pipeline previously proposed by the authors at the concept stage has been implemented as a working system and evaluated quantitatively. Raw Sentinel-2 Level-0 granules are ingested without ground correction, radiometrically calibrated, coregistered and coarsely georeferenced on board, screened for cloud, reduced to three spectral bands, analysed by a 23,714-parameter network quantised to 29.9 KB, and converted into georeferenced alerts of approximately 10 KB that are transmitted in place of the imagery. On an 18-granule catalogue the system reaches

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IJARCCE.2026.1510125

granule-level accuracy of 100 % with no false detections, and under leave-granules-out evaluation it classifies 240 of 240 unseen patches and 4 of 4 unseen granules correctly, including a scene type absent from training.

Three findings are of wider interest than the particular numbers. Coarse spatial coregistration is the enabling step rather than an optimisation: it is the cheapest operation in the chain at 0.128 ms, yet removing it while holding the model fixed costs 16.67 percentage points of accuracy, because a misaligned window no longer presents the inter-band relation on which the detection depends. Integer quantisation is a favourable trade in this regime, reducing the model threefold and the latency eightfold while changing no granule-level decision. And the inexpensive impulse-repair operation inside the radiometric stage, rather than the network, is what confers tolerance to single-event upsets, holding accuracy at 100 % up to a 6 % corruption rate where its removal costs a third of the accuracy at a rate of 0.2 %.

Future work is determined by Table VI and by the limitations of Section IV-M, in the following order of priority. The immediate requirement is to replace the synthesised catalogue with real THRawS granules [9], for which the import path already exists, and to re-evaluate every result reported here; until that is done no accuracy figure in this paper transfers to the field, and the anthropogenic thermal sources that dominate operational false positives remain entirely untested. The second requirement is execution on representative flight hardware, which would convert the latency figures from indicative to meaningful and would expose the migration cost reported in [4]. The third is pixel-level segmentation in place of bounding, for which a segmentation-annotated dataset is the prerequisite. Thereafter, evaluation against thin cirrus and sun-glint, and a fault-injection study at the level of model weights rather than input pixels, would address the remaining aspects of the radiation environment described in [1].

REFERENCES

[1]. G. Furano, G. Meoni, A. Dunne, D. Moloney, V. Ferlet-Cavrois, A. Tavoularis, J. Byrne, L. Buckley, M. Psarakis, K.-O. Voss, and L. Fanucci, “Towards the use of artificial intelligence on the edge in space systems: Challenges and opportunities,” IEEE Aerospace and Electronic Systems Magazine, vol. 35, no. 12, pp. 44–56, Dec. 2020.

[2]. G. Giuffrida, L. Fanucci, G. Meoni, M. Batic, L. Buckley, A. Dunne, C. van Dijk, M. Esposito, J. Hefele, N. Vercruyssen, G. Furano, M. Pastena, and J. Aschbacher, “The Φ-sat-1 mission: The first on-board deep neural network demonstrator for satellite Earth observation,” IEEE Transactions on Geoscience and Remote Sensing, vol. 60, 2022, Art. no. 5517414.

[3]. G. Guerrisi, F. Del Frate, and G. Schiavon, “Convolutional autoencoder algorithm for on-board image compression,” in Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2022, pp. 151–154.

[4]. G. Guerrisi, F. Del Frate, and G. Schiavon, “Artificial intelligence based on-board image compression for the Φ-sat-2 mission,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 8063–8075, 2023.

[5]. P. Di Stasio, A. Sebastianelli, G. Meoni, and S. L. Ullo, “Early detection of volcanic eruption through artificial intelligence on board,” in Proc. IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 2022, pp. 714–718.

[6]. G. Meoni, R. Del Prete, F. Serva, A. De Beusscher, O. Colin, and N. Longépé, “Unlocking the use of raw multispectral Earth observation imagery for onboard artificial intelligence,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 12521–12537, 2024.

[7]. F. Massimetti, D. Coppola, M. Laiolo, S. Valade, C. Cigolini, and M. Ripepe, “Volcanic hot-spot detection using SENTINEL-2: A comparison with MODIS–MIROVA thermal data series,” Remote Sensing, vol. 12, no. 5, p. 820, 2020.

[8]. ESA Φ-lab, “PyRawS: Python for raw Sentinel-2 data,” software repository, 2023.

[9]. G. Meoni, R. Del Prete, F. Serva, A. De Beusscher, O. Colin, and N. Longépé, “THRawS: Thermal hotspots in raw Sentinel-2 data,” dataset, Zenodo, 2023.

[10]. European Space Agency, “Sentinel-2 user handbook,” ESA Standard Document, rev. 2, 2015.

[11]. P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperGlue: Learning feature matching with graph neural networks,” in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 4938–4947.

[12]. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, “scikit-image: Image processing in Python,” PeerJ, vol. 2, p. e453, 2014.

[13]. B. Jacob, S. Kligys, B. Chen, M. Zhu, M. Tang, A. Howard, H. Adam, and D. Kalenichenko, “Quantization and training of neural networks for efficient integer-arithmetic-only inference,” in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 2704–2713.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510125

[14]. ONNX Runtime developers, "ONNX Runtime: Cross-platform accelerated machine learning," software, 2021.
[15]. J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, "JAX: Composable transformations of Python+NumPy programs," software, 2018.
[16]. Smithsonian Institution, "Global Volcanism Program: Volcanoes of the world database," National Museum of Natural History, Washington, DC, USA, 2023.
[17]. Copernicus Emergency Management Service, "Rapid mapping and risk and recovery mapping portfolio," European Commission, Joint Research Centre, 2023.
[18]. A. G, C. M. H, H. K. P, and Mallikarjuna, "ROMEO: Raw onboard multispectral Earth observation," Project Phase-1 Report, Department of Computer Science and Engineering, East West College of Engineering, Visvesvaraya Technological University, Bengaluru, India, 2025–26.
[19]. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.

 This work is licensed under a Creative Commons Attribution 4.0 International License